Data-Driven Modelling of Routing and Scheduling in Freight Transport

Nadi Najafabadi, A.

DOI
10.4233/uuid:6d71f79a-5cb5-4b64-a256-2906dbbedff6

Publication date
2022

Document Version
Final published version

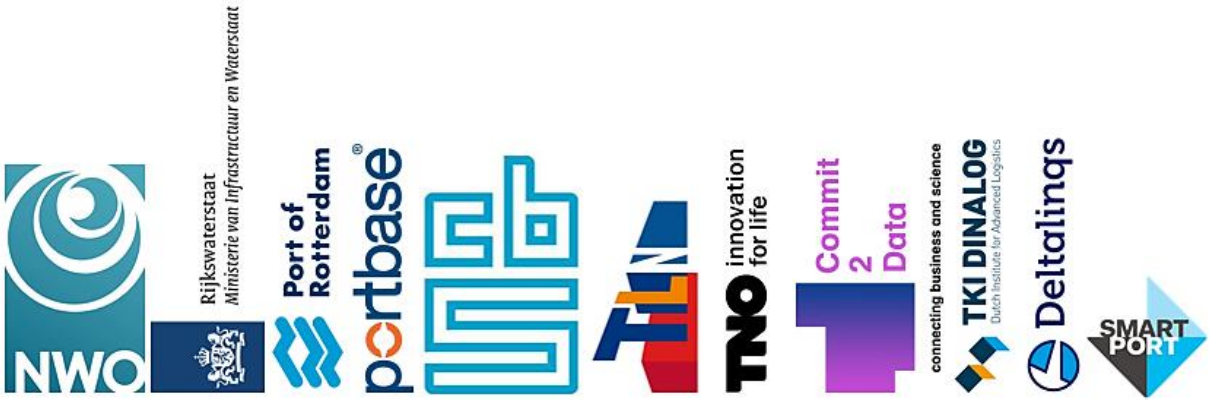
Citation (APA)
Nadi Najafabadi, A. (2022). *Data-Driven Modelling of Routing and Scheduling in Freight Transport*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:6d71f79a-5cb5-4b64-a256-2906dbbedff6

Data-Driven Modelling of Routing and Scheduling in Freight Transport

Ali Nadi Najafabadi

Delft University of Technology

This doctoral dissertation was funded by the Dutch Research Council (NWO), TKI Dinalog, Commit2Data, Port of Rotterdam, SmartPort, Portbase, TLN, Deltalinqs, Rijkswaterstaat, and TNO under project ‘ToGRIP-Grip on Freight Trips’ with grant number 628.009.001.



Cover illustration: Getty Images, rudzhan Nagiev, bestbrk

Data-Driven Modelling of Routing and Scheduling in Freight Transport

Dissertation

For the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates,
to be defended publicly on
Monday 17 October 2022 at 12:30 o'clock

by

Ali NADI NAJAFABADI
Master of Science in Civil Engineering – Road and Transportation,
K.N. Toosi University of Technology, Iran
born in Esfahan, Iran

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof.dr.ir. L.A. Tavasszy	Delft University of Technology, promotor
Prof.dr.ir. J.W.C van Lint	Delft University of Technology, promotor
Dr. M. Snelder	Delft University of Technology, copromotor

Independent members:

Prof.dr.ir. J. Holguin-Veras	Rensselaer Polytechnic Institute, USA
Prof.dr.ir. M.J. Roorda	University of Toronto, Canada
Dr. E.I. Vlahogianni	National Technical University of Athens, Greece
Prof.dr.ir. S.P. Hoogendoorn	Delft University of Technology
Prof.dr.ir. B.H.K. De Schutter	Delft University of Technology, reserve member

TRAIL Thesis Series no. T2022/14, the Netherlands Research School TRAIL

TRAIL
P.O. Box 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-317-6

Printed in the Netherlands

To all who have never stopped trying!

Acknowledgments

I'm now at the final stage of obtaining a Doctor of Philosophy (Ph.D.). I believe this journey has commenced years ago when I was a kid with an infinite passion to learn, a vivid desire to understand, and an aspiration deep inside me that never runs out. Obtaining the Ph.D. is not the end of this adventure but is a definite milestone in my life. In this book, you can see some contributions to my field, if you read the lines, and many contributions to my life, if you read between the lines. For all that, I would like to take this opportunity to thank all the contributors.

With Prof.dr.ir. Lóránt (Lori) Tavasszy , Prof.dr.ir. Hans Van Lint, and Dr. Maaike Snelder, I was extremely fortunate to have the most complementary supervisory team. I would like to begin by expressing my deepest gratitude to Lori who has brought in his abundance of knowledge in Freight transport and logistics. I met Lori four years ago in his office when I was just about to start my research. Since then, we have had ample time for discussions every single of which was full of inspiration. Lori, You have always intrigued me with your thoughtful ideas and have been a continuous source of encouragement for me. All the great things that I have learned from you will remain with me forever and I will try my best to hand them over to others after me. I cannot be more thankful for your support, and all the opportunities that you gave me to grow over the course of my Ph.D. My sincerest appreciation, then, goes to Hans who has brought in his invaluable knowledge in the traffic domain. He enlightened my path with his brilliant mind and kept me moving with his enthusiasm. Hans, I have always enjoyed and learned a lot from our conversations. Thank you for believing in me and supporting me from the very beginning to the end. Last but not least, I owe a debt of heartfelt gratitude to Maaike who has brought in her expertise in transportation research. Maaike, you make all problems look so easy and were there to support me both scientifically and mentally. I could not end this journey without you. Thank you for being a great support. In all, I want to thank the three of you for all the professional support that you have provided me with throughout these years and for all the memorable moments that we have made together.

I would like to thank Prof. Niel Yorke-smith who has supported and guided me during my research with his invaluable knowledge of AI and optimization. Our collaboration in the Amazon challenge was one of the highlights and a truly unforgettable experience during my Ph.D.

I would like to sincerely appreciate my doctoral committee members, Prof. Jose Holguin-Veras, Prof. Roorda, Prof. Vlahogianni, and Prof, Hoogendoorn, for all the thoughtful reviews and constructive remarks that helped me to improve the quality of my dissertation and for all your fantastic scientific tracks that inspired me during my research.

There is nothing that can make a working environment professional, productive, and enjoyable than enthusiastic colleagues who you work with. It has been my utmost pleasure to work with such amazing teams in Freight and logistics Lab and Dittlab - Salil, Raeed, Xiao, Rodrigo, Michiel, Shahrzad, Merve, Zhenjie, Tin, Zahra, and Panchamy. Special thanks go to Salil and Raeed for all the great collaborations we had. We started our journey together and had a lot of brainstorming, mind mapping, and coding sessions but other than that we also had a lot of fun and memorable moments. I always admire you guys and learned a lot from you. I have to thank my dear friend Peyman for all the off-topic discussions on almost everything during breaks around the corners (plane seats). You are one of the few people I know that can/would listen to my philosophical jibber-jabbers with enthusiasm. You should let me know if you hate it, otherwise, I won't stop. Many thanks to my best friends in Iran- Mehdi, Ruholla, Hessam, and Nahid for cheering me on and supporting me from the other side of the world.

I would like to thank my sisters Shiva and Neda and my brothers Navid and Saeed and my brother in laws Amir and Ayoub. You always have been my greatest source of energy. Thank you for all the love and support. Special thanks go to Saeed. You have always had a strong belief in me and you saw potential in me when I didn't. You taught me coding at the age of 13, which changed my life, and guided me in all the important decisions states. Where I'm standing now is definitely a result of your guidance and support. Thank you for all of that.

There is no word that can sufficiently convey my gratitude towards the greatest people in my life – my love (Razieh) and my parents (baba jan Mostafa and maman Behjat). I want to thank my father for giving me confidence, responsibility, and resistance in work and for teaching me how to stand on my own. You are my role model in forgiveness, munificence, and diligence. I wish I can represent your personality as a good man. I want to thank my mother for all the caring, praying, and patience. You always wished I could remain a sweet kid playing around but at the same time completely dedicated yourself to supporting me grow. I want to tell you that you should not be worried, I'm still a sweet kid inside and will remain as such. And Razieh, nothing can explain the chemistry between us. You have brought love into my life and your love has opened many doors for me. You and I together planned this journey 17 years ago and you have always been a part of all the tears and fears and enjoyments on this path so this is as much your achievement as it is mine.

Finally, I would like to thank the one inside me who has never let me give up in difficult situations, the one who has always whispered in my head that I can do it!

Ali Nadi,
Delft, September 2022.

Contents

Chapter 1

Introduction

1.1 Research motivation

Logistic activities and freight transport form the backbone of the trade in goods within the international economy. The performance of traffic networks determines the waiting times, handling times and travel times in the network and thus has a significant economic impact (Tavasszy and Reis, 2021, De Jong et al., 2014). There is a lack of research, however, regarding how logistics and freight transport activities interact with a highly dynamic traffic system, where travel times are continuously changing. This research aims to unravel this interaction by considering the spatial and temporal dynamics of both systems.

This doctoral research is part of the ToGRIP-Grip on Freight Trips project. The project is funded by the Dutch Research Council (Nederlandse organisatie voor Wetenschappelijk Onderzoek or NWO)[1] with as overall objective: "To develop a data-driven integrated traffic and logistics model that can be used to design interventions (executed by the Port, Road, and other authorities) to combat travel time unreliability and to improve logistics operations."- Snelder (2016). ToGRIP is based on two pillars: pre-trip and on-trip decisions of carriers. These two pillars are spread over four work packages: empirical analysis; integrated traffic and logistics model; interventions and knowledge utilization; and project management. This research fits in the first pillar and concerns the pre-trip decisions of carriers. The second pillar is studied by Salil Sharma as a part of his doctoral research.

This chapter provides an introduction to the research by first scoping the research area and background to identify research gaps. Then, we formulate our main research question followed by research methodology and subquestions. Next is an expression of our scientific and societal contributions and finally, we outline the structure of this book.

[1] grant number 628.009.001

1.2 Scoping and background

In a logistics system, ports play a vital role as linking pins within international and national supply chains. The port of Rotterdam as the largest port in Europe facilitates the needs of the hinterland to hundreds of millions of consumers all over Europe. Transporting containerized freight from port to hinterland and hinterland to port is the outcome of decisions of different decision-makers whose characteristics and activities change across time and space. Many of these decisions will have an impact on freight flow patterns. Subsequently, truck movements through transport networks represent these freight flow patterns. Whereas these trips can have a large impact on parts of the road network by contributing to congestion, road network performance can also affect trip-related logistics decisions (see **Figure 1-1**). The overall scope of our focus in this research will be on these decisions, in the context of freight flows, truck movement, and road network performance.

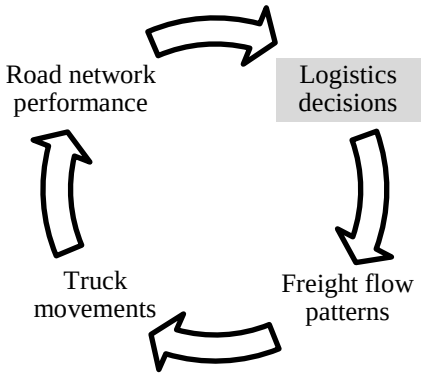


Figure 1-1:Impact cycle from logistics decisions to road network performance

Given the impact cycle represented in **Figure 1-1**, it is difficult to predict tactical-level logistics reorganization responses to changes in the performance of the road network. These reorganization responses are driven by many logistics decisions. A comprehensive framework for logistics decisions is presented by Langevin and Riopel (2005). Long-term structural decisions are related to investments physical facilities and communication networks while operational level decisions relate to demand forecasting, inventory management, production, supply management, transportation, product packaging, material handling, warehousing and order processing. Each of these decision categories consists of many sub-decisions, leading to an impressive number of 48 logistics decisions. The complexity of this set of decisions is daunting. The transportation category consists of 8 sub decisions: (1) transportation modes (2) types of carriers (3) carriers selection (4) degree of consolidation (5) transportation fleet mix (6) assignment of the customer to vehicles (7) vehicle routing and scheduling and (8) vehicle load plans.

In this thesis, we focus on one specific decision from this set. From the above, the routing and scheduling decisions (7) have the most direct and tractable impact on the contribution of freight to traffic flows. The reason is that these decisions mainly relate to the tour planning of carriers of goods who bundle several commodities into one pickup and delivery tour to maximize their capacity utilization and reduce their transport cost.Transport costs contain two key components with time-space variations. The first component relates to constraints imposed by firms in pickup and delivery locations. Examples of these logistic constraints are inventory of distribution centers, cargo handling of transshipment terminals, stock replenishment of retails, and certain characteristics of producer or consumer of goods. The time-space variation of this

cost component relates to the industrial activity of firms and hence connects the routing (space) and scheduling (time) decisions to the logistics systems. The second component relates to the travel time between pickup and delivery locations. Congestion and unreliability of travel time on road networks affect the logistics system by imposing additional transportation costs and will thereby affect tour planning. The space-time variation of travel times on road networks connects the routing and scheduling decisions (tour planning in general) to the traffic system (see **Figure 2-1**).

Traffic is also a dynamic system in which its state changes across time and space. The dynamic of the traffic mainly depends on the characteristics of the infrastructure (supply), the existence of vehicles (demand), and their behavior on the road networks. Trucks, which are a result of tour planning decisions, are part of this demand. Therefore, the tour planning decisions in time (when to depart tours) and space (where and to what order pick up or deliver a commodity) can change the percentage of trucks on road networks at a particular time of day and hence can contribute to changes in flows, speed, and travel times on road networks. **Figure 1-2** illustrates how the time and space dimensions in these two systems are tied together.

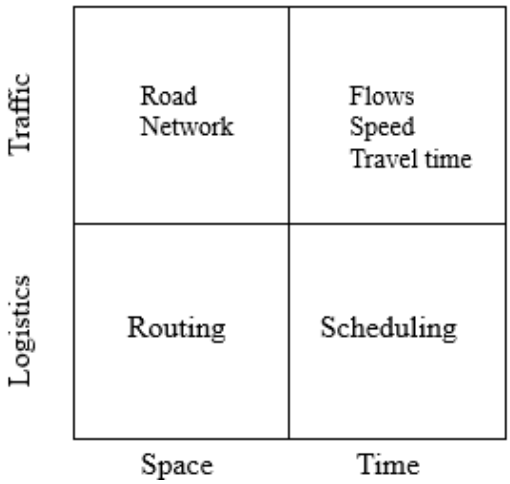


Figure 1-2: Time-space characteristics of logistics and traffic system

In sum, our general area of work concerns the interactions between logistics systems and traffic systems, which encompass key logistics decisions (routing and scheduling) that affect traffic conditions on one hand, and traffic congestion (and the resulting delays) affecting these decisions on the other (**Figure 1-3**).

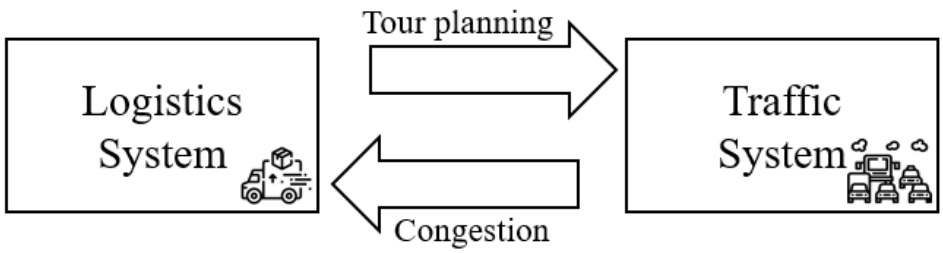


Figure 1-3:Interactions between logistics and traffic systems

Although we know that there is such a mutual relationship between logistics and traffic systems, traffic and logistics systems have mostly been studied separately. The evidence presented thus far supports the need for integrated logistics and traffic modelling to unravel the complex relationship between these two systems (see **Figure 1-4**). Examples of this evidence are the highly congested road networks, especially around port cities like Rotterdam, which would lead to a major loss in logistics systems and society. For example, according to (TLN and EvoFenedex., 2020), in 2019, congestion on Dutch motorways led to a total economic loss of 1.5 billion euros for freight transport.

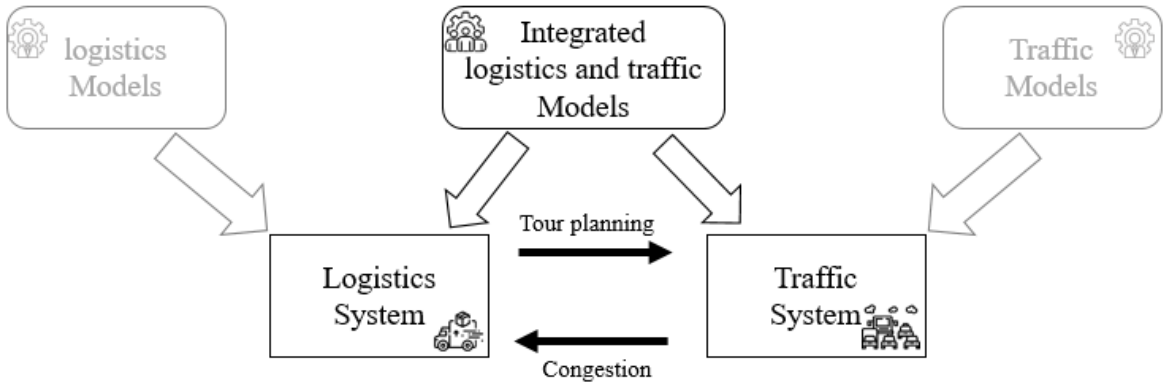


Figure 1-4: Integrated logistics and traffic models

This need for an integrative approach to logistics and traffic modelling requires new research. Freight transport modelling has followed for decades a similar traditional approach as its counterpart the passenger transport system. This approach, which is known as four-step transportation modelling, includes four sub-models that are applied sequentially. These steps are trip generation, trip distribution, modal split, and trip assignment. Previous research in the field of freight transport modelling has identified the differences between these two systems and introduced several other components to freight transport modelling like freight generation, tour-scheduling, and tour formation (De Jong et al., 2014, Tavasszy and De Jong, 2013). In 2012 Tavasszy et al. presented a review on freight demand modelling identifying research opportunities for incorporating logistics decisions in freight models. An earlier example of such incorporation was SMILE (see Tavasszy et al. (1998) which describes freight modelling in three levels production, inventory, and transportation. The framework proposed by Tavasszy (2008) also resembles the four-step framework of passenger transport but includes logistic operations and decisions which are specifically related to freight transport (e.g. storage). This framework consists of three layers: (1) exchange of goods, which is a representation of the freight economy and describes production, consumption and trading patterns; (2) a logistics behavior layer, which models the decisions that are being made to find a trade-off between transporting goods or storing them in a warehouse (inventory decisions); and (3) freight trips and transport organization. The last layer is related to those models that consider decisions about transport modes, vehicle type, shipment size, routing and scheduling decisions. **Figure 1-5** summarizes the freight transport modelling from the generation of freight to the representation of vehicles on a congested road network. Here, the transport movements or trips are the central phenomena to observe, as they connect the logistics and traffic processes. The logistics process relates to the pre-trip characteristics, while the traffic process concerns phenomena that on-trip characteristics have been established.

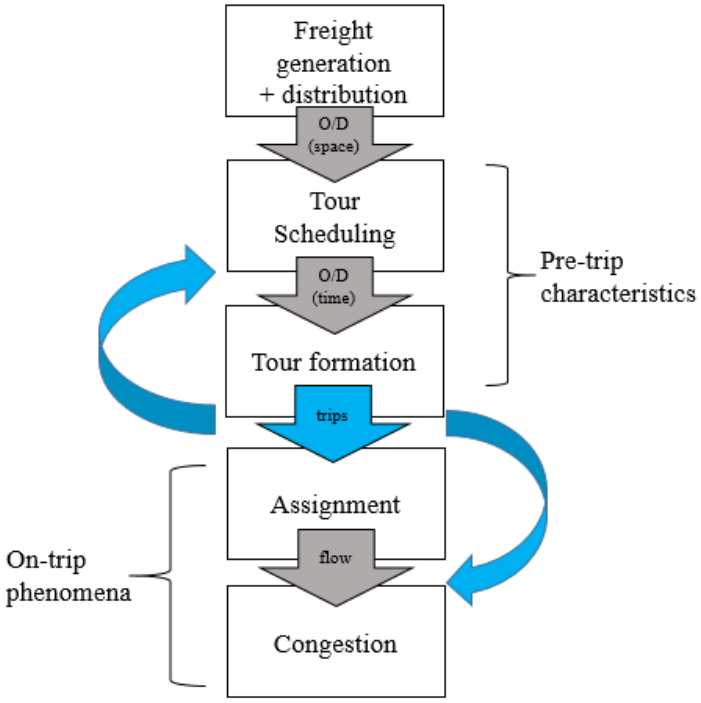


Figure 1-5: The freight modelling framework

The many trips made by trucks are not all independent as they are often structured in tours. Once these interdependencies can be observed or inferred from observations, the connection between the 2 subsystems can be made inside freight transport models. There are three major methods in the literature for modelling pre-trip tour characteristics. First is the tour growth approach, in which tours are constructed incrementally through an econometric approach by estimating conditional choices of the next destination, given previous stops (Hunt and Stefan (2007)). In this approach, trips are the unit of analysis and are considered independent. (2) Second is the commodity-based approach which is similar to the tour growth approach but the unit of analysis is commodities that are assigned to the vehicles incrementally e.g. in Thoen et al. (2020). None of these two approaches, however, deals with the interdependencies of trips and scheduling of tours accurately and thus have difficulty connecting to the after-trip phenomena tangibly. The key issue with these approaches is the behavioural assumption of the tour formation process which is not the context in which tours are often formed. Tour planning, is much rather the end result of interdependent routing and scheduling optimization process than simply the result of sequential discrete choices of decision-makers. Finally, the third approach concerns normative models that apply the vehicle routing problem (VRP) like in Donnelly (2009). Opposed to the first two approaches, VRP models consider the interdependency of trips and scheduling of tours. However, the result of these models deviates from the real world, due to the heterogeneity in objectives and constraints of transport markets. This suggests a need for data-driven approaches that can deal with understanding the pre-trip characteristics and connect well to the after-trip traffic phenomena.

Despite the general awareness of these systematic approaches, we still have a very poor understanding of the freight transport system and its interaction with the traffic system. This, to a great extent, is owing to the lack of data about this system (Holguín-Veras et al., 2014). As

shown in **Figure 1-5**, trips establish this important connection, and hence freight trip data can help to enrich our knowledge about pre-trip decisions i.e. routing and scheduling characteristics of logistic firms. It also can help us better understand the after-trip phenomena like the dynamics of traffic on road networks associated with these trips. This provides us with the opportunity to turn complex modern trip data into valuable knowledge that can support innovations in logistics and traffic models. These innovations require data-driven modelling of freight transport and traffic system which begins with monitoring and analyzing available data in these systems for identification, comprehension, prediction, optimization, and then tactical planning for (re)designing policies and services.

The data-driven model building in this thesis is based on combining and analyzing different data sources that represent real-world spatial-temporal characteristics of logistics and traffic systems. Previous studies were limited to local surveys and expensive and yet aggregate data collected from different parts of these systems (Ben-Akiva et al., 2016). In recent years, more data from multiple sources have become available due to advanced automatic data collection technologies and data-sharing platforms. The main needs include the following data:

(1) Freight transportation data that include micro-level information about tours, trips, and shipments. An example of this data is the XML microdata that is collected by the Statistics Netherlands (CBS). Every record is tagged with an ID that can relate shipments to the tours and the trucks and includes details about the departure time, loading and unloading locations (GPS), distances, and commodity characteristics like weights and type.
(2) Logistics data that include information about firms and their logistic activities. Examples are the Dutch Firm Establishment Survey dataset including information about all firms registered, their industrial classification, location, number of employees, and several other attributes, and the Portbase Port Community System dataset which contains container processes from vessel arrivals and unloading containers from vessels to loading them on trucks in chronological order. And finally,
(3) traffic data that can present vehicle counts and speeds (both trucks and passenger cars) on road networks.

As mentioned, these data come from various sources and have been collected by different organizations, at different times and by different standards. Utilizing these data for developing integrative freight transport and traffic modelling requires decent data preprocessing, gap filling, and data matching. The literature, however, also lacks systematic approaches to combining and using these data for understanding the connection between logistics and traffic.

Finally, with the models and data described above, it becomes possible to develop prescriptive approaches to design transport policies that focus on logistics, aiming to reduce congestion. Earlier studies and simulation experiments like Mahmassani and Jayakrishnan (1991), Yoshii et al. (1998), and Thorhauge et al. (2016b) reported that temporal demand spreading like shifting departure time of commuters is a more effective traffic mitigation policy as compared to policies that steer drivers to choose alternative routes. As opposed to passenger transport, research on peak avoidance policies for freight transport through departure time shift is relatively few and limited to urban freight deliveries (Sánchez-Díaz et al., 2017). Until now, no single study exists that considers the impact of such policy around a logistic hub like container terminals taking into account multiple stakeholders' activities, logistics operations, and traffic systems.

1.3 Knowledge gaps

Given the overview of the research background above, we identify 4 main research gaps that we highlight as follows:

- the understanding of the logistics determinants of freight transport flows and their interaction with the traffic system;
- integrative freight transport and traffic modelling with support from real world-data;
- approaches for systematic logistics, freight transport, and traffic data matching and preprocessing, for their combined use in freight transport modelling; and
- knowledge of the effectiveness of transport policies that target logistics decisions with the aim to reduce impacts on traffic flow.

1.4 Main research question

Following the knowledge gaps for understanding the interrelation between logistic and traffic systems and the need for developing data-driven freight transport models for integrating them, we can now highlight the main research question:

How can we develop data-driven freight transport models that can predict and control the mutual impact of tour planning (vehicle routing and scheduling) and traffic states?

In the next section, we explain the research methodology that helps us systematically find the answer to this key research question.

1.5 Research methodology and sub-questions

Our research method to develop data-driven freight transport models follows four sequential steps as illustrated in **Figure 1-6**. It begins with processing multiple raw data to acquire valuable knowledge that can describe pre-tour-planning decisions. These decisions are the type and departure time of tours. After identifying the type of tours, a data-driven routing and scheduling model can generate tours of each type and predict truck flows between zones. Next is a traffic model that can predict the impact of these time-dependent truck flows on the state of traffic on road networks. Finally, it includes some control mechanisms that can support decision-makers to define and assess policies. The following of these four steps breaks down the main research question into 7 sub-questions helping us to step by step reach our research objective.

1) Preprocessing and data matching are prerequisites for modelling data-driven freight transport models. Therefore, we need to first develop a systematic and standard way of combining freight trip data with logistics and traffic data in order to enrich one from the other. This leads us to the first sub-research question. *How can freight trip data sources be enriched with logistics and traffic data?* **[chapter 2]**

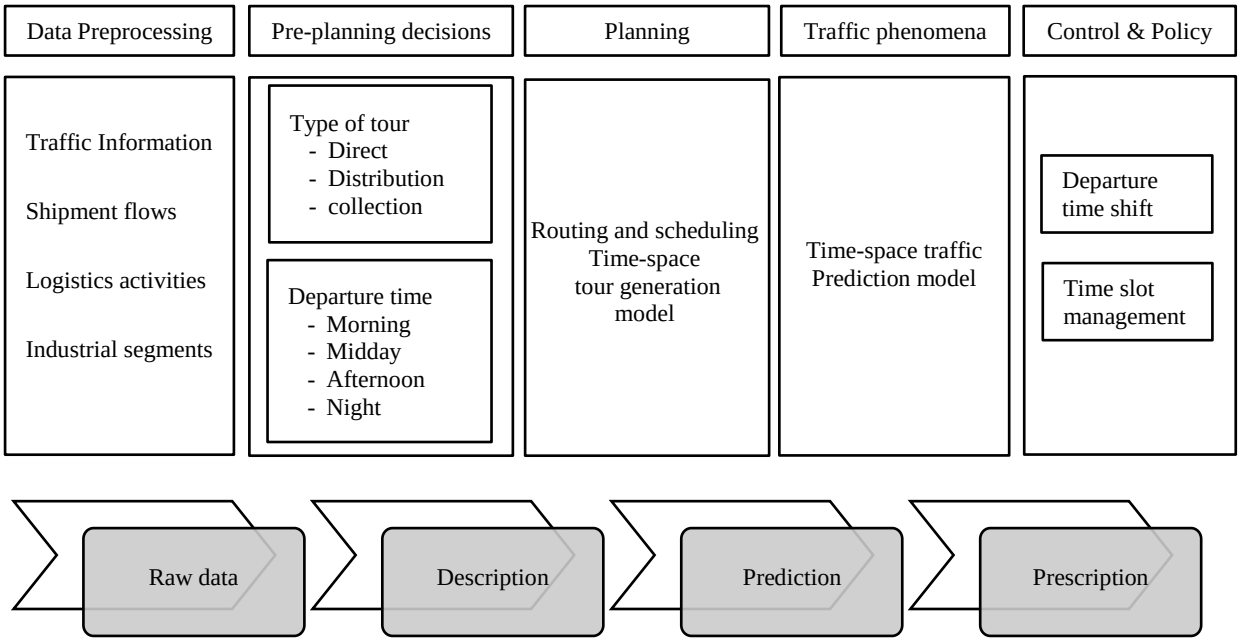


Figure 1-6: Research methodology

2) After data integration, we can use these data to develop models that help us understand and describe the characteristics of transport-related decisions before planning tours. These decisions describe the high-level characteristics of tours that are mainly made before planning a tour. We, therefore, need an approach that can extract empirical evidence/knowledge from the enriched freight tour data sources. *How can patterns in departure time (scheduling) and space (type of tours) be understood from mining freight trip data sources?* **[chapter 2]**

3) Once we understand the high-level characteristics of tours, we can develop models to generate the general tour patterns, as close as possibly resembling and reproducing observed tours. *How can we develop a generalized data-driven vehicle routing and scheduling model for the tour-based representation of time-dependent freight transport activities?* **[chapter 3]**

4) After modelling tours from trip data, it is important to predict the impact of tour schedules on the dynamics of traffic on the road network. To this end, we first need to look for empirical evidence from both traffic and trip data sources that proves the short-term relation between schedules of trips and dynamics of vehicle flows on the road network. *How can we model the relation between freight trip scheduling and the volume of freight vehicles on a motorway?* **[chapter 4]**

5) Once the relation between freight trip schedules and the volume of freight vehicles is clear, we need to predict the propagation of this impact on traffic dynamics across the entire road network. To this end, it is important to integrate time and space dimensions in modelling the impact of freight activities on traffic networks. *How can we incorporate spatial and temporal dependencies in predicting traffic dynamics from freight trip schedules?* **[chapter 5]**

6) Now that we have tools to predict the impact of freight trip schedules on traffic networks, we can design a policy evaluation mechanism using our method that is to shift the scheduling of freight trips and assess its impact on the traffic system. This leads us to the next research question. *What are the impacts of truck departure times shift policy on traffic conditions on road networks?* **[chapter 5]**

7) Shifting truck departure time, however, requires a design of an intelligent and context-aware mechanism that can control and manage truck arrival times at terminal gates. This system can mitigate congestion at terminal gates by assigning time-slots to trucks for their container pickups and drop-offs. For this, this system should consider scheduling preferences of trucks, terminal operation, and traffic on the road networks. *How can we design a time slot management system for seaport terminals to mitigate congestion at terminals' gates considering road networks conditions, scheduling of trucking companies, and terminal operations?* **[chapter 6]**

1.6 Contributions

Our main contribution to the field of freight transport and traffic modelling is developing a data-driven modelling approach that connects the pre-trip decisions of carriers to the traffic dynamics on road networks. We provide tools and modelling frameworks for using available freight trips, logistics, and traffic data for descriptive, predictive, and prescriptive analysis of these systems. This paves the way for the integration of logistics and traffic systems where the costs and benefits of both can be seen together. Our study offers the following contributions categorized under four perspectives:

Freight trip data preprocessing

- A new systematic data preprocessing approach for integrating logistics, freight transport, and traffic data. [chapter 2]

Knowledge discovery from freight trip data

- New empirical insights into the influence of congestion on the structure of tours and time of day dynamic tour features. [chapter 2]
- A new decision tree modelling approach for the analysis of freight tour structures identifying tour type patterns. [Chapter 2]
- New methods based on association rule mining for transport market segmentation with commodity pick-up and delivery analysis. [Chapter 2]
- New insights into the impact of freight trips schedules on traffic intensities and its resulting delays. [chapter 4]

Space-time tour patterns and traffic predictions

- A new data-driven vehicle routing problem and parameter estimation method based on Bayesian optimization to capture preferences of planners in routing and scheduling of freight vehicles from partially observed tour information. [Chapter 3]
- New insights into routing and scheduling decisions of carriers. [chapter 3]
- A new analytical framework that makes use of an artificial neural network for prediction and NSGA-II for feature extraction to predict truck traffic from freight trip schedules in a container terminal. [Chapter 4]
- Utilization of real-world container pick-up time schedules, for the first time, for a network-wide short-term prediction of traffic on truck-intensive motorways. [Chapter 5]
- A new graph-based, modular deep neural network, using novel message passing and neighborhood aggregation rules that follows traffic flow theory to capture interpretable spatial and temporal patterns in traffic. [Chapter 5]

Prescriptions on freight and traffic control

- New modelling for an optimized truck departure time shifts policy that is linked to overall road network performance. [Chapter 5]

- New insights into the possible benefits of shifting trucks' schedules from peak hours. [chapter 5]
- A new method to develop a context-aware decision support system for a time slot management system that controls truck arrivals and departure to/from marine terminals considering multiple stakeholders' preferences and operations in the port and hinterland. [Chapter 6]
- New insights into the costs and benefits of time slot management for port and hinterland operations. [chapter 6]

1.7 Social relevance

This research can be of interest to port authorities, container terminal operators, transport policy makers, and logistics service providers. The result of this projects can also be beneficial to road authorities and traffic control centres who are responsible for congestions which is attributed to inefficiency in freight transport.

1.8 Thesis outline

We outline the chapters of this thesis in this section. These chapters are based on papers that are published or under review. The text in each chapter is identical to the publications. An overview of chapters is illustrated in **Figure 1-7** with each chapter belonging to two or three of the Identifications, Prediction, and Prescription parts.

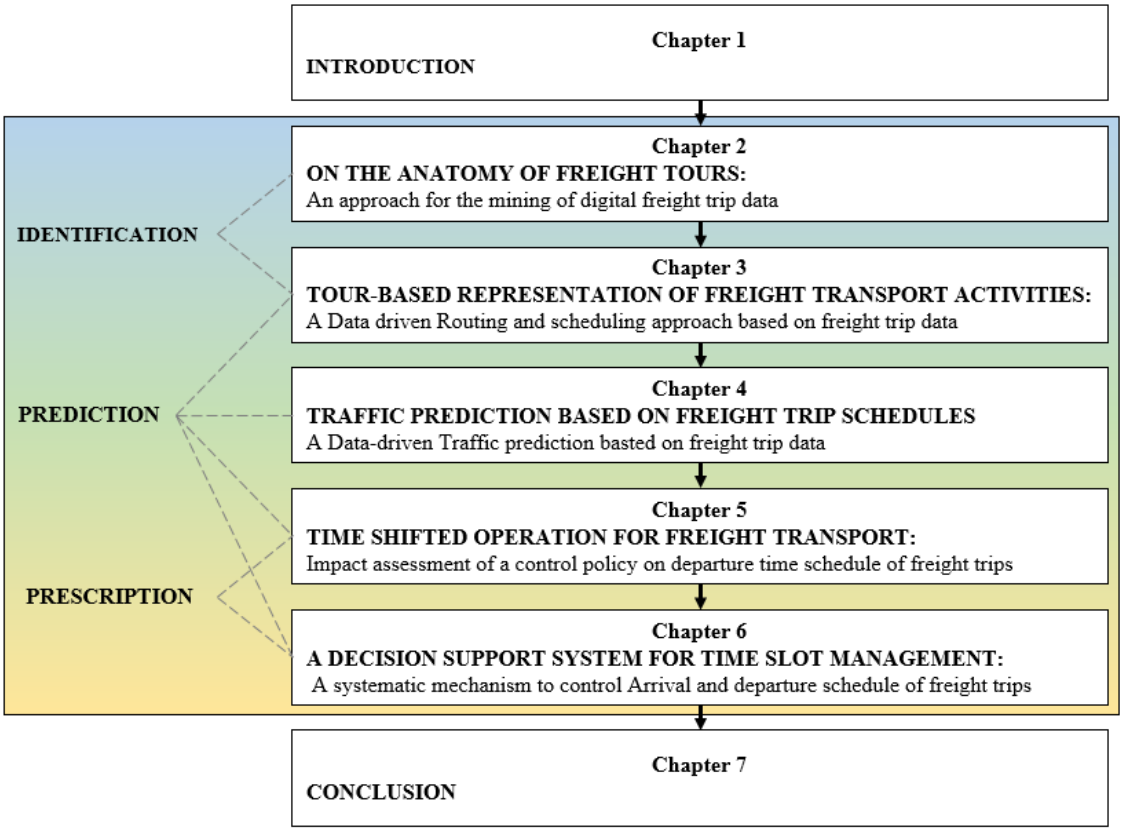


Figure 1-7: The layout of the thesis

Part I - Identification provides empirical insights into the pre-trip decisions of carriers. Chapters two and three belong to this category. Chapter 2, presents a data-driven modelling pipeline and methods for knowledge extraction from a large set of planned tours. This chapter presents clusters of transport markets linking them to the industrial markets using their frequent associated pickups and deliveries. With this, we could increase homogeneity in tour patterns that would help further analysis of tours. This addresses a problem that is rarely discussed in the literature, which results in the poor performance of models due to the heterogeneity in transport markets. Furthermore, this chapter characterizes the structure of tours from time of the day and type of tour perspectives and specifically explores the correlation between the structure of tours with zonal congestion levels. Given the empirical insight into aggregate characteristics of planned tours, we can develop models that can reproduce observed pickup and delivery tours. In chapter 3, we present a data-driven vehicle routing and scheduling algorithm with learning parameters that can learn tour planning decisions from observations.

Given the empirical insight into aggregate characteristics of planned tours, we can develop models that can mimic the pickup and delivery patterns close to observed tours. In chapter 3, we present a data-driven vehicle routing and scheduling algorithm with learning parameters that can learn tour planning decisions from observations.

Part II – Prediction which provides predictions on collective tour patterns and dynamic traffic states. Chapters 3, 4, 5, and 6, are dedicated to this category predicting the pre-trip decisions and the impact of these decisions on road networks. This requires linking trip schedules to the dynamics of traffic. In chapter 3, the data-driven routing and scheduling model can predict collective truck flows between zones and link these flows to the overall tour planning decision. In Chapter 4 we explore the correlation between tour planning decisions and traffic dynamics finding delayed short-term patterns in truck volumes on road networks resulting from the schedules of trucks. Chapter 5 expands this method to a network-wide traffic prediction model that can predict short-term traffic states on a given road network based on given truck schedules. This chapter proposes a graph convolution deep neural network using a special attention mechanism that can capture meaningful spatial and temporal patterns on a congested road network and link these patterns to the trucks' schedules. Next, we use this method to predict the impact of a truck departure time shift policy on the traffic system and accordingly optimize the gain of this policy. Although chapter 5 provides insights into the possible societal gains from traffic systems due to an optimized time-shifted transport policy, shifting the time of trucks is not easy as it could have several consequences on several stakeholders.

Insights from chapter 5 and further analysis in chapter 6 contribute to Part III – Prescription of this thesis where we propose a context-aware time slot management system that aims at controlling the arrival and departure of trucks to a logistic hub considering operations in the logistic hub, Truckers scheduling preferences, and traffic in port and hinterland. This chapter describes the application of integrated traffic and logistic models and finally helps to achieve the main aim of this thesis.

Based on the mentioned studies in chapters 2 to 6, chapter 7 summarizes the main findings of this thesis, highlights our contributions, and gives recommendations for future research.

Chapter 2

Spatial and temporal characteristics of freight tours

This chapter presents a modelling approach to infer scheduling and routing patterns from digital freight transport activity data for different freight markets. The structure of tours for different submarkets, the interactions between industries, and the sensitivity to traffic conditions are important characteristics of freight transport that have not yet been analyzed in detail. This research aims to help fill this gap by developing a series of models that help explain the mechanisms, based on disaggregate shipment and trip observations. We provide a complete modelling framework including a new discrete-continuous decision tree approach for extracting rules from the freight transport data. We apply these models to collected tour data for the Netherlands to understand departure time patterns and tour strategies, also allowing us to evaluate the effectiveness of the proposed algorithm. To support the research community, an open-source implementation of all algorithms is made available.

This chapter is based on the following papers:

1. Nadi, A., Snelder, M., van Lint, J.W.C., & Tavasszy, L. (2022). Spatial and temporal characteristics of freight tours: A data-driven exploratory analysis - submitted to a journal.

2.1 Introduction

Two key structural characteristics of tours in freight transportation are schedules for pick-up/delivery and the routing patterns connecting the stops. Identifying the structure of tours is useful for freight transportation decision-makers and traffic managers, as they support the descriptive and predictive analysis of the freight transportation system. It is important for freight and traffic management, as these structures are conditioned upon and also determine the state of the traffic on motorways. From the demand management perspective, modelling time-of-day choices casts light on the scheduling of departures and modelling tours elucidates routing patterns of commercial vehicles. As opposed to passenger travel, researchers have devoted relatively little attention to modelling freight transport trips. The reason for this is the complexity of this part of the traffic system, particularly as freight transportation includes many submarkets (Khan and Machemehl, 2017a, Holguín-Veras and Patil, 2005). Previous research has addressed multiple dimensions of freight trips (Garrido and Mahmassani (2000), Nuzzolo and Comi (2014), Wisetjindawat et al. (2012), and Figliozzi (2010)). To the best of our knowledge, however, the structure of tours for different submarkets, the interactions between industries, and the sensitivity to traffic conditions have not yet been analyzed in detail. This research aims to help fill this gap by developing a series of models that help explain the mechanisms, based on disaggregate shipment and trip observations.

The opportunity for this work arose from access to a large and rich database of micro-level observations of freight transport activities in the Netherlands. As the main direction of the work, we consider three dimensions of tour characteristics i.e. time, type, and size. These entail departure times of tours, the sequence of loading and unloading activities, and tour length. Most of the existing works of literature are used to model tour structure through behavioral modelling using micro-simulation of classic discrete choice methods. These choice models can be used to study the preferences of decision-makers when making a single (discrete) choice among a set of alternatives. In our study, however, we are more interested to identify patterns and the circumstances in which tours occur from a given dataset, without enforcing any theory or behavioural assumptions on this process ex-ante. We believe that viewed across a larger population of firms, the executed tours will display similarities under particular circumstances. Knowledge of these patterns is valuable for practitioners to know and exploit. Here, we aim to distill these patterns using a data-driven approach.

In this chapter, we propose a machine learning approach to describe the structure and timing of freight tours, based on association rule mining and decision trees. We use rule mining to investigate relationships between different clusters of freight transport markets, and decision trees to discover tour and timing patterns from a large database of freight transport activities. We use the information extracted from 9 identified transport markets. Specifically, we ask the following questions:

1. How do tour strategies and external circumstances affect the peak-hour occurrence of tours?
2. How do the types of logistic nodes (e.g. distribution centers, transshipment terminals) influence departure times and tour structures?
3. How do transport costs influence the structure of tours?
4. How does the number of shipments (i.e. from empty containers to containers with many shipments) affect the scheduling of tours?
5. How does the type of vehicle affect the routing and scheduling of shipments?

These research questions are answered using a two-step approach. In the first step, we use decision trees to explore patterns in departure times of tours. Earlier studies have identified that freight tours encompass multiple inter-related decision variables, such as the type of tour pattern and the number of stops in each selected tour type (Khan and Machemehl, 2017a). In the second step, we mine probabilistic rules that (a) describe the various pick-up and delivery strategies and (b) predict the average number of stops per tour for each strategy simultaneously. Prediction of a mixture of target variables is needed in the second modelling task. However, most machine learning techniques can only predict one type of target variable at a time. To deal with this problem, we introduce an enhanced decision tree algorithm that can predict joint mixed target variables.

These two models together explain the anatomy of tours taking both scheduling and routing activity of various transport markets into account. The main contributions of this paper are the following:

1. We introduce a new data processing framework to support the process of inferring patterns in big, disaggregate freight tour databases.
2. We extract rules from observations to cluster the freight transport market and investigate the interaction of multiple industries, through an analysis of their tour activities.
3. We propose a new decision tree algorithm that predicts mixed discrete and continuous target variables to classify tour structures and predict the number of stops per tour strategy. This innovative algorithm can be used for modelling similar cases with other mixed-type target variables.
4. We provide new insights into how transport companies plan their tours in the presence of congestion.

The chapter is organized as follows: Section 2.2 provides a brief overview of the existing studies on the topic of this paper. Section 2.3 presents the methodology with which we characterize freight activities, including the description of the data structure and the fusion of multiple data sources. Section 2.4 presents the findings from the descriptive analysis of the freight routing and scheduling patterns. Section 2.5 sets up a discussion about findings, provides validation against literature and suggests directions for further research. Section 2.6 offers concluding insights.

2.2 Literature review

This section presents a brief review of the literature that concerns freight tour activity modelling. The literature reports several important factors to take into account. Primarily, it is the operational decisions of shippers, senders, receivers, and carriers that determine truck flows (Khan and Machemehl, 2017a). Logistics hubs (e.g. distribution centres and transshipment terminals) may add complexity to the tours due to their different functionalities for storage and cross-docking of freight. Tour generation is dependent on important dimensions that are often difficult to measure due to privacy issues. Examples are vehicle type, weight, capacity, container type and dimension, departure and arrival times, number of stops, and tour duration (Khan and Machemehl, 2017a). Besides, freight transportation demand is highly variable over time and space (Garrido and Mahmassani (2000)). The heterogeneity in transport markets and industrial sectors creates additional diversity in tour characteristics.

To cover the above issues, disaggregate activity-based modelling has been developed in the literature capturing daily activity patterns of goods movements. Recently, more research is

focusing on agent-based micro simulation of multiple actors' decisions in freight transport systems. Examples are MASS-GT which is an agent-based simulation system for urban goods transport. MASS-GT simulates the choices of suppliers and receivers, transport service providers and carriers (de Bok and Tavasszy, 2018). These choices include distribution channel selection, carrier selection, vehicle type, shipment size, and routing and scheduling decisions. Another example is SimMobility developed by Sakai (Sakai et al., 2020) which can simulate disaggregate interactions of multiple agents that are engaged in freight-related activities. These decisions include commodity contracts, overnight parking place choice, carrier selection, and vehicle operation choices. Lately, agent-based freight simulators are also integrated with agent-based passenger simulators like MATSIM to model the impact of freight transport decisions on traffic system (Schröder and Liedtke, 2017). For example, Mommens et al. (2018) used TRABAM, which is a freight simulator linked to MATSIM, for the impact assessment of traffic management scenarios like night distribution.

All in all, the common choice models that exist in the most agent based freight simulators include (a) tour purpose and vehicle type choice; (b) next stop destination choice; (c) next stop purpose choice; (d) joint tour-type and number-of-trip choice and (e) departure time choice. Among all, time-of-day (i.e. scheduling) and type-of-tour (i.e. routing) are the two main characteristics of freight transport activities that are believed to be conditioned upon the level of congestion on motorways (Khan and Machemehl, 2017b) and thus are of interest for traffic and transport managers.

Time-of-day modelling is relevant as variations in departure time of freight flows may have large impacts on motorway congestion; whereas in turn, those congestion problems may also heavily affect logistics operations (Figliozzi, 2010). The application of a time-of-day model helps researchers to evaluate shifts between peak and off-peak periods for freight transport as a means to avoid congestion on roads. In spite of its significance, there exists little research on freight transport departure time policies. The earliest time-of-day modelling frameworks utilized Monte Carlo simulation in which departure times were averaged from a sample of (limited) observations (Hunt and Stefan, 2007). Probably, the study of Nuzzolo and Comi (2014) is one of the first studies that modeled the departure time as a part of the freight demand modelling framework. In this study, like most similar studies, a discrete choice model was estimated on urban freight trip survey data to calculate the probability of a delivery tour departure time from an origin. The results indicate that the departure time of a tour is strictly correlated with the number of stops per tour. A recent time-period choice model based on stated preferences for road freight transport can be found in (De Jong et al., 2016). They estimated a multinomial logit model on SP data of 158 receivers of the goods. This model captures the sensitivity of peak and off-peak delivery choices to changes in travel time and cost. It shows a stronger sensitivity to travel costs than to travel time. This study was followed up by Vegelien and Dugundji (2018) who used revealed preference data (GPS tracking data of trailers) to model time-of-day choice, with a nested logit model. This model includes trip duration and product type as explanatory variables. Khan and Machemehl (2017b) proposed a discrete-continuous probit model to recognize trucks' time-of-day preferences and predict the vehicle-mile traveled at a specific time of the day. This model was estimated on commercial vehicle survey data. It shows the contribution of spatial and temporal factors including vehicle type, commodity type, unloading weight, characteristics of intermediate stops speed, and service time. However, this study does not include the impact of traffic on time-of-day preferences which has our interest.

Similar to the time-of-day model, a small number of studies specifically investigate tours. An early study of commercial vehicle delivery strategy can be found in (Burns et al., 1985). They formulate transportation by truck and inventory costs of freight to evaluate the cost trade-off between the direct shipping and collection or distribution type of tour strategies, based on shipment size and the number of customers. They show that in the direct shipping strategy the optimal shipment size follows the economic order quantity (EOQ) rule while in collection or distribution, the shipment size is close to a full truckload. Ruan et al. (2012) followed this study further using mixed and multinomial logit models to model multiple types of tours. This model helps to understand decisions regarding distribution strategies and tour chaining. Zhou et al. (2014) studied tour patterns of urban commercial vehicles based on the number of trips made for delivery and pickup activities. They defined four tour-type alternatives bundling different ranges of the number of trips with different tour types. Then, they used multinomial and nested logit models to identify daily trip chaining behavior. Khan and Machemehl (2017a) considered the number of trips as a continuous dimension of tour activities. This is more appropriate since it refers to the size of tours and there is ordinal relation between tours with different sizes. They develop a multiple discrete-continuous choice method to model a joint distribution of the type of tour strategy and the number of trips for commercial vehicles. Based on their findings, the number of trips and type of tour are two inter-related joint decisions and unified discrete-continuous modelling is a more appropriate modelling framework to capture commercial vehicle movement patterns. The results of this study also indicate the impact of trip features, commodity characteristics, and attributes of base and intermediate stop locations on choice of the type of tour and the number of stops per each tour chain strategy. Most recently Siripirote et al. (2020) proposed a statistical approach to estimate truck activities, commodity-related trip chains, and the status of legs (e.g. empty, full, or partially loaded trips) from collected GPS data. This class of demand modelling uses different units of analysis, either trips, as in Hunt and Stefan (2007) or shipments (as in Thoen et al. (2020) and Nuzzolo and Comi (2014)) and utilizes choice models to explain daily tour patterns.

In summary, the mentioned studies underpin the importance of modelling the time-of-day and type of tour strategies of commercial vehicles in urban areas. We note several gaps, however. Firstly, to the best of our knowledge, existing studies on time-of-day and type of tour analysis relate to urban logistics and do not consider freight transport activity between large suppliers, manufacturers, and logistics hubs. Secondly, all of these studies are built upon a limited sample of survey data from the sender or receiver of goods. The literature lacks a systematic data-driven modelling pipeline that allows us to estimate models on large-scale real tour data. Thirdly, the interaction between multiple industries and their impact on time and type of tour analysis has not yet been presented in the literature. Finally, all existing studies used classical choice models to study the preferences of decision-makers, focusing on a choice between some discrete alternatives. The currently dominant approach to descriptive modelling of tours assumes discrete choices; either using a “tour growth approach” which assumes stop-by-stop discrete decisions for the next trip; or an enumeration approach suggesting that the planner has multiple full tours to choose from. In reality (i.e. overwhelming majority of cases) tours are not created this way. A data-driven approach allows extracting valuable knowledge without making such potentially problematic assumptions about the real-world decision process.

To help fill these gaps in research, we develop a new modelling approach utilizing and enhancing rule-based machine learning techniques to explore the departure time, type of tour, and the number of stops per tour pattern, using an extensive freight transport tour database. The next section elaborates on the methodology of the research.

2.3 Modelling framework for knowledge extraction from tour data

In this section, we explain the methods and data used for the analysis of the freight tour databases. Figure 2-1 shows the overall rule-based modelling framework that allows us to extract appropriate rules about departure time and type of tour strategies in freight transport activities. This framework includes 3 main steps from the preprocessing of data to the full analysis of the structure of tours.

1. In the first step, data is collected from databases that include the trip survey data and contextual data relating to speeds as well as the locations of important hubs including distribution centers and transshipment terminals.
2. In the preprocessing step, data is prepared through the fusion of databases and clustering of transport markets. Also, the proximity to congestion zones is modeled to prepare for the tour structure analysis.
3. In this third step, an enhanced decision tree algorithm is used to model time-of-day patterns, type-of-tour strategies, and the number of stops.

The next subsections describe the details of the approach in this order.

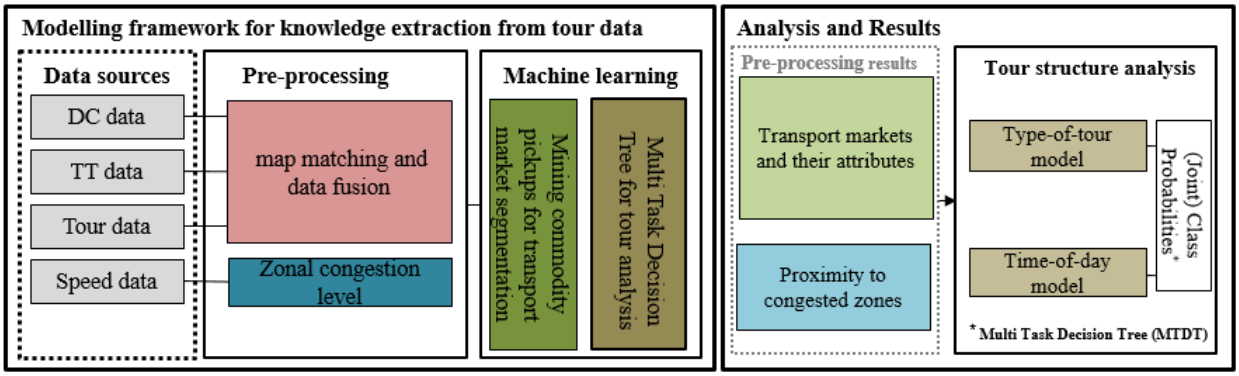


Figure 2-1: Procedure for modelling and extracting knowledge from freight tour data

2.3.1 Data sources

In this study, We make use of a unique and detailed dataset of truck diaries provided by the Centraal Bureau voor de Statistiek (CBS), or Statistics Netherlands. In total CBS provides a set of digitally collected 2.7 million records of data related to freight tour activities from the year 2015. For each commodity in a tour, there is a record in this dataset that contains geographic information regarding (un)loading locations, commodity type, vehicle types used, and other tour characteristics. The data however lacks logistic activity of the visiting locations or the sender or receiver of the shipments. We require this information to understand how different types of logistics activity may change the structure of tours. To enrich the database, we fused it with two external data sources. The first dataset contains over 1600 registered distribution centers along with their 6-digit postcode, size, and sectors. This dataset was sourced from Directorate-General for Public Works and Water Management (Rijkswaterstaat) in the Netherlands. The second dataset contains transshipment terminals (TT) from the IDVV-binnenvaart game (IDVV inland shipping game) containing information regarding 76 terminals and their postcode, size, and annual throughputs. Finally, we used national firm establishment data that includes firms' registration dates, sizes, and business types. More information about these data and the use of it can be found in Zhang and Tavasszy (2012). As described in the

next section, we used a matching algorithm to match the pickup and delivery location of vehicles to the right logistic activities using these two data sources.

2.3.2 Pre-processing

Map matching and data fusion

To impute the activity type of loading and unloading locations in CBS tour data we propose a hierarchical assignment approach on national geographical divisions based on both 6 (PC6) and 4 (PC4) digits postcodes. Having L ={l_1,l_2,…,l_n} as a set of all geographical locations, we can define $L_{6-digits} \subset L_{4-digits} \subset L$ where $L_{6\text{-}digit}$ has the highest resolution in national geographical divisions. We give precedence to the PC6 because the divisions are smaller and thus the assignment takes place with higher certainty. We then aggregate the information at this level for use in PC4 where the imputations are more probabilistic.

The algorithm, first, uses firms establishment data and available national statistics to calculate the probability $P_{s,c}^{a,l}$. This shows the probability that a shipment *s* that picks up or delivers commodity type *c* at its loading or unloading location *l* belongs to a firm with activity type *a*.

$$P_{s,c}^{a,l_{6-digit}} = \frac{N_{a,l^{6-digit}} \times P_{a,c}^{make/use}}{\sum_{i \in A} N_{i,l^{6-digit}} \times P_{i,c}^{make/use}} \qquad \forall s \tag{2.1}$$

Where $N_{a,l}$ is the number of firms with any of the activity types A={ distribution centres, transshipment terminal, producer, consumer } in each location *l* and $P_{a,c}^{make/use}$ denotes the probability that firms with activity type *a* are making (if *l* is loading location) or using (if *l* is unloading location) of the commodity type *c*. Then the algorithm uses the Monte Carlo simulation to assign each activity type to the loading and unloading locations of a shipment based on $P_{s,c}^{a,l}$.

Please note that in cases that the loading and unloading location of a shipment is reported in a higher resolution i.e. 6-digit, the $P_{s,c}^{a,l}$ is in many cases 1 or close to one as there is a lower chance of having multiple firms with different activity types in a 6-digit zip code. Therefore, the assignment process in these cases is more deterministic than probabilistic. However, this is the opposite when the reported locations are in 4-digits and the assignments would have a high level of uncertainty. To reduce this uncertainty level for 4-digits locations, the algorithm uses the results from assignments in 6-digit resolution. First, the probability that a shipment with commodity type *c* goes to a zone $l_{6\text{-}digit}$ inside the reported larger geographical division (4-digit) is calculated.

$$P_{c,l_{6-digit}} = \frac{M_{c,l_{6-digit}}^{in/out}}{\sum_{l \in L_{6-digit} \subset l_{4-digit}} M_{c,l}^{in/out}} \qquad \forall l_{4-digit} \tag{2.2}$$

Where $M_{c,l}^{in/out}$ is the number of shipments with commodity type c that go inside or come out of the zone *l*. Then a Monte Carlo simulation is used to select and assign each shipment in a tour to an $l_{6\text{-}digit}$ location inside the reported $l_{4\text{-}digit}$. Afterward, a similar approach as in 6-digit postcode zones (i.e. Equation (2.1)) is used to assign an activity type to the visited location.

There are cases where the commodity type is missing or not reported in the dataset. However, our algorithm calculates the assignment probabilities without considering the term $P_{a,c}^{make/use}$ and indices *c* in Equations (2.1) and (2.2). In other words, the assignment probabilities only depend on $N_{a,l}$. Please note that these cases are not considered for further analysis in this research.

Zonal congestion level

Besides logistics data, we also required traffic data to understand what the traffic conditions were at the shipment time. Below, we explain the methods used to obtain traffic conditions by making use of traffic speed data.

Carriers usually use planning software to schedule the tours. This software uses the average travel time while making routing and scheduling tables. To understand how the congestion can affect tour characteristics, we require a perception of congestion at the time of deliveries. We make use of vehicle speed data derived from loop detectors to obtain congestion information for the pickup and delivery locations. We use the same method introduced by (Christidis and Rivas, 2012) to calculate the aggregate congestion level for each geographical zone. In this approach, a moving average of delays is determined as the congestion indicator. To calculate this delay, we first calculate the moving average speed of each loop detector on a road over a time filter T to smooth very short-term fluctuations. Note that the speed in Equation (2.3) is instantaneous.

$$\bar{v}_i = \frac{1}{T}\sum_{j=i}^{j+T} v_j \tag{2.3}$$

Where *i* is the start of the period, and v_j is the average speed obtained from sensors on the link for the time step *j*. While we define v_{min} as the lowest average speed for the segment, the free flow speed v_{free} is the maximum measured average speed:

$$v_{\min} = \min_{j\in[t_0,t]} \bar{v}_j \tag{2.4}$$

Where t_0 and *t* are the start and end of a period *P*. Then the average maximum delay during period *P* for each segment is:

$$d_P = 60\times\left(\frac{1}{v_{\min}^p} - \frac{1}{v_{free}^p}\right) \tag{2.5}$$

We use d_P as an indicator to define the congestion level for period *P* measured in minutes per kilometre. For cases where v_{min} =0, we hard coded the value of the v_{min} to a very small number to prevent generating undefined delays. One can specify the *P* parameter to limit the delay estimation for a specific period. We set *P =1* hour for each departure time specification i.e. morning peak , midday, evening peak hour, and rest of the day. Based on the average delay indicator, We calculate the 1-hour aggregate of the congestion level (CL) for every zone (zoning system is based on the 4 digits postcode):

$$CL_{z,P} = \frac{\sum_{r\in z} l_r d_{r,p}}{\sum_{r\in z} l_r} \tag{2.6}$$

Where l_r is the length of segment r in zone z. We consider 10 seconds delay for every kilometre in the peak period, as is recommended by Christidis and Rivas (2012) to identify if a zone is congested.

$$CI_{z,P} = \begin{cases} 1 & CL_{z,P} > 10s \\ 0 & otherwise \end{cases} \quad (2.7)$$

Where $CI_{z,P}$ is the congestion indicator which is 1 if the zone is congested and 0 otherwise. In section 2.4.1, we explain how this indicator can come into use for analysing tour structure.

2.3.3 Machine Learning

Mining Commodity pickups for transport market segmentation

Clustering the activity of carriers is an important preliminary step in understanding disaggregate patterns of tour structures, due to the heterogeneity of decision rules and conditions in different transport markets. In this section, we adopt a machine learning method that is used in the analysis of structured data collected from a sequence of activities of individuals such as in market basket analysis (Raeder and Chawla, 2011). We use this method for clustering transport markets considering their tour activities. To the best of our knowledge, this method, which we adapt in several ways for our purposes, has not been applied to transportation. The Classification of Products by Activity (i.e. CPA-classification) is an international standard that consists of a complete hierarchy of divisions, groups, and classes. This standard classifies companies based on their economic activity. The CPA-classification enables governments to describe companies according to the type of product that they produce. Since 1967, the standard goods classification for transport statistics (NST-R) has been used and linked to the CPA-classification at the European level (Liedtke and Schepperle, 2004). The NST-2007, for example, consists of 81 goods categories which are aggregated under 20 main commodity type chapters. Although NST-R gives a direct impression to the economic sectors, it is not clear how carriers can work with different sectors. One carrier may just adopt his utilities to carry agricultural products while others may use other facilities to transport mineral products. These carriers may have different decision rules in planning their tours due to the sectors' conditions and protocols. However, some carriers work with multiple sectors. This happens when there is an interaction between two or more sectors in the supply chain. Therefore, CPA-classification alone cannot be used to cluster transport markets accurately. To understand accurately how carriers' tour planning rules differ, we need to cluster them based on the different sectors that they work with. In this section, we use an association rule mining technique to introduce a commodity pickup analysis. This analysis is to extract a set of frequent rules in pickup and delivery data to explain how carriers interact with multiple sectors. In this analysis, we have a set of Tours $T=\{T_1,T_2,..,T_n\}$ in the CBS freight transport diary database. Having a set of commodity items $I=\{I_1,I_2,\ldots,I_3\}$ which are the goods categories in NST-2007, each tour T_i consists of a set of pickups with commodity types I_i (i.e. $T_i \subseteq I$). We aim to extract strong rules like $A \Longrightarrow B$ which implies that carriers who transport commodity $A \subset I$ are also willing to transport commodity $B \subset I$ in their planned tours (A and B are two disjoint subsets where $A \neq \phi, B \neq \phi$, and $A \cap B = \phi$). By definition, a rule is strong if the usefulness and certainty of that rule are higher than minimum thresholds. We can measure the usefulness and certainty of a rule using *Support* and *Confidence* indicators respectively.

$$Support(A \rightarrow B) = P(A \cap B) \quad (2.8)$$

$$Confidence(A \rightarrow B) = P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{2.9}$$

Where *Support* is the probability that a carrier transports both commodity types A and B, and *Confidence* is the probability that a carrier works with sector B given that they also work with sector A. Although *Confidence* and *Support* can identify strong rules in data, some strong rules may happen randomly and thus we should not necessarily consider them as important rules. Therefore, another indicator that can capture the correlation between picking up commodity *A* and picking up *B* should be considered. This indicator is known as *Lift*.

$$lift(A \rightarrow B) = \frac{P(A \cap B)}{P(A) \cdot P(B)} \tag{2.10}$$

The *Lift* indicator can get any value greater, equal, or less than 1. *Lift* equal to 1 means that the rule happened randomly while *Lift* greater than one implies that the occurrence of picking up commodities A and B has a higher chance than just a random occurrence. In other words, the *Lift* indicator considers the correlation between the occurrence of picking up commodity types A and B. Therefore, the greater the *Lift* is, the higher chance the rule has of occurring.

The complexity of finding all strong and important rules is $2^{\mathrm{m}} - 1$ for m items in the item set. This means that verifying all the possible rules in the database is, in some cases, impossible. However, algorithms like Apriori (Agrawal and Srikant, 1994) use the close-set and maximal-set concepts to limit the search space and solve such complex problems. This algorithm iteratively uses a sequence of pruning and joining processes to capture the most prevalent rules with minimum *Support* and *Confidence*. Apriori algorithm is a non-parametric algorithm that can capture frequent patterns in a structured dataset. Despite other advanced machine learning algorithms like Bayesian networks, this algorithm makes no assumptions about the random variables and the dependencies among them. This makes it a good candidate for our analysis especially when the number of items in a set (here customers' in a tour) is not too large. Based on the most important rules derived from this method, we can understand how carriers interact between different sectors and we can cluster transport markets into a number of submarkets with similar pickup and delivery patterns. We use the 'arule' package in R programing language (Hornik et al., 2005, Hahsler and Chelluboina, 2011) for implementation and discuss the result of this analysis in section 2.4.

Multi-task decision tree for tour analysis

Given all previous pre-processing steps, we can now take a step forward towards the development of a method for explaining the anatomy of tours for each of the submarkets. Several machine learning techniques can be used to classify or predict the characteristics of tours. There is a non-linear relation between feature space and dependent variables in this context. Therefore, we need a nonlinear method like kernel-based support vector machines or neural networks that can deal with such non-linearity. However, these methods usually suffer from a lack of interpretability. Among other machine learning techniques, there are two classes of methods that can deal simultaneously with non-linearity and interpretability. The first class is the generalized additive model (GAM) that was introduced by Hastie (1987). GAM is a sum of weighted smooth functions that apply transformations in the linear regression models. The focus of parameters in this approach is on the inference made by these smooth functions on the predictors. The second class is the Decision trees (DT) model which is a supervised learning method with a wide range of applications in both regression and classification problems (Breiman et al., 1984). This method obtains a set of rules by taking the interaction of covariates

and can explain the variability of the response variable by recursive partitioning of all the data according to the most significant covariate (Loh, 2014). In other words, a decision tree is a graphical representation of all paths to a decision under certain conditions. These models are not only simple but also powerful to build descriptive models. They usually outperform other classic statistical models like linear regression (in regression problems) and logistic regression (in classification problems) if the relationship between covariates and response variables is highly non-linear. Compared to other black boxes non-linear machine learning predictors with high prediction accuracy, DT is fully explainable and descriptive. The other advantage of the Decision Tree method is that it makes no assumptions regarding the probability distribution of variables (as we do for instance in naive Bayesian classifiers) but is still able to identify explanatory variables and detect interactions among them. All these reasons made us choose to employ and enhance this methodology for our use. In the next section, we explain how and why this method is enhanced to be used for the structure of tour analysis.

Most machine learning methods including decision trees usually support single target variables. However, In some use cases, it is important to model multiple target variables simultaneously. In the modelling terminology, if the target variables are categorical, then it is called multi-label or multi-target classification, and if the target variables are numerical, then we call it multi-target regression (Borchani et al., 2015). Here, however, we want to construct a decision tree that can predict mixed discrete and integer target variables. The application of this problem is to predict the type of tour strategy (a categorical quantity), and its associated number of stops (an integer quantity) that take place within that strategy. There are some techniques to handle this problem:

1) Create multiple single models each for one of the target variables. This method however does not consider the interaction and possible correlation between target variables.
2) One can also create multiple single models each for one target variable with other target variables appended to the feature space. Although this method takes the interaction of target variables into account, there are some drawbacks. First, creating multiple models is difficult to interpret. Besides, in most of the statistical and machine learning models, we assume the attributes are independent. Therefore, putting target variables into the feature space may contradict this assumption. Also, this method can be interpreted as one target variable being conditioned on the others. It cannot be interpreted as a joint occurrence of them which is desired in our case.
3) The third approach builds a single classification model that classes are a pairwise combination of the categorical variable with each level of an integer variable. This increase the number of class in the model and the model may not be tractable. A solution for this is to categorize the integer variable into a few classes. However, finding the right number of clusters for integer variables is still challenging.
4) Using a multivariate distribution like Copula is another approach which in the case of integer-discrete variables may lead to a large number of classes with an imbalanced distribution which is challenging for most classification algorithms to search in such feature space.

To handle the mentioned problems, we propose a new approach to build one single (unified) DT model with both integer and discrete target variables. In this case, a DT has to learn multiple tasks, i.e. regression and classification, simultaneously. Therefore we called this algorithm a Multi-Task Decision Tree (MTDT). This approach neither requires discretizing target variables nor makes a large number of classes. It also does not make any assumption about the distribution of target variables and yet can take the dependency between them into account. To this end, we adopt the idea from ID4.5 algorithm proposed by Quinlan (2014) to introduce a new version of this algorithm, that can handle multiple mixed integers and categorical target

variables. We applied this algorithm to model the type-of-tour and the number of stops simultaneously. The ID4.5 algorithm starts splitting from the root node and continues on further nodes. The first step in a decision tree is to decide where to split. In other words, which feature should be used to split the data. For classification purposes, information gain (IG) is used to decide on a cut-point. Information gain is the measure of impurity or randomness and is the decrease in Shannon's entropy after the data set is split based on an attribute.

$$H(C) = -\sum_i p(C = c_i)\log(p(C = c_i)) \tag{2.11}$$

$$H(C \mid X) = -\sum_j \sum_i p(c_i \mid x_i)\log(p(c_j \mid x_t)) \tag{2.12}$$

$$IG = H(C) - H(C \mid X) \tag{2.13}$$

Where x_i is the possible levels in covariate X, and c_j is the possible classes in target variable C. However, in regression models, the minimum sum of squared residuals (SSR) is used to reduce the variance in each of the subsets. In general, constructing a decision tree is all about successively finding the most significant attributes and splitting data in a way that returns the highest information gain (for classification) or minimum SSR (for regression). Since the measure of splitting in DT is different for discrete (information gain or IG) and continuous (variance reduction or SSR) variables, the node selection, and splitting phase becomes a multi-objective decision-making process. In other words, we construct a tree finding attributes and splitting data in such a way that maximizes IG and minimizes the SSR at the same time, as in

$$\min F_1 = \min(SSR) \tag{2.14}$$

$$\max F_2 = \max(IG) \tag{2.15}$$

These two decision-making criteria may contradict each other in some cases. For example, the attribute that increases information gain also may increase the sum of squared residuals. In this case, we may not be able to choose among some attributes since they have no priority against each other. To solve this problem, we used the non-dominated ranking approach proposed by (Deb et al., 2000). They introduced this approach to solve multi-objective optimization problems. Here, we use the same line of thinking to rank attributes based on the number of times each attribute is dominated by the other attributes regarding both splitting criteria in modelling mixed target variables. We have a set of attributes A={x_1,x_2, …, x_n} for which we have to determine their position as a node in the tree. By definition, dominance is

$$x_i \, dom \, x_j \Leftrightarrow \forall d \;\; F_d(x_i) \le F_d(x_j) \;,\; \exists d' \;\; F_{d'}(x_i) < F_{d'}(x_j) \tag{2.16}$$

Where d and d` represent dimensions of decision criteria F for attributes x_i and x_j.

Algorithm 1: required steps to implement MTDT

1. for each attribute in set A we calculate F_1 and F_2
2. check the dominance of attributes
3. rank attributes based on the number of times they are dominated and select Rank 1 set
4. select the one attribute from the Rank 1 set based on a secondary criterion.
5. split data based on the selected attribute
6. check the pruning criteria to avoid overfitting
7. repeat the process for each split

Although this algorithm can rank attributes based on multiple splitting criteria, still multiple attributes can be ranked similarly. To choose one of the attributes that are ranked 1, we need a secondary criterion. We introduce two approaches. 1) the modeler can give precedence to one of the criteria, i.e. either to the IG to prioritize classification or to the SSR to outweigh the regression. 2) the second approach is to select the attributes with the minimum distance to the ideal points. If there is one attribute in Rank 1, we select that one as the splitting criteria. If there are two attributes in the Rank 1 set, then we select the attributes with the closest distance to the zero points (0,0) which is the ideal point. In the case of more than two attributes in the Rank 1 set, we select the one with minimum distance to the middle point of the values of the objectives. Pseudo Code 1 gives a more systematic explanation of the entire process of including the ranking algorithm and selection of attributes. We implemented this algorithm in MATLAB R2019 using the statistics and Machine learning toolbox.

Pseudo Code 1:Pseudo code Multi-task decision tree (MTDT)

```
A= list of attributes A={x1,x2, …, xn}
F= calculate  information gain and  the minimum sum of squared error for all xi
R=Rank(A,F)
Rank1= R{1}
Best_Node = use the secondary distance criterion to choose the best attribute from the Rank 1 set.
Split  data on Best_Node and repeat the process for each split until the leaf node

Function  R=Rank (A, F)
        Sp=[]                   % is a list of attributes that are dominated by attribute p
        np=0                    % is the number of times that attribute p is dominated by other attributes
        R{1}= []
        for p in  A
            for q in A
                check if p dominates q
                then add p to the Sp list
                and np =np+1
                otherwise
                add q to the Sq list
                and nq=nq+1
            end
            check if np =0      % check which attributes are not dominated by any other attributes
            then R1=[R1  p] and Rank of p is 1
        end
        k=1
        while true
            Q=[]
            for p in R1             % for all attribute that have rank 1
                for q in Sp
                    nq=nq-1
                    check if nq =0
                    then Q=[Q  q] and Rank of q is k+1
                end
            end
            check if Q is empty then Exit
             else
                Rk+1=Q
                k=k+1
         end
end
```

Figure 2-2 clearly explains how we can select between attributes with similar ranks. F1 and F2 are respectively the objective values for SSR and IG and d is the minimum distance to the

middle point (i.e. F^{mid}) of the values of the objective functions. Assume that we have *i* attributes that are ranked from 1 to 3. If we wanted to choose one attribute, the best would be the one with minimum d^i. The F^{mid} and *d* in Figure 2-2 are computed using formulas 2.17 and 2.18, where F_j^{max} and F_j^{min} are the maximum and minimum of the objective *j*.

$$F_j^{mid} = \frac{F_j^{\max} - F_j^{\min}}{2} \tag{2.17}$$

$$d^i = \sum_{j=1}^{2} | F_j^{mid} - F_j^i | \tag{2.18}$$

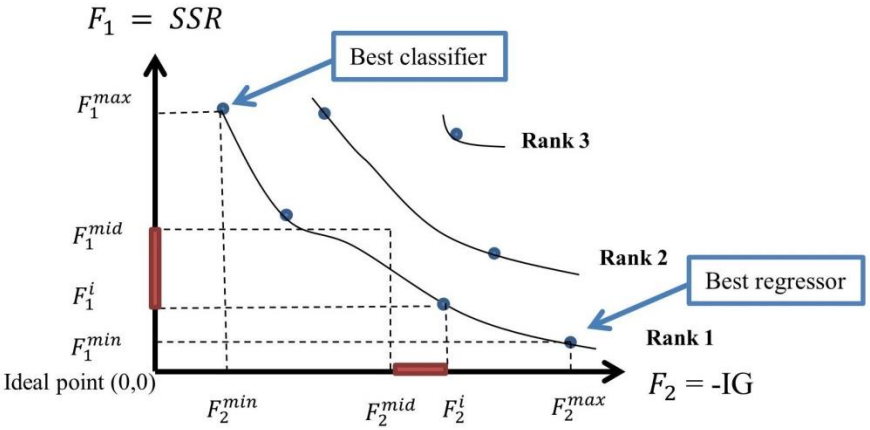


Figure 2-2: Graphical representation of selecting attributes based on the secondary criterion

We also used grid search with cross validation to control the depth of the tree and avoid overfitting. In summary, this algorithm is new in the following respects:

1- It can simultaneously learn two tasks i.e. classification and regression
2- It considers the correlation between discrete and continuous response variables and can predict a joint mixture
3- It uses a non-dominant sorting approach to rank and select attributes under two splitting criteria

As a measurement for the goodness-of-fit, we used the approach proposed by Arentze and Timmermans (2003) and (Kim et al., 2018). We list these measurements in Appendix A for presentation purposes.

We used this algorithm for two use cases in this paper i.e. the time-of-day model and the type-of-tour strategy model. In the next section, we present the result of the analysis based on the proposed rule-based modelling framework.

2.4 Analysis and Results

In the current section, we present the result of utilizing the zonal congestion level indicator. Secondly, we explain findings from the commodity pickup analysis and introduce 9 transport markets with their attributes. We also provide descriptive statistics of explanatory variables.

Finally, we explain results for the anatomy of tours for these transport markets through tour structure analysis.

2.4.1 Preprocessing results

Proximity to congested zones

To use the congestion level indicator defined in Equation (2.7) in our analysis, first, we calculate the proximity of every zone to the congested zones, Because trucks may have to cross several congested zones to reach a destination that is close to or surrounded by congested zones. Proximity is defined as the Euclidean distance between the center of zones. We used the Jenks natural breaks classification method to identify different ranges of proximity to the congested PC4 zones. Figure 2-3 shows a heat map based on the proximity of the PC4 zones to congested zones for the morning peak period. Then we mark each zone as congested if it is close enough to a congested zone using a predefined threshold. It is important to keep this threshold small to only consider real congested zones and their surroundings. For this study, we selected the range 5653-6423 (see Figure 2-3) as the threshold. That is all the zones with a proximity of fewer than 6423 meters to the center of congested zones are also considered congested. Decreasing this threshold to 0 means that there is no proximity considered and hence the effect of congested zones on their neighbors cannot be modeled. On the other hand, increasing the threshold would add more uncongested locations to the list of congested zones due to their proximity. This adds uncertainty to the model. The threshold, therefore, is selected in such a way that only 5% of uncongested locations will be added to the list of congested zones list due to their proximity. This way, there are relatively small to medium zones that are very close to the congested zones.

Finally, we define two binary variables identifying if the first and later pick-up or delivery locations are in a congested zone at the time of arrival. This indicator gives us information on the variation of freight activity patterns based on the perception of the congestion.

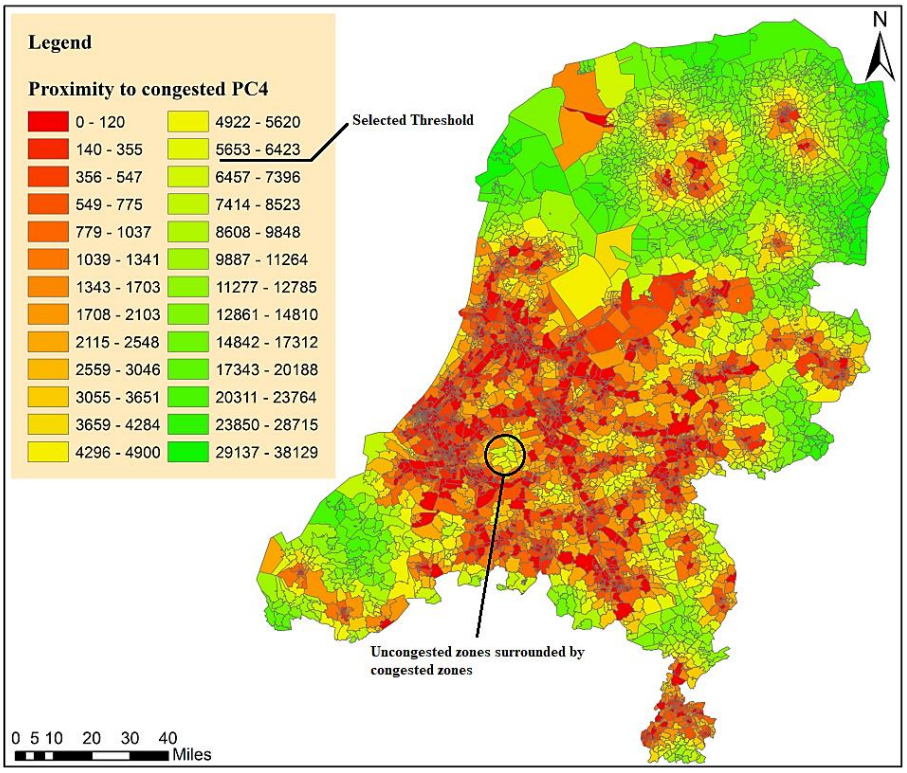


Figure 2-3: Zone clusters based on their proximity to congested zones (PC4 - postcodes)

Transport markets and their attributes

In this section, we explain how we clustered transport markets into classes of homogeneous patterns, based on their tour activities. This leads to more accurate further modelling where specific patterns can be found inside each of these submarkets.

We use the commodity pickup analysis proposed in section 2.3.3 to identify carriers that work with multiple sectors. The minimum size of rules is set to 2 to exclude rules that identify carriers working with only one sector. Besides all these carriers that work with only one sector category (i.e. one NST-2007 categories), we find 57 important rules in the diary of carriers that work with multiple sectors based on the minimum support and confidence thresholds. Figure 2-4 shows all these 57 rules with their support, confidence, and lift. Although all these rules are considered as important, those with higher *Lift*, *Confidence*, and *Support* are the most important and certain rules.

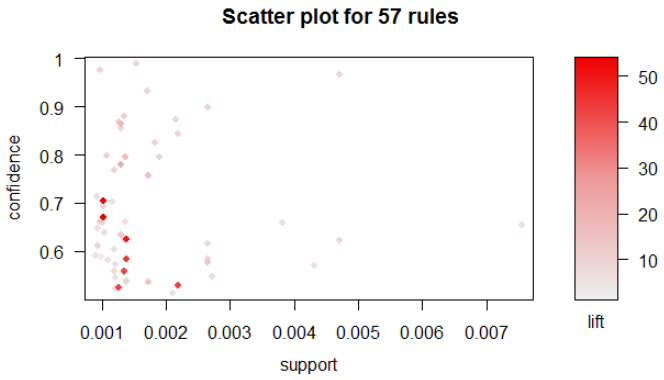


Figure 2-4: Specification of extracted rules from the diary of carriers

Figure 2-5 provides a graphical representation of the extracted rules with more than 0.7 confidence. There are a significant number of carriers that are active in both the food and agricultural industries. Although wood, rubber, grouped products, and fabricated metals are in different NST-R categories, we can see from the commodity pick-up analysis that these products have a high chance to be planned within the same tour of some carriers. This is because sometimes the raw material of industry belongs to a specific good category while the end product of it belongs to another class of goods. The extracted rules clearly show that carriers who are used to deliver raw materials to a firm also pick up the end products to deliver to a consumer or distribution center. The same holds for chemical and petroleum products. For example, the basic organic chemical products are the raw material for pharmaceutical industries and we can see from Figure 2-5 that there is a high chance that these carriers who pick up basic organic chemical products also pick up pharmaceuticals.

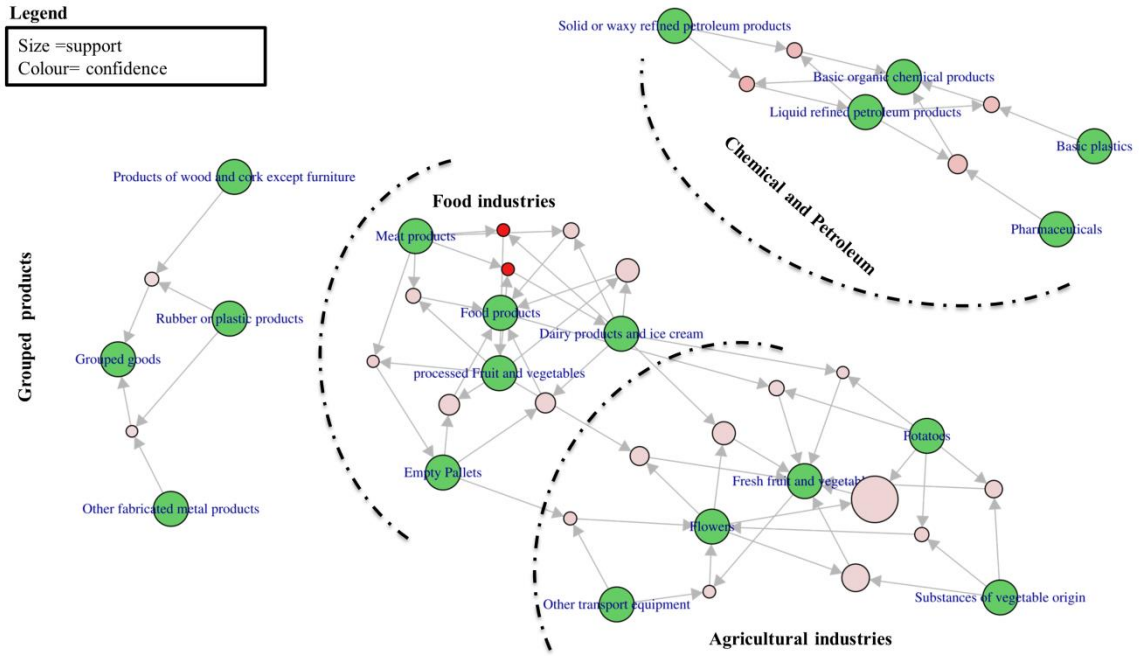


Figure 2-5: Clusters of transport markets with inter-cluster and intra-cluster interactions

In Figure 2-5, we can see the connections between Agricultural and Food products. This indicates how carriers connect industries between these two clusters as well as industries inside each of them. An example of the inter-cluster connection is substances of vegetable origins (e.g. seeds, stems, and roots) which are raw materials for potatoes, fresh fruit, vegetable, and flower industries. An example of an intra-cluster connection is the fact that fresh fruit and vegetables are resources for processed fruit/vegetables and food industries.

We can use the results of this algorithm to explore the probability of occurrence for all significant rules that can explain activities regarding one selected good category. For instance, Figure 2-6 shows the industries that produce trips for empty pallets on their tours.

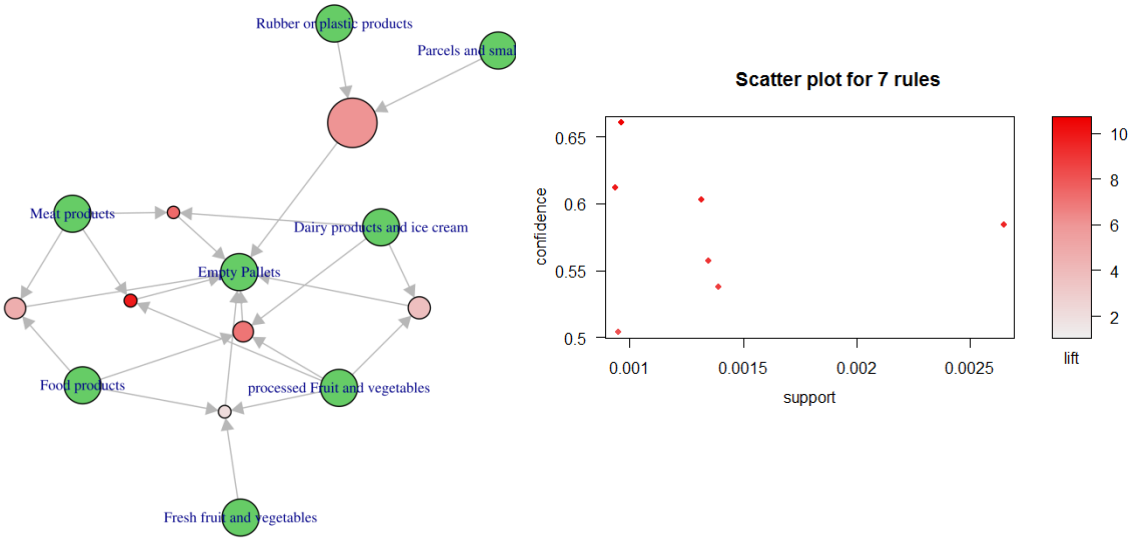


Figure 2-6: Industries with a high probability to produce empty trips

Trips with empty pallets are an important part of tours. Some industries avoid having such empty trips on their tour while it may be indispensable for other industries. From a transportation management point of view, it is important to understand which industries have a higher chance to produce such trips. Figure 2-6 shows 7 significant rules that indicate sectors with a higher chance of having empty trips in their tours. For instance, trips with empty pallets have a higher chance to occur in transporting parcels and rubber or plastic products. The same holds for meat products, food industries, and processed and fresh fruits and vegetables.

As we can see from the above analysis, analyzing the structure of tours requires an understanding of the interaction between multiple sectors. Even carriers that work with one category of goods type, may have interaction between multiple industries within that category. Having extracted these interactions from the tour database, we can now cluster the transport market into the 9 most frequently occurring transport categories in the Netherlands. For each of these, we can then analyze the structure of tours of carriers. Table 2-1 shows the frequency of each of these submarkets in the database after data cleaning and pre-processing.

Table 2-1: Number of observations in each transport market.

Transport markets	Number of observation
Total number of tours	16891
Tours with fresh fruit and vegetables	3304
Tours with Flowers and live plants	772
Tours with Agricultural Products	3768
Tours with Food industries	1172
Tours with Agricultural-food	1290
Tours with Chemical and petroleum	1977
Tours with Construction materials	1038
Tours with Parcels & small packages	1693
Tours with Miscellaneous goods	1877

2.4.2 Explanatory variables

Table 2-2 indicates the sets of explanatory and response variables. They are classified into five categories including temporal variables, spatial variables, shipment characteristics, vehicle characteristics and operational attributes.

Table 2-2: List of variables and their description

Variable	Type of variable	Description
Temporal variables		
Day of week	Explanatory	0:Monday, 1:Tuesday, 2: Wednesday, 3: Thursday, 4: Friday, 5: Saturday, 6: Sunday
Departure time	Response	1 : Morning (6:13-10:38), 2: Midday [10:38-14:53), 3:Afternoon [14:53-19:32), 4: Night [19:32-6:13)
Spatial variables		
Visit distribution centers	Explanatory	1: if any distribution center is visited within a tour, 0: otherwise
Visit transshipment terminals	Explanatory	1: if any transshipment terminal is visited within a tour, 0: otherwise
Congestion status of first intermediate stop	Explanatory	1: if the visited location is marked as congested, 0: otherwise
Congestion status of other intermediate stops	Explanatory	1: if the visited location is marked as congested, 0: otherwise

Table 2-2 (continued)

Shipment characteristics		
Weight factor of first loading location	Explanatory	(0,1]
The number of commodities	Explanatory	Min=1, max=825
Empty container/pallets	Explanatory	1: if the tour includes trips with empty pallets or empty container
Vehicle characteristics		
Vehicle Type	Explanatory	0: truck, 1 Trailer
Trip/tour operational attributes		
The average tour length	Explanatory	Min= 0, max ≈150
Type of Tour	Response	Direct, collection, distribution
Number of stops	Response	Min=1 Max= 33

The temporal attributes include the day of the week as an explanatory variable and departure time of tours as a response variable. The departure time of the tour is a continuous variable reported in minutes (0 to 1440). To keep the model tractable and reasonably simple, we categorized this variable into four clusters. We used the distribution of observed departure times to systematically come up with these categories. For this, we used a density-based clustering technique that converts a continuous variable into factors using frequency and/or density of the variables (for more information, see the discretize function in the arule package in R programming language (Hahsler et al., 2021)). This way, the categories follow the observed pattern in the distribution of the data. This method suggested five categories:

1- [0,373) minutes which is equal to [0:00, 6:13), labelled as ***Night***
2- [373,638) minutes which is equivalent to [6:13,10:38), labelled as ***Morning***
3- [638,893) minutes which is equivalent to [10:38,14:53) ,labelled as ***Midday***
4- [893,1172) minutes which is equivalent to [14:53,19:32), labelled as ***Afternoon***,
5- [1172,1439] minutes which is equivalent to [19:32,23:59) labelled as ***Night***.

For better presentation, we label these categorized times of the day departure time as Night, Morning, Midday, and Afternoon. Note that, we combined the first and the fifth category as we have very few observations in these categories and they both belong to the night or early morning transport when it is outside of the regular working hours.

The spatial variables relate to the characteristics and land use of the visited location in a tour. This includes the activity type of the visited firm and the congestion state of the visited zone where a visited firm is located in. T the visited firm could have three types of activity i.e. transshipment terminal, distribution centre, or a producer/consumer of the goods. Dummy variables are created from the activity types of the visited firms. The state of the congestion for the first and intermediate visiting locations are also coded as binary variables. In section 2.3.2, we explained how the congestion state of each zone is calculated for each hour of the day. To make use of this variable, we calculated the arrival time to each visited zone based on the observed departure time of the tour assuming that trucks took the shortest path between two pairs of zones. The status of each visited zone is considered as 1 if the zone is congested at the time of arrival and 0 otherwise.

The shipment characteristics relate to the size of the shipment. Shipment size can be explained by three aspects i.e. dimensions, weight, and the number of commodities. Among these three, the dimensions of the commodities are not observable to us from the available data. Therefore, we only used the last two for our analysis. We normalized the weight of shipments based on the maximum weight of similar commodities transported in a market.

$$WF = \frac{w_{\max} - w}{w_{\max}} \tag{2.19}$$

This is a factor between (0,1] where the shipments with a weight factor (WF) closer to 0 are considered heavy weighted. This variable only accounts for the shipments loaded in the first loading location in a tour. Note that we are not modelling the trip sequences in a tour in this study. Therefore, the weight of shipment in all the pickup and delivery locations cannot be a factor in this analysis. Amongst all these visiting locations, however, the weight factor of the shipment in the first loading location can play a significant role in identifying the type of tours and the number of stops.

The number of commodities varies between 1 and 825 depending on different transport markets. We also considered shipments that are aimed at transporting empty containers or pallets to depots. This variable indicates whether or not such shipment is included in a tour.

We considered vehicle type as a candidate explanatory variable for vehicle characteristics. Trucks and trailers are two reported vehicle types in the tour data. We created a dummy variable based on the type of vehicle.

Tour/trip operational attributes relate to the operational characteristics of the trucking companies. We define one explanatory and two response variables in this category. The explanatory variable is the average tour length and the response variables are the type of tours and number of stops respectively.

Shipments represent the flow of goods between loading and unloading locations. In practice, shipments are transported with commercial vehicles in an optimum way. This makes vehicle operations more complex as compared to direct shipments between loading and unloading locations. Vehicle operations consist of trips and tours. Trips are the movement of a commercial vehicle between two logistic nodes and a tour is a sequence of multiple trips. In previous research (Ruan et al., 2012, Alho et al., 2019), tours were classified into various kind including (1) direct; (2) distribution, and (3) collection tours. Figure 2-7 shows the relationships between shipments and vehicle flows.

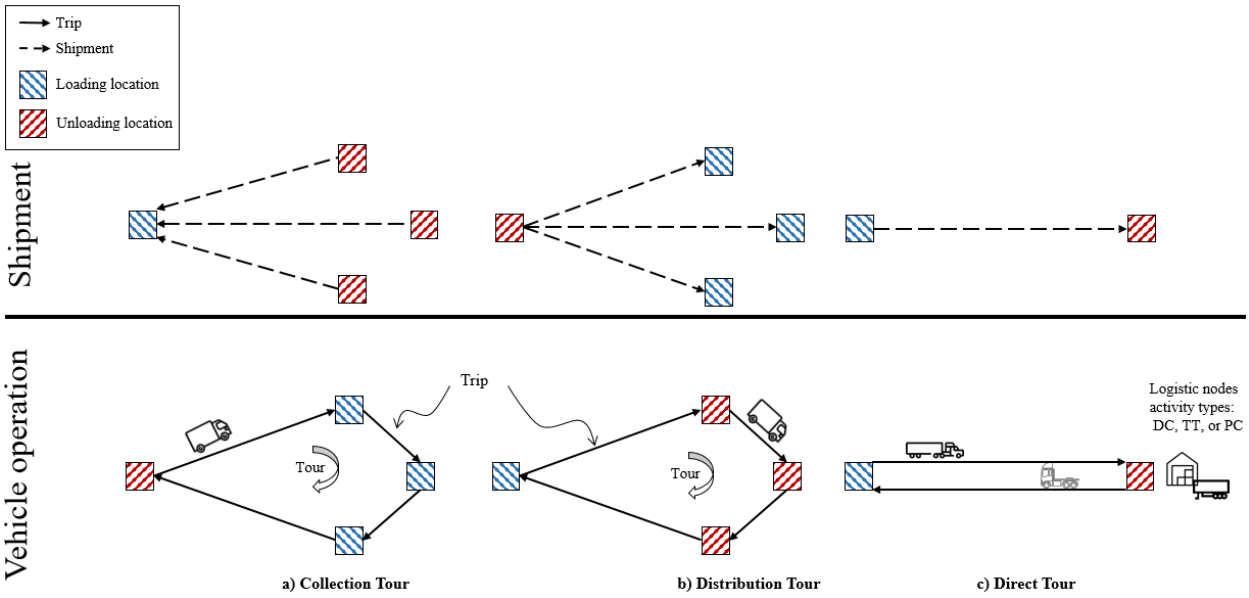


Figure 2-7: Shipments and vehicle operations for 3 types of tour

Tours are of different sizes and their size can be related to the number of stops/visits that occur in a tour. The number of stops varies between 1 and 33 depending on the transport market. The average tour length is a variable that captures the relative distance of a group of consumers to the start location of a tour. This indicates how far or close a set of customers are to the carrier. This variable varies depending on different tour type categories, different trucking companies, and different transport markets and distinguishes between local and long haul transport patterns. Please note that this variable is not the average trip length for an observed tour since this quantity is not clear to the planners before planning the tour. Instead, this is a simple inference-based location of the costumers that indicates the relative spatial relationship between each carrier and its customers. For example, some companies have longer average tour length in each of the tour type categories and some other companies serve more local customers.

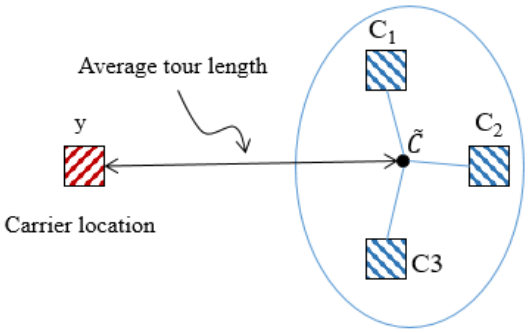


Figure 2-8: Presentation of the average tour length

Tour length is the distance between carrier and a hypothetical costumer $\tilde{C}$ representative of all the customers that have to be visited on a tour. The $\tilde{C}$ is located in the geometric median of the customers. That is a place that has a minimum sum of the distance to all the customers (see Figure 2-8).

$$\arg\min_{\tilde{c}} \sum_{i=1}^{m} \| c_i - \tilde{c} \|_2 \tag{2.20}$$

$$Average\,tour\,length = \| \tilde{c} - y \|_2 \tag{2.21}$$

Given the above variable descriptions, we propose a method to analyze the structure of tours for each of these submarkets in the next section.

2.4.3 Extracted knowledge from tour structures

In this section, first, we evaluate the performance of two decision tree models, developed to study the structure of tours in freight activity of the 9 freight transport markets segmented as described above. Then, we discuss the rules identified from these two models. The first model explains the patterns in departure times of tours, while the other model provides an understanding of the identification of the type of tours. Additionally, this latter model simultaneously predicts the average number of stops per tour type using the MTDT approach proposed in section 2.3.3.

In this study, we used a random sample of 80% of the data to estimate the models and the rest to test the (predictive) performance of the models. For the training set, we used 10-fold cross-validation with 10% validation set to control the depth of the trees as well as the bias-variance trade-off. Table 2-3 provides the performance and accuracy of the selected models with the global commodity type as the root node. The metrics are the average of 100 runs of the fitted trees on the test data.

As we can see from this table, Cohen's kappa statistics are within the substantial agreement range (0.6-0.8) for both models. The lower Macro-F1 than Micro-F1 scores suggest that the class distributions are slightly imbalanced for both models. This is a common problem for time-of-day models since the departure time of vehicles are not equally distributed over different time slots. However, we used Synthetic Minority Over-sampling Technique (Chawla et al., 2002) to reduce this effect as much as possible. With this technique, a synthetic tour departure time is generated for the minority class by randomly drawing from the existing tours and perturbing the attributes towards the average of the class attributes. The one-vs-all average accuracy (72 %) for the time-of-day model is relatively acceptable according to our application domain. The weighted random guess (WRG) accuracy indicates that both models perform far better than a model with predictions by chance. The results show the goodness-of-fit ρ=0.6 for time-of-day and ρ=0.78 for the type-of-tour model. It implies that the type of tour model is easier to predict as compared to the time-of-day model. Both models suggest a relatively high improvement in goodness-of-fit compared to that of the root models.

Table 2-3:Models performance and accuracy for the out-of-sample validation set

	Macro-F1	**Micro-F1**	**One-vs-all**	**kappa**	**WRG**	**ρ**	**P_{root}**	**ρ_{incr}**	**R2**
Time-of-day model	0.57	0.65	0.72	0.79	0.28	0.60	0.23	0.48	-
Type-of-tour model	0.73	0.81	0.92	0.68	0.41	0.78	0.55	0.51	0.71

Among the extracted rules with the two decision tree models, we can seek answers to questions like 1) how do external conditions affect peak-hour tour occurrence? 2) how do different logistics activities influence departure time and tour structure? 3) how do transport costs determine the structure of tours? 4) how does the number of commodities affects the scheduling of tours? and 5) how does vehicle type affect routing and scheduling of shipments? In order to demonstrate how these questions can be addressed, we first discuss one example in detail: food industries and agricultural products. Next, the results for other transport markets are presented.

Agricultural freight consists of products from farmlands to shops, markets and exports to other countries. Such freight encompasses a wide range of perishable goods such as vegetables, fruits, potatoes, tomatoes, cereals, livestock, raw milk, honey, and fish. The transportation of agricultural goods often faces challenges as it requires tight time schedules, high capacity and specialized equipment for loading, unloading and keeping products fresh. On the other hand, the quality, quantity and safety of food products rely on the efficiency of the food distribution system. This requires specialization not only in processing but also in transporting food products. For example, the food transportation market has adopted its functionality of keeping food fresh for an extended period through refrigerated containers.

Agricultural farming and food industries are part of the food supply chain system which ranges from seeds, fertilizers and pesticides, farming, processing, and distribution to satisfy the demand for food. Therefore, understanding the significance of food supply chains requires comprehensive modelling since they include much more than a simple transport consideration, and involve coordinated activities. We focus here on movements from agricultural farming to the food industry. We build DT models to understand the main time-of-day patterns and tour patterns independently. Figure 2-9 and Figure 2-10 present the trees. Each DT model is composed of nodes, which are numbered from the root, e.g. the top-node 1 in Figure 2-9, via internal nodes to the so-called leaf nodes (the bottom row of nodes: 3, 5, 8, 9, etc.). Every leaf node consists of the probability distribution of target variables. This tree reveals the rules related to congestion, logistics nodes, transport costs, number of commodities per tour and empty trips, and type of vehicle used. As such it helps to answer our main questions. We discuss these in the remainder of this section.

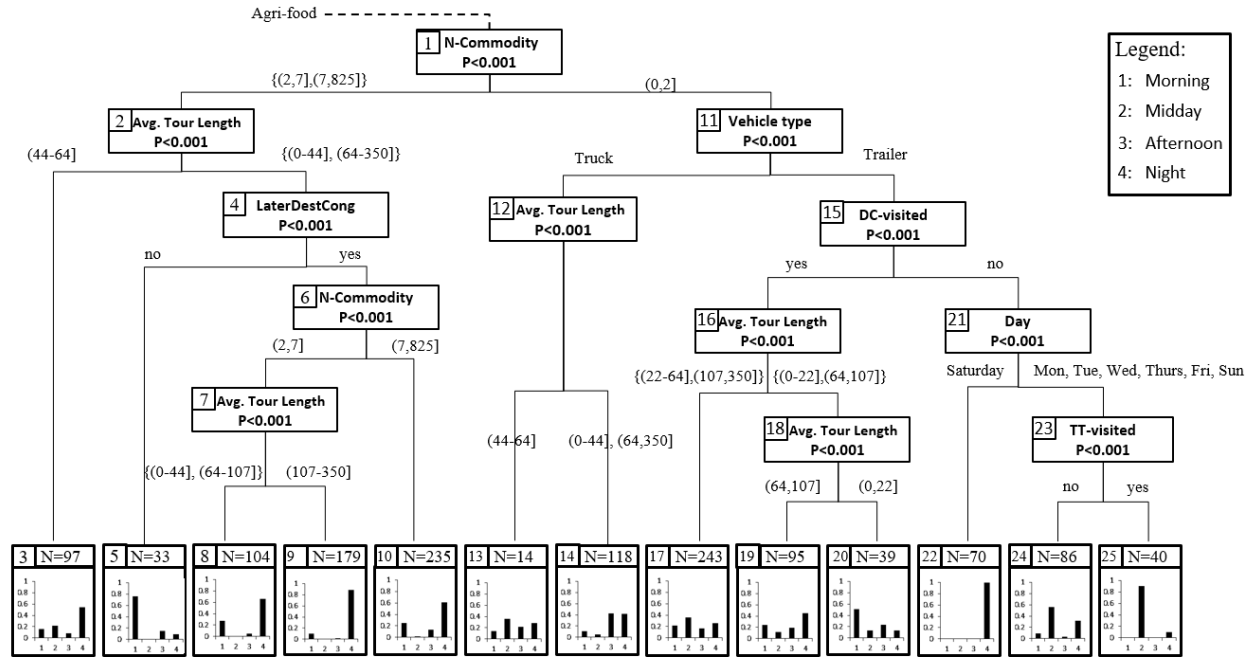


Figure 2-9: Estimated tree structure for time-of-day modelling of Agricultural-Food transport.

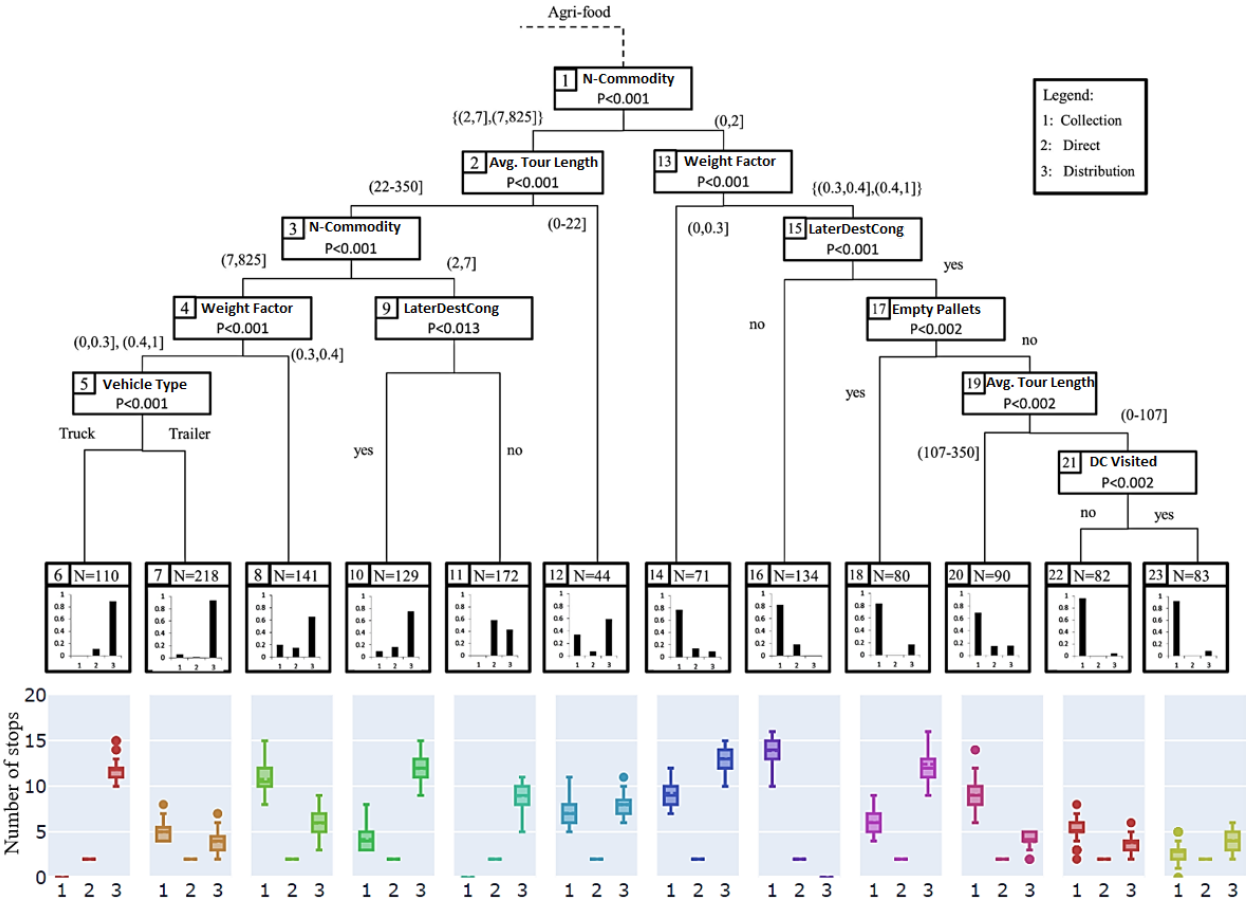


Figure 2-10:Estimated tree structure for type-of-tour modelling of Agricultural-Food transport.

Spatial variables: Congestion related Rules

To comprehend the time of day strategies in this market, we look at nodes related to the congestion indicator of the first and later intermediate stops. Both models indicate that the congestion state of the first pickup or delivery locations does not significantly change the departure time of tours in this market. Significant rules that identify the structure of tours facing later congested zones in their tour are:

- Comparing leaf node 5 with leaf nodes 8, 9, and 10 in Figure 2-9 indicates that the chance of scheduling a tour before morning peak hours (before 6:00 AM) increases if visiting pick-up or delivery locations are located in congested zones.
- In general, planners in this market schedule direct or collection tours (depending on the number of commodities) more often if there is no congestion (nodes 11 and 16, Figure 2-9). However, customers are served in distribution-type tours if they are located in congested zones (node 10, Figure 2-9).

Spatial variables: Logistic Nodes specific rules

In this study, we considered three types of logistics nodes that are often visited within scheduled tours: distribution centers (DC), transshipment terminals (TT), and producers/consumers. From model 1, we obtain two rules regarding the departure time of tours visiting distribution centers or terminals (Figure 2-9, nodes: 17, 19, 20, 25).

- If DC is nearby (less than 22 km) →Tours have a higher chance to depart during the morning peak
- If DC is far away (more than 22 km) → Tours have a higher chance to depart off-peak (Midday or Night)
- If TT visited during a tour → Tours have a higher probability to depart off-peaks (Midday or Night)

The second model does not show any significant rule for transshipment terminals. This is because of the diversity in types of tours for the export and import of agricultural-food products. It does identify tours visiting DC as ones that are usually planned in a collection tour, with 3 stops on average.

Shipment characteristics

One of the most important features to predict time-of-day and type of tour activities is the shipment characteristics. In this study, shipment characteristics are translated into the number of commodities, weight factor, and transportation of empty pallets/containers. The following rules are obtained:

- From Figure 2-9, nodes 1 and 6, we can see that planners generally plan tours with more commodities in the off-peak periods. One possible explanation for this rule is that the loading time is higher in this case and they must be scheduled in a way to avoid the peak periods and arrive on time. on the contrary, the chance of peak period departure time increases if the number of commodities decreases.
- Tours between food manufacturers and agricultural producers are often planned in either distribution or collection types. However, the probability of planning direct tours between them slightly increases for loads of lighter weights (comparing nodes 6 and 7 with 8 and 14 with 16). We elaborated on this rule further in the discussion section.

- Despite the time-of-day model, the type-of-tour model suggests that transporting empty pallets in a planned tour is more likely to take place in a collection tour with a relatively low number of stops (7 stops on average).

Vehicle characteristics

The type of vehicle is also one of the significant explanatory variables in both models. We have two kinds of vehicles, i.e. truck and trailer, in our sample data. The following rules are obtained from the models.

Mode 1 (Figure 2-9):

- Although tours are often planned in off-peak, trucks still have a high chance of starting a tour during the Afternoon peak depending on the Tour length.
- However, trailer combinations usually start a tour off-peak either in Midday or Nights when there is no congestion. Planners also may plan a tour with a trailer in the morning period if the average tour length is relatively low (< 22 km).

Model 2 also indicates that planned tours with trailers have fewer stops (2 on average) than trucks (9 on average) in distribution tours.

Tour/Trip operational attributes

In this study, we used the average trip length as a measurement to capture the relative adjutancy of customers to each other and the carrier. This variable distinguishes between local and long-distance transport and its impact on the time of day and type of tour patterns. We obtain the following rules to address question 3.

- In general, Figure 2-9 shows that tours with long average trip lengths are more likely to depart at night. Carriers should deliver commodities during working hours. Therefore, they commence the tour at night/early morning to both avoid congestion and arrive during working hours.
- We can see from Figure 2-10– nodes 20, 22, and 23 that the shorter the average trip length, the shorter the number of stops in the collection type of tour. In other words, planners may serve fewer local customers in one tour in a short distance and more customers over long distances.

All the transport markets

Similar to this analysis for transport movements between the agricultural and food industries, we have applied the models to extract knowledge for all 9 transport markets. Although decision trees are self-explanatory and can be easily understood following the paths from the root of the tree to leaf nodes, it is of interest to evaluate the quantitative effect of each of the covariates on the response variables. This helps the researcher to get a more global understanding of the model's parameters. As rule-based systems such as decision trees are non-parametric we cannot directly capture such quantitative effects. To cope with this problem, we utilized the measures proposed by Arentze and Timmermans (2003) and used in many applications such as Kim et al. (2018). These indicators quantify the magnitude and the direction of the impact of the covariates in the predicted response variable. These measures are calculated using a confusion table derived from the prediction results of each of the models. We calculate the magnitude of the impact (MI) of the covariates on the response variables as follows:

$$MI_{vi} = g(f_{vi}, f_{vi}^{'}) \tag{2.22}$$

Where g is a function that measures the distance (e.g. Chi-square, likelihood, etc.) between two tables of quantities. We used Chi-square statistics in this study for the function g. f_{vi} is the predicted frequency table for covariate v on response variable i and the f'_{vi} is the expected frequency table assuming that the covariates v has no impact on the response variable. The overall impact of the covariate v on the response variable is the sum of the impacts across all the alternatives in the response variable.

$$MI_v = \sum_i MI_{vi} \tag{2.23}$$

Besides the magnitude of the impact, the direction of the impact is also important for the interpretation of the models. The measure for the direction of the impact is as,

$$DI_{vi} = \frac{\sum_{j=2}^{J} (f_{i,j} - f_{i,j-1})}{\sum_{j=2}^{J} |f_{i,j} - f_{i,j-1}|} \tag{2.24}$$

In this formulation the $f_{i,j}$ is the frequency of the predicted alternative i in the response variable given the j[th] alternative in covariate v. This measurement can take any values between -1 and 1. If one covariate has a monotonically increasing impact, the DI_{vi} equals 1. It would be -1 provided that the covariate v has a monotonically decreasing impact. The DI_{vi} between -1 and 1 indicates that the impact of the covariate on v on the alternative i in response variable is non-monotonous regarding its positive or negative sign. This measurement, however, has no meaningful interpretation for covariates that are unordered (such as weekdays (S. Kim et al., 2017). We calculated the magnitude and direction of the impact for all the significant covariates derived from time-of-day and type-of-tour models. These measures are normalized across all covariates and the complete results are presented in Tables B-1 and B-2 in Appendix B. These results indicate that the principles behind the structure of tours in freight activity can largely be understood under the contexts of time of day and type of tour and number of stops. The most salient findings however are discussed here. From Figure 2-11, we select a number of the most noticeable findings for further discussion. weight factor turns out to be irrelevant in time-of-day patterns for all sectors. Whereas it is an important factor with a relatively high magnitude of impact on type-of-day patterns in all the sectors except flowers and construction materials. The congestion level of the visiting locations is among the top three most important variables that have a high impact on both the time of day and type of tour activities.

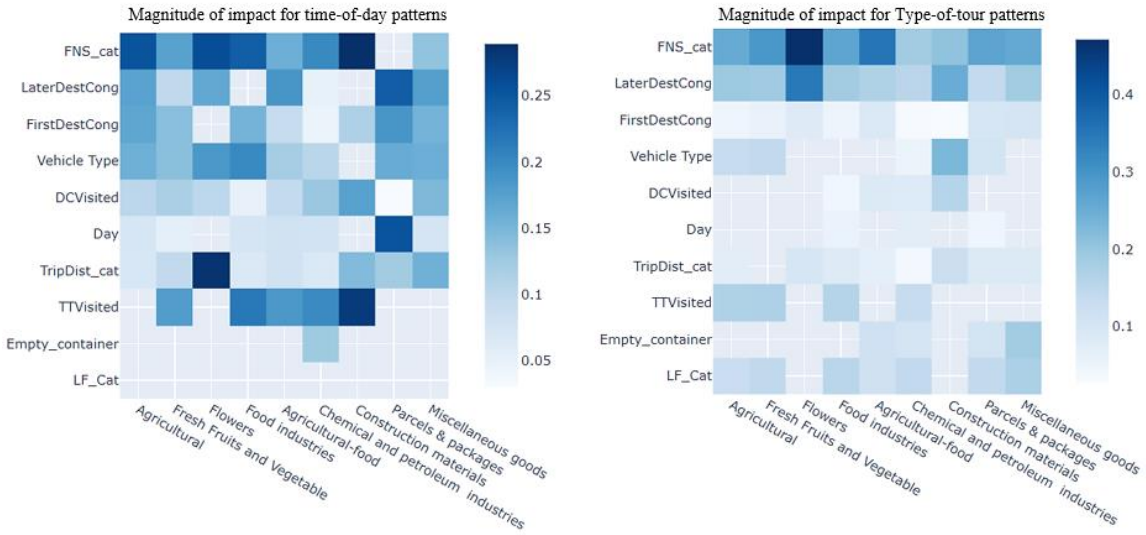


Figure 2-11:The overall magnitude of the impact (MI_v) of the covariates on the response variables across all sectors.

Regarding the direction of the impacts on time of day, the sign of impact (DI_{vi}), for the Agricultural, construction, and food industry transport markets (see Figure 2-12(a)), suggests a negative impact of visiting a congested zone on departure time slot (i.e. DI1: Morning) while the impact is positive for the night/early morning departure time (i.e. DI_4). On contrary, the signs of impact for parcels and package markets are all positive with a high magnitude for the morning and afternoon peak periods. That is the chance of departure is high for these periods even if the locations are congested. This is mainly because the parcels and packages should often be delivered before or after working hours of households which matches the time slots DI_1: Morning and DI_3: Afternoon. Regarding the time of day patterns, visiting transshipment terminals (TT) is also one of the most important variables (see Figure 2-11).

Figure 2-12(b) compares five different transport markets concerning visiting TT. Visiting TT has a negative impact on all the time slots for food and fresh fruit and vegetables with a larger impact on morning and night departures. This implies that there is a higher chance for departure at Midday and Afternoon for these industries. However, For the chemical and petroleum, and construction materials transport markets, visiting TT has a positive impact on the time slot Afternoon.

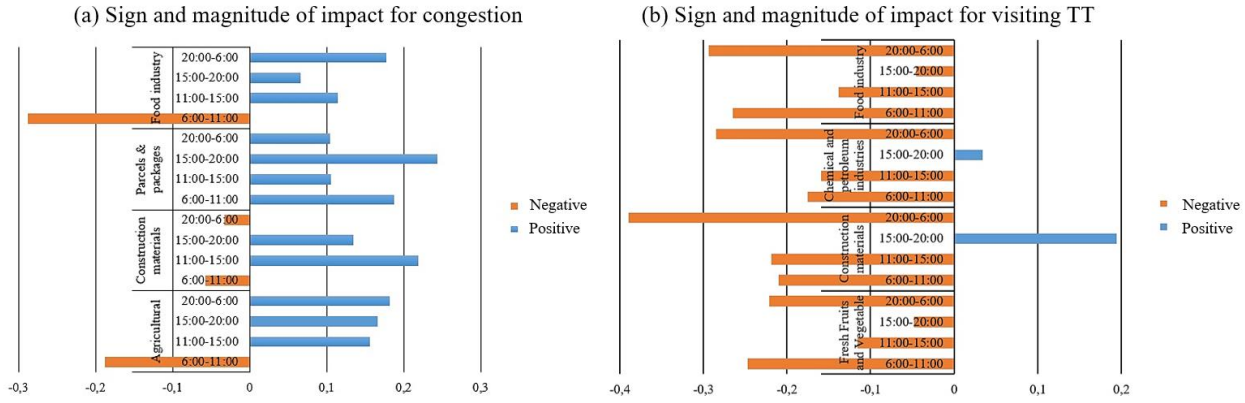


Figure 2-12: Time-of-day interpretation of sign and magnitude of impact for congestion and visiting transshipment terminals

Regarding the impact of congestion on type-of-tour strategies, we can see from Table B-2 (Appendix B) that in all markets, facing a congested location has a negative impact on direct tours. This can imply that carriers often avoid planning direct tours when their customers are in congested zones. In sum, these signs and magnitudes are in line with the findings from the structure of the decision trees and our expectations. We summarize all these findings in Table 2-4. It provides a catalog of results for the different commodities (table rows), by summarizing the dependence structure of tours for the various contextual factors (columns), that result from the analysis. In the next section, we highlight the overall and most salient findings from the analysis.

Table 2-4: Comparison of the tour characteristics and strategies for the segmented transport markets.

Transport market	Dealing with congestion	Influence of logistic nodes	Relation with tour length	Size of shipments	Vehicle types
Fresh Fruit and vegetables	• Planners avoid planning tours in the morning if the zone is congested. • Tours are often planned off-peak. • Facing congested zones→ plan in collection or distribution type of tours.	If DC visited, tours are often planned in the Afternoon. If TT visited type of tour is often direct.	Larger tour length > 64 km → type of tour: direct or collection	• More shipments > 7 the higher chance for distribution type of tours. • lighter weights (higher weight factor) lead to a lower probability of direct tours.	Trailers are less used for collection types of tours.
Flowers and live plants	• Tours depart morning even if the customers are in congested zones. • Congested zones→ collection or distribution type of tour • Uncongested zones → direct type of tours	No significant rule.	• Tour length > 107 → increases the chance for planning tours in early morning/night or daytime off-peak hours(midday). • Tour length < 107 → plan tours in the morning even if customers are in congested zones.	• The more shipments the higher the chance for early morning/ night departure time. • The weight factor is not significant neither in time of day nor in the type of tour patterns.	• Trucks are often scheduled for off-peak hours departure schemes. • Trailers usually depart in the early morning except on Sundays.
Agricultural Products	• congested zones→ Tours depart in the early morning or at night if the average Tour length visiting locations is more than 64 km. • congested zones→ tours with a lower number of commodities usually do not avoid the peak period. • congested zones→ tour type: collection or distribution. • not congested zone→ tour type: direct.	• departure time→ $\begin{cases} \text{Morning}, & \text{Tourdist} \leq 64 \text{ km} \\ \text{Afternoon}, & 64 \text{ km} < \text{Tourdist} \leq 95 \text{ km} \\ \text{Night}, & \text{Tourdist} > 95 \text{ km} \end{cases}$ • Tours visiting transshipment terminals are usually planned in a collection type of tour with four stops on average.	• Long tour distances → night/early morning departure. • Short Tour lengths→ higher the number of stops.	• Tours with more than 2 shipments planned in the early morning or night delivery. • No rule for empty trip or collection of empty pallets • Tour type→ $\begin{cases} distribution, & N_{commodity} > 7 \\ collection, & N_{commodity} \leq 7 \end{cases}$ • lighter weights→ lower the probability of direct tours.	• Trucks: start a tour in the afternoon. • Trailers: usually depart in the morning and also in the afternoon when there is no congestion.
Food industries	• congested zones→ before the morning peak departure. • Congested zones → collection or distribution tours • Uncongested zones → direct tours	• If DC visited → the chance for Morning departure time is higher. • If TT visited → the chance for Midday departure time is higher. • If TT and DC both visited → collection type of tour with a low number of stops.	• Tour length > 44 km → early morning or night departure time. • Tour length < 64 km → the chance for direct tour increases • Tour length > 64 km → distribution type of tour	• The higher the number of shipments the higher the chance for a morning departure time. • lighter weights→ lower the probability of direct tours.	• Planners plan tours for a trailer in the early morning/night departure scheme. Whereas trucks are usually planned between 11:00 and 15:00.
Agricultural-food	• congested zones→ tours depart early morning. • congested zones→ tour type: direct or collection • not congested zone→ tour type: distribution	• DC is nearby →Tours depart during the morning peak. • DC is far away → Tours depart off-peak (midday or night). • TT is visited → Tours depart off-peaks (midday or night). • DC has visited → the collection with 3 stops.	• Long Tour lengths → night/early morning departure. • Short Tour lengths→ lower the number of stops. • Long Tour lengths → more number of stops.	• Tours with more shipments are planned in the off-peak period • Empty pallets are planned in collection type of tours with a lower number of stops	• Trucks: stating a tour during evening peak depending on the Tour length. • Trailers: usually start a tour off-peak. • Trucks: relatively higher number of stops.

Transport market	Dealing with congestion	Influence of logistic nodes	Relation with tour length	Size of shipments	Vehicle types
				• lighter weights → higher the probability of direct tours.	
Chemical and petroleum products	• The probability of departure schedule for morning is high even if the vising locations are congested for the local customers. • Tours are planned in off-peak early morning or night if the customer is in the congested zone and the Tour length is higher than 22 km • Congested zones→ type of tour distribution or collection based on other conditions. • Uncongested zones→ direct	• Planners plan in the afternoon to visit Transshipment terminals with a direct type of tour. • If DC visited → for local customers (distance < 22 km) tours depart in the morning and for Tour lengths > 22 km but less than 107 km tours start at early morning/night time.	• for Tour length > 107 km→ tours starts between 11:00 and 15:00 • longer Tour lengths reduce the number of stops in distribution type of tours	• More than 7 shipments → morning off-peak departure. • If empty container → morning off-peak departure in distribution type of tour with 2 stops on average. • lighter weights→ lower the probability of direct tours.	• Trucks are more likely to depart in the morning or afternoon depending on the congestion level of the visiting zones. • Trailers are more likely to depart between night or early morning. • Trailers: tour type: direct • Trucks: direct tours, but the chance for distribution tours increases
Construction materials	• pick up point is congested → the chance for departure time schedules in Midday increases. • congested zones→ tour type: distribution or collection depending on the number of commodities and vehicle types • not congested → direct tours	• If DC visited →The chance for morning peak departure decreases especially if the visiting locations are congested • If TT visited → planners typically schedule tours at midday. • If DC visited → tour type: direct with a slight increase in the chance of distribution tours.	• Long Tour lengths greater than 107 km→ night/early morning departure. • Short Tour length less than 42 → Morning departure.	• more than 7 shipments → morning off-peak departure. • less than 7 shipments → type of tour: direct • more than 7 shipments → type of tour: distribution • The weight factor is not significant.	• No significant rules for departure time. • Planners use trucks for trips with 14 stops on average whereas trailers are often used for tours with 4 stops on average.
Parcels & small packages	• The chance for morning or evening peak period departure is relatively high even if the visiting locations are in congested zones. • Congested → Type of tour: Delivery or collection • Uncongested → direct tours	If DC visited → Although parcels were transported mostly in the evening peak period, the chance for night delivery increases.	The longer the Tour length → The higher chance for night departure time.	• Empty pallets are collected in tours with 4 stops on average. • The probability of collection tours increases for lighter weights.	• Trailers have a higher chance of departing during the evening peak period as compared to trucks. • Trailers are used for direct and collection tours whereas trucks are used for distribution.
Miscellaneous goods	• Congested → although the chance for early morning departure increases, the majority of tours still depart in morning or evening peak hours. • Congested → type of tour: collection or distribution • Uncongested → type of tour: direct	• If DC should be visited→ the chance for morning and evening peak period departure increases even if the location is congested. • If DC is not visited→ planners often schedule tours in midday or early morning/night period. • No significant rules for the type of tour strategies.	• Long Tour lengths greater than 107 km → night/early morning departure unless DC should be visited. • A short trip with less than 107 km → tours are often scheduled within the morning or evening peak period. • Tours with longer than 64 km → direct tours	• The larger number of shipments (greater than 7) in the first pickup point makes dispatchers plan tours in the early morning/night scheme. • Empty pallets are carried mostly in the distribution type of tours. • The probability for direct tour decreases for lighter weight.	• Despite trucks, the departure time of trailers is usually planned in midday or early morning. • Dispatchers plan distribution tours for trucks whereas direct tours are often planned for trailers.

2.5 Discussion

In order to assess the ability of the proposed methods to extract meaningful and acceptable knowledge about the anatomy of tours, we summarize our main findings below and include an initial validation of outcomes against the existing literature.

The case of analyzing the structure of tours demonstrated that the new method and model proposed is capable of distilling valuable knowledge from large freight tour databases. This is the first main finding from our research. The accuracy of the proposed model is better as compared to similar approaches in the literature. For example, Khan and Machemehl (2017a) reported Rho-Square and adjusted Rho-Square of 0.56 and 0.52 respectively for the type-of-tour model using discrete-continues extreme value modelling. Whereas, the Rho and adjusted Rho for our model is 0.78 and 0.51 respectively. Similarly, for the time of day modelling reported Rho-Square and adjusted Rho-Square are 0.275 (our method reports 0.6) and 0.271 (our method reports 0.48) using the multinomial Logit model (De Jong et al., 2016). One reason for these differences is the homogeneity in tour data made by successfully applying our proposed method for the segmentation of transport markets taking tours and the activity of carriers into account (results are presented in 2.4.1).

The results of analyzing the structure of tours for all transport markets (see Table 2-4) shows that most of the transport markets are sensitive to congestion except carriers transporting miscellaneous goods, parcels and small packages or flowers and live plants. They still plan tours during morning peaks even if the visiting locations are congested. One possible reason for this is less demand flexibility and tight time windows in these industries. Another reason may be that these markets (especially parcels and small packages) often have high demand and hence not enough trucks to serve all demand in off-peak. This finding is in line with Holguín-Veras et al. (2008) who report that carriers transporting food, wood, metal, or construction materials are relatively responsive to off-peak period policies to avoid congestion or extra transport costs.

Interestingly, our result showed that exclusive of the Agricultural-food markets, planners prefer direct tours if they face no congested zone, and they prefer distribution or collection otherwise. This result is intuitive since in most sectors direct tours are more efficient for large shipment sizes which require special facilities and more handling time for loading and unloading. In cases where customers are in congested areas, however, it would be more efficient for carriers to un consolidate shipments into a smaller size to reduce waiting time at congested areas, and instead, bundle multiple customers to maximize their capacity utilization. Nonetheless, for the Agricultural-Food markets, the evidence seems to suggest otherwise. Although Khan and Machemehl (2017a) did not consider the impact of congestion on the type of tour, they reported, in general, that distribution and collection is the most likely tour chain pattern to ship farm products and foods, which aligned well with our findings.

The Chemical and petroleum transport markets are more sensitive to congestion for local customers (e.g. gas stations) than for long-distance trips. This is because the destinations of the long-distance trips are often factories or industries with tight time windows. Carriers avoid peak hours in all industries while visiting transshipment terminals. However, the rules for visiting DCs differ from one industry to another. Visits to DCs in most of the transport markets occur during morning or afternoon peak periods, even if the locations are congested. While some markets, like fresh fruits and vegetables, have tours in the afternoon, other markets, like food industries and agricultural-food markets, plan tours in the morning. A plausible explanation is

that fresh products are moved during the day before they are sold to be available for clients early in the morning, whereas food products can stay in distribution centers for a longer period. There are no freight time-of-day models in the literature that allow a systematic comparison for multiple products. The only relevant study is De Jong et al. (2016) who reported that the period 19:00 to 05:00 (night delivery) is the most preferred delivery period for wholesaler receivers on long distances. This aligns well with our finding (see Table 2-4) that in all transport markets, except Miscellaneous goods, larger Tour lengths to visit distribution centers increase the chance for night/early morning departures.

All the transport markets are sensitive to average tour length in all dimensions (i.e. time, type, and the number of stops). Tours with a higher average tour length are more likely to depart at night or early morning. One possible reason is that carriers have to deliver the shipments to the customers during working hours. Besides, they may want to decrease travel time costs by avoiding peak period transport. The result is that local tours with short Tour lengths often have a larger number of stops. Finally, planners plan direct tours for customers with high total transport costs. As we can observe, direct tours usually have large shipment sizes or volumes. For these, the unit transport costs become low, which drives planners to arrange for a direct shipment.

The weight factor, as a shipment characteristic, is found to be insignificant for the time of the day models across all transport markets and hence is excluded from our analysis. One possible reason for this is the level of aggregation that we took into account for transport markets. Although carriers in each transport market are homogeneous, they transport goods between heterogeneous industries with different replenishment cycles and economic order quantities. In addition, our study is limited to a low resolution for the time of the day to make the model simple and tractable. For this level of resolution, the general pattern of replenishment of receivers overlaps and hence makes it difficult for the model to find a pattern. If the impact of weights on the time of day transport operations is desired, we recommend considering a lower level of resolution both in transport market segmentation and time of the day discretization.

Regardless of the time of the day, the weight factor is significant for the type of tour decisions in all transport segments except flower industries and construction materials (see Figure 2-11 or Table B-2 in Appendix B). For the rest of the transport segments, shipments with light weights decrease the probability of direct transport except in Agricultural-food and parcel markets. For transporting goods between a producer of agricultural products and a food manufacturer, the shipment mainly consists of light-weighted vegetables which have to be directly transported to the manufacturer to keep them fresh. On the other hand, processed food in cans are relatively heavy and has to be distributed to multiple retailers or customers. For the parcel delivery market, lighter weights would lead to a higher probability of collection tours. Postal or delivery companies like DHL usually have several pickup points where packages are received from disaggregate senders. They will plan tours to collect separate shipments with relatively low weights and deliver them to a central parcel depot, where packages will be consolidated together (shipment weight increasing because of the bundling) and after possible transport to a next hub, re-sorted and delivered to the receivers by a distribution type tour. Therefore, this is logical that lighter weights are planned in collection tours whereas heavier weights are transported more in distribution type of tours.

2.6 Conclusion

This paper investigates the anatomy of tours in freight transport and proposes a new modelling approach for understanding the complex process of transporting goods within different sectors of industry. We propose a new enhanced decision tree algorithm (Multi-Task Decision Tree) that predicts multiple discrete and continuous target variables. This algorithm can be used in transport modelling where both discrete and continuous decisions are relevant. This study also has produced new knowledge about the interaction between various markets through the analysis of the tour activities that take place between them. The measurement of these interactions allows for clustering transport markets. We explored the structure of tours in three dimensions i.e. time-of-day, tour type, and the number of stops. Based on these three features, we extracted rules from an extensive tour database to identify existing strategies for the routing and scheduling of freight vehicles, in different transport markets.

The results provide important insights into the preference of transport markets for within-peak or off-peak travel, and for the type of tour applied when facing congested zones. We showed that some markets have strict rules regarding the time of day and type of tour when visiting logistic hubs like distribution centers or transshipment terminals. All transport markets are sensitive to transport costs, as tours are more likely to depart at night or early in the morning to avoid morning peak hours. In addition, tours serving local customers with short Tour lengths often have a larger number of stops. Finally, planners tend to plan direct tours more often for customers with large distances. From a congestion avoidance perspective, planners prefer direct tours if they visit the non-congested zones and rather plan distribution or collection tours otherwise. As could be expected, these rules differ from one market to another. Our paper brings these differences and similarities to the discussion by applying a generic model to different sectors.

This research leads to the following recommendations for future work and applications in practice. From the research perspective, one can explore the application of the MTDT algorithm to model other desired mixtures of discrete and continuous target variables and capture more rules from big transport databases. Examples are the mixture of vehicle type and shipment size preferences, and the tour type and vehicle miles traveled per type of tour. we recommend also further exploration of the relationship between departure time of tours and structure of tours and number of stops since these dimensions are modeled independently in current research. This method gives a set of certain rules distilled from frequent activity patterns of transport markets but does not uncover the underlying reasons behind these rules. This topic could be investigated further through interviews with sectors or experts. Finally, transport policymakers can also benefit from this analysis, to get a better grip on the functioning of transport markets and create more effective strategies for freight demand or traffic management. Examples of such policies are controlling (shifting) the departure time of specific transport markets or changing the tour type pattern of the logistics segments under certain extreme road network conditions. From a methodological perspective, a comparative analysis between MTDT and other nonparametric machine learning methods like GAMs, random forests, and SVM on benchmark problems can provide more insights into the accuracy of the proposed MTDT method. Finally, a robust implementation of the MTDT method in open-source programming languages and statistical tools like python and R could be useful for practitioners and researchers.

Chapter 3

Tour-based representation of freight transport activities

In the previous chapter, we aimed to understand the underlying characteristics of planned tours. As a next step, we develop descriptive models that can regenerate observed tour patterns. In this chapter, we propose a Bayesian optimization method to estimate the parameters of a data-driven vehicle routing and scheduling problem that can be used for simulating freight tour planning. This model seeks to capture the preferences of planners and replicate the tour-based activity of carriers and their tactical decisions. The model can learn and synthesize the most likely tour flows from shipment flow data, total transport cost and time per tour, and travel time information between origins and destinations per time interval. Thus the model allows for descriptive analysis of freight transport and, eventually, policy assessment.

This chapter is based on the following papers:

1. Nadi, A., Yorke-Smith, N., Tavasszy, L., van Lint, H., & Snelder, M. (2022). A Data-Driven Routing and Scheduling Approach for Activity-Based Freight Transport Modelling. Accepted for presentation at TRISTAN XI.
2. Nadi, A., Yorke-Smith, N., Snelder, M., van Lint, J.W.C., Tavasszy, L. (2022). Data-Driven Preference-Based Routing and Scheduling for Activity-Based Freight Transport Modelling - submitted to a journal.

3.1 Introduction

Truck flow patterns can best be understood through the study of freight activities on a transportation network. As opposed to passenger models, freight transport activities have been less researched and hence the literature in this field is relatively limited. There are various reasons for this, among which is that transport modellers often lack observations on firms' activities, whereas disaggregate tour data collection is very expensive. Additionally, modelling freight activities is very complex due to the heterogeneity in the transport markets (Khan and Machemehl, 2017a, Holguin-Veras and Patil, 2007), variety of objectives, its dynamic nature, and the ambiguity in the multi-actor decision-making environment (You et al., 2016, Gonzalez-Calderon and Holguín-Veras, 2019). Researchers have long recognized that there should be distinctions between freight and passenger transport modelling due to the complexity associated with logistic decisions (Tavasszy and De Jong, 2013, Gonzalez-Calderon and Holguín-Veras, 2019). Recently, there has been an increasing interest in descriptive agent-based freight simulation models to study these logistic decisions in the context of freight transport. Examples of this trend are TRABAM (Mommens et al., 2018), MASS-GT (de Bok and Tavasszy, 2018), and SimMobility (Sakai et al., 2020).

Tour planning is a tactical operation that firms undertake to transport commodities from the point of production to the point of consumption. Therefore, it has become a crucial component in all the existing microscopic freight simulation models. Most research to date has tended to rely either on statistical/econometric methods (Hunt and Stefan, 2007, de Bok et al., 2020) or on normative operational research tools (Schröder and Liedtke, 2017, Sakai et al., 2020) to generate tours for simulation purposes.

Models based on econometric methods provide statistical insights on tour formation through sequentially constructing tours based on choice models that predict the next trip destination given the current location of the vehicle until a decision is made to end the tour. The most important drawback with this (trip-based) approach is the ex-ante assumption that is required about the end result of the tours before being able to use the model to generate tours. Examples are the number of tours generated from zones and the number of stops. Recent research like Nuzzolo and Comi (2014), and Thoen et al. (2020) have addressed this issue by introducing a shipment-based tour modelling where the construction of tours is based on assigning shipments to a tour rather than focusing on trips. The choice of shipment assignment is based on a generalized cost that the next shipment adds to the tour. Although the shipment-based model is the most promising tour model so far, such a model is not combinatorial and cannot capture the interdependent routing and scheduling characteristics of tours. That is, the structure of tour and the way tours spatially and temporally form cannot be completely understood with such models.

Normative combinatorial optimization models, on the other hand, are from a similar context to which the tours are planned in reality, and thus are able to find the optimal routing and scheduling of shipment pickups and deliveries considering real-world constraints. Two main drawbacks of such models are their computation time and lack of generalization. The first issue is more manageable due to the recent advances in both hardware technologies and soft computing algorithms. The latter, however, requires a new methodology development that can deal with heterogeneity in objectives, constraints, and preferences of tour planners in various logistic sectors. The main contribution of this research is to bring the advantage of data-driven modelling to the normative discrete optimization models capturing the preferences of planners/or even drivers in vehicle routing and scheduling of shipments. This leads to a more realistic tour model for use in freight simulators.

This paper is structured as follows. In the next section, we view related literature on tour modelling. Then we propose a methodology for a data-driven vehicle routing and scheduling model. Afterwards, the model is applied to the case of freight transport in the Netherlands. We conclude the paper by discussing the findings from the tour activities of various industries.

3.2 Related research on tour modelling

Previous research has developed several tour-based freight transport models. We identify three main classes of tour modelling approaches in the literature. The first class of freight tour modelling is based on the entropy maximization theory in which the most likely tour flow is estimated based on freight trip generation, truck counts on the road network, and total transportation cost or time. In this method, the Lagrangian multiplier associated with zones and transport cost and time can be used to interpret the aggregated tour flow pattern on a network (Sánchez-Díaz et al., 2015). Gonzalez-Calderon and Holguín-Veras (2017) extended this method by adding a heuristic to the model to identify the number and locations of the traffic counts that should be used in the estimation. They also tested the sensitivity of the model to the locations of the traffic counts under different scenarios. Although the entropy maximization method provides valuable insights into general time-dependent tour flow patterns on an aggregate level and bypasses the need for expensive surveys, the method is not able to generate disaggregate tour sequences (Gonzalez-Calderon and Holguín-Veras, 2017).

The second class of tour models relies on the assumption that firms are rational profit maximizers whose behaviour can be predicted based on the theory of utility maximization. These approaches can focus on three units of analysis, either trip, as in (Hunt and Stefan, 2007), shipments as in (Thoen et al., 2020, Nuzzolo and Comi, 2014), or tours (Khan and Machemehl, 2017a). Choice modelling is the basis for all these approaches helping to explain behavioural daily tour patterns.

Of this second class of tour models, trip-based models are often estimated based on disaggregate trip data collected through surveys or GPS. In these models, tours are constructed through incremental trip chaining in such a way that the next destination in a tour is estimated based on the conditional probability of the current stop (Hunt and Stefan, 2007). These types of models are relatively easy to implement and provide transport modellers with descriptive statistics and insights into freight transport systems. However, there are several issues with these types of models. Discrete choice methods are not subject to constraints (like time windows) and therefore hardly can capture spatial-temporal characteristics of tours (Heinitz and Liedtke, 2010). Additionally, trip chaining decisions are made once at the tactical level and hence incremental reconstruction of the tour at the operational level is not the context in which carriers plan tours in reality.

Shipment-based tour models, similar to the trip-based models, stack a set of choice models to reconstruct tours. The difference is that these choice models consider whether or not assigning a shipment to a tour will maximize the utility of the planner based on a generalized cost. The most comprehensive shipment-based tour modelling is developed by Thoen et al. (2020). This model consists of two binary choice models. The first model is the choice of selecting a shipment to be added to a tour; the second model is the ‘end of tour’ choice model that makes sure if the tour should be ended due to duration constraints and or capacity limitations. For reconstructing tours, a nearest neighbour search algorithm is used to find the next stop. The shipment-based architecture allows inclusion of several logistical constraints, such as shipment size, in tour formation. In practice, tours are often the result of an optimization process where

optimal time and sequence of trips are treated interdependently. The shipment-based architecture, although promising, does not capture scheduling decisions and also does not take the combinatorial challenge of deciding the visiting order of locations into account.

In order to deal with the issues mentioned regarding the second class of tour-based freight transport modelling, some researchers have proposed using a family of vehicle routing problems (VRP) and simulation in order to directly study the tours. These VRP formalisms model the pickup and delivery behaviour of carriers within a multi-agent micro-simulation framework (Donnelly, 2009, Van Heerden and Joubert, 2014, Wisetjindawat et al., 2012, Sakai et al., 2020, Siripirote et al., 2020). Although such normative VRP models can perfectly capture space-time constraints, their outcome could deviate from observed tours due to heterogeneity in the tour planning preferences of planners (Canoy and Guns, 2019) or other differences between the stylized model and reality. In some sectors, executed tours deviate from the planned tours because of tacit knowledge of truck drivers about the receivers' conditions and/or externalities that are not recognizable to the planners or are not easy to put in the objective function of VRP models (Mandi et al., 2021).

Only a few studies have touched upon this problem, introducing a new family of VRP which can be called preference-based routing. In this class of VRPs, the objective is to minimize the perceived costs or maximize the utility of planners and/or drivers (Mandi et al., 2021). From a methodological perspective, these studies can be divided into two groups. The first group of studies uses a matrix of transition probabilities between the stops instead of a matrix of distance or travel time in conventional VRPs. Canoy and Guns (2019) propose a maximum likelihood routing problem that maximizes the joint transition probabilities while planning tours. They use Markov chain models to estimate the transition probabilities. Mandi et al. (2021) propose a similar approach but adopt neural networks to estimate transition probabilities. Using neural networks allows adding contextual variables to the prediction of transition probabilities. Both these studies show that this maximum likelihood routing with Markov transition probabilities can learn tours from a set of historical solutions with the reasonable route and arc differences. A problem with these approaches is that they require a large history of complete tour solutions to learn from. Such information may be available in practice for individual firms, where such tools have already been implemented for firms to help planners plan tours. However, generalizable data is hardly accessible for scientific purposes and particularly so for freight simulation. This is mainly because individual firms are not willing to disclose their activities exhaustively due to their customers' privacy.

The second group of studies on preference-based routing in shipment-based tour models employs an inverse optimization approach to calibrate a family of multi-objective VRP models. You et al. (2016) use the method of successive averages to estimate the weights of a weighted sum of multiple objectives from a set of observed tour diaries. The calibrated VRP can be used in any micro-simulation framework to simulate the activity of the freight carriers. Similar to the first group, parameter estimation of these methods mostly requires fully observed truck movement patterns from GPS trajectories (Siripirote et al., 2020). However, tour data, if available, is often privately-owned and only partially available to traffic agencies and policy makers, if at all. As opposed to the maximum likelihood routing approach, the parameter estimation of the inverse optimization approach can be adapted in such a way that it can learn from partially observed tour information just for cases where full information is not available. Until now, the literature lacks an efficient parameter estimation method for these cases.

In summary, the mentioned studies underline the importance of data-driven routing in freight transport modelling. However, existing methods are mostly based on either the data collected from a limited and expensive survey or on fully observed truck movements with trip sequences and schedules, obtained from GPS. In either case, existing work has neglected the role of scheduling of tours. Here, a method to estimate disaggregate dynamic tour planning spatially and temporally based on partially observed tour data will be a requirement. The regular inverse optimization technique as proposed in You et al. (2016) cannot be adopted when the tour information is not fully available. To the best of our knowledge, the development of a method to calibrate a time-dependent VRP model based on shipment flows with partially observed tour data has not yet been explored and is therefore the subject of the current paper. To address this research gap, we propose an efficient parameter estimation method to build a data-driven time-dependent VRP model based on partially-observed tour information.

This study advances the state-of-the-art in literature in three ways:

1- It designs a new freight routing and scheduling model for transport modelling that can accurately link freight tour activities to the truck flows per time of day in between firms or traffic analysis zones.
2- It utilizes partially available tour information, to make sure that the model is as close as possible to reality, by generating the collective tour planning patterns of various carriers.
3- It proposes a new and efficient parameter estimation method that combines the strength of both optimization and machine learning approaches, to provide statistical insights with a certain confidence about tour patterns learned from data.

3.3 Method for data-driven routing and scheduling

Recall that our objective is to infer from historical data the collective pattern of pre-trip routing and scheduling process of freight transport. In the tour planning process, planners often use optimization software that can produce pickup and delivery plans that minimize tour length and travel time. Such optimization problems are known as traveling salesman problems (TSP), for one vehicle, or vehicle routing problems for multiple vehicles.

To explore the potential deviation of planners from TSP tour sequences, we use a large tour database collected by the Central Bureau for Statistics (CBS) in the Netherlands. After data pre-processing, a set of 16,171 tour activities of the 720 vehicles from the largest carriers are selected for this study. We provide more information about this dataset in Section 3.4.

We use the open-source Google OR-Tools solver to solve a benchmark TSP for all the tours in the dataset. Figure 3-1 compares the length of the observed tours with the TSP routes. From this experiment, we identify that tours are on average 18.5% longer than TSP solutions. This reveals that planners or drivers, in many cases, do not prefer the tours with the lowest travel times or costs. the data finds that humans are apparently not rational with respect to the TSP optimization model (Hofstede et al., 2019).

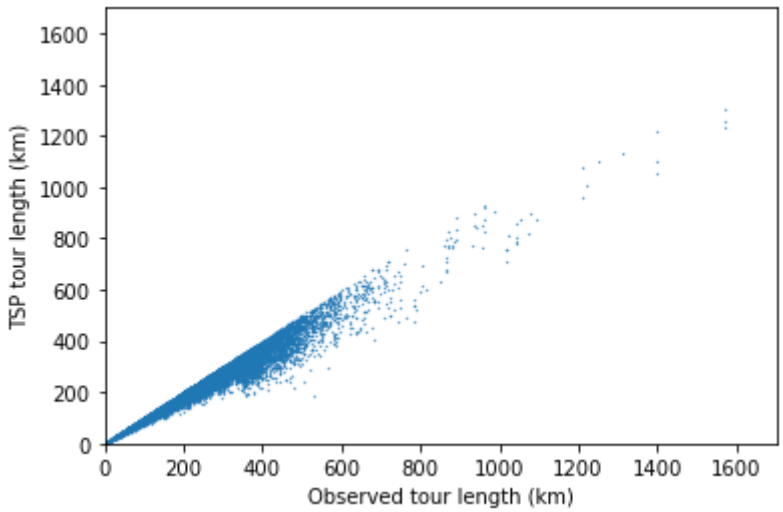


Figure 3-1: Deviations between TSP route and executed tour lengths

In some cases, planners might use other objectives like fuel consumption and/or the cost of carbon emissions in the objective of the optimization problem. Further, our estimate of times and costs may differ from those used by the firms. However, such objectives and other real-world constraints or their perception of time and costs are not observable to us.

To deal with this problem and capture the preferences of planners and drivers in planning tours we propose a data-driven routing and scheduling problem where the preferences of the planners can be identified through a set of features associated with zones (or firms). To achieve this, we include a linear weighted sum of these features in the cost of routing and scheduling decisions. We use the Benders decomposition method (Castellucci et al., 2021) with an adaptive large neighbourhood search (ALNS) algorithm to find the optimal sequence of visited zones/firms, and utilize a Bayesian optimization algorithm to iteratively update the parameters associated with the zonal features. We next give details of the mathematical model and the decomposition-based approach.

3.3.1 Tour modelling formulation with learnable parameters

We formulate tours as a time-dependent VRP with pickup and delivery and capacity constraints. In our data, tours are often starting from the base location of carriers or firms to visit consumers of goods, transshipment terminals, manufacturers, retailers, or distribution centres. For each firm in the area of study, this VRP is defined on a graph $G=\{V, A\}$ where V is the set of all nodes and A is the arcs' set. Vertices of this network consist of carriers' depot $\{0\}$, pick-up points P and delivery points D. Each link is associated with c_{ij}^{m} which is a function of features related to the pickup or delivery points and the link between them in time interval m. The higher the c_{ij}^{m} the lower the inclination of planners to visit j after i in the sequence. The order of visiting locations, however, depends on a combination of the planner's tendency towards all other i and j, and hence can be identified through a combinatorial optimization. Each pick-up point i is also associated with demand quantity $q_i>0$ and its corresponding delivery point $q_{i+m}=-q_i$. For better reference to the indices, parameters, and decision variables, we them as follows:

V	A set of pickup and delivery locations
i,j	Index of pick up or delivery locations $i,j \in V$
K	The set of available trucks at the firm/zone

M	A set of time intervals
cap	Maximum capacity of vehicle *k*
L_i	A load of a vehicle visiting location *i*
x_{ij}^m	If a vehicle travels from *i* to *j* in interval m, then $x_{ij}^m = 1$, otherwise $x_{ij}^m = 0$
β	Vectors of parameters to be estimated for the cost of the links
c_{ij}^m	Cost of the link between *i* and *j* in the interval $m \in$ M (with learnable parameters)
t_{ij}^m	Travel time between *i* and j in the interval m ∈ M
a_i	Arrival time to location *i*
s_i	Loading/ unloading duration in location *i*
q_i	Demand quantity in location *i*
v_i	Index of the first node in the route that visits node $i \in$ V

We adopt a compact form of a standard two-index formulation of the time-dependent pickup and delivery problem from Castellucci et al. (2021) and Furtado et al. (2017). In this paper we are not proposing a new VRP formulation; rather, we propose a method that can be used to make any class of VRP learnable from a set of real-world historical tours. We choose the following formulation because it is aligned with the structure of tours that is observable to us. One can adapt this formulation based on the need and observability of real-world objectives and constraints.

$$\min \sum_{m\in M}\sum_{i\in V}\sum_{j\in V} c_{ij}x_{ij}^m \tag{3.1}$$

subject to :

$$\sum_{m\in M}\sum_{i\in V} x_{ij}^m = 1; \qquad \forall j \in P \cup D \tag{3.2}$$

$$\sum_{m\in M}\sum_{j\in V} x_{ij}^m = 1; \qquad \forall i \in P \cup D \tag{3.3}$$

$$\sum_{i\in V} x_{ij}^m = \sum_{i\in V} x_{ij}^m; \qquad \forall m \in M \tag{3.4}$$

$$\sum_{m\in M}\sum_{j\in V} x_{0j}^m \le K; \tag{3.5}$$

$$a_j \ge a_i + s_i + t_{ij}^m - M(1-x_{ij}^m); \qquad \forall i \in V, j \in V, m \in M \tag{3.6}$$

$$L_j \ge L_i + q_i - M(1-x_{ij}^m); \qquad \forall i \in V, j \in V, m \in M \tag{3.7}$$

$$\max\{0, q_i\} \le L_i \le \min\{cap, cap + q_i\} \tag{3.8}$$

$$a_{i+n} \ge a_i + s_i + t_{i,i+n}^m \qquad \forall m \in M, \forall i \in P \tag{3.9}$$

$$v_{i+n} = v_i \qquad \forall i \in P \tag{3.10}$$

$$v_j \ge j.x_{0j}^m \qquad \forall j \in P \cup D, m \in M \tag{3.11}$$

$$v_j = j.x_{0j}^m - n(x_{0j}^m - 1) \qquad \forall j \in P \cup D, m \in M \tag{3.12}$$

$$v_j \ge v_i - n(x_{ij}^m - 1) \qquad \forall i, j \in P \cup D, m \in M \tag{3.13}$$

$$v_j \le v_i - n(1-x_{ij}^m) \qquad \forall i, j \in P \cup D \tag{3.14}$$

$$a_i + s_i - t(m-1)x_{ij}^m \geq 0 \quad \forall i \in V,\ j \in V \setminus \{0\}, m \in M \tag{3.15}$$

$$a_i + s_i \leq t.m + (1 - x_{ij}^m)(T - t.m) \quad i, j \in V, m \in M \tag{3.16}$$

$$x_{ij}^m \in \{0,1\} \tag{3.17}$$

The objective function presented in Equation (3.1) minimizes the perceived cost of routing and scheduling for planners/drivers. Equations (3.2) and (3.3) make sure that each node is visited exactly once. Equation (3.4) ensures the degree of balance for all nodes. Equation (3.5) represents the constraint for the total number of vehicles. The load and time consistency is guaranteed with constraints presented by Equations (3.6) and (3.7). The inequality constraint proposed in Equation (3.8) ensures the capacity limitation of vehicles. Equation (3.9) guarantees the precedence of pickup and delivery locations, as pick-up points always have to be before delivery points. Equations (3.10) to (3.14) reflect the pairing relation which makes sure that all related pick-up and delivery locations are on the same tour. We defined the lower and upper bound on departure times in equations (3.15) and (3.16) ensuring that departure time from each location is linked with its associated interval *m*. Finally, Equation (3.17) defines the domain of the decision variable. The above formulas represent in general how tours can be planned by firms based on a dataset. We did not include the time windows constraint. As mentioned previously, not all the tour information is available for transport modellers. Real-world constraints like time windows, if available, can add to the accuracy of the model.

To capture the spatial and temporal preferences of tour planners from a set of partially observed tours information, we define c_{ij} (presented in Equation (3.1)) to represent the cost of visiting j after *i* from the planner's perspective. We link this cost to a set of features that relate to the characteristics of visiting zones or firms and also the link between them.

$$c_{ij}^m = \sum_{f=1}^{F} \beta_f^m \chi_{fij}^m + b^m \tag{3.18}$$

In Equation (3.18) β_f^m represents the importance of feature *f* and χ_{fij}^m represents the value of each feature in time interval m from the planners' perspective and b^m is a bias term that captures the aggregate costs that planners may consider but is not observable to us. We introduce the features that can contribute to the planner's preferences. Note that with respect to the *m*, departure time of the tours, these features are alternative specific attributes. This implies that we estimate different parameters for similar features at different tour departure times. These introduced features are as follows:

- χ_{1ij}^m: travel time cost per hour between firm/zone *i* and *j* when the trip starts at time interval *m*. This can be derived by multiplying a fixed cost per hour by the travel time between *i* and *j*.
- χ_{2ij}^m: transport cost per kilometre between firm/zone *i* and *j* when the trip starts at time interval *m*. This can be derived by multiplying a fixed cost per kilometre by the distance between *i* and *j*.
- χ_{3ij}^m: vehicle load while traversing from *i* to *j* and the trip starts at time *m*.
- χ_{4ij}^m: ratio of the number of commodities for being picked up or delivered in *j* over *i* and the trip starts at time *m*.
- χ_{5ij}^m: ratio of the weight of commodities for being picked up or delivered in *j* over *i* when the trip starts at time *m*.

- χ_{6ij}^{m}: transport cost per kilometre between depot and *j* over transport cost per kilometer between the depot and *i* when the trip starts at time *m*.
- χ_{7ij}^{m}: travel time cost per hour between depot and *j* over travel time cost per hour between the depot and *i* when the trip starts at time *m*.
- χ_{8ij}^{m}: transport cost per kilometre between *j* and depot over transport cost per kilometre between the depot and *i* when the trip starts at time *m*.
- χ_{9ij}^{m}: travel time cost per hour between *j* and depot over travel time cost per hour between *i* and depot when the trip starts at time *m*.
- χ_{10ij}^{m}: 1 if *j* is a producer/consumer of goods and *i* is the distribution center and 0 otherwise.
- χ_{11ij}^{m}: 1 if both *i* and *j* are distribution centers and 0 otherwise.
- χ_{12ij}^{m}: 1 if the commodity between any *i* and *j* is of type NSTR 0 (Agricultural products) and 0 otherwise.
- χ_{13ij}^{m}: 1 if the commodity between any *i* and *j* is of type NSTR 1 (Food products) and 0 otherwise.
- χ_{14ij}^{m}: 1 if the commodity between any *i* and *j* is of type NSTR 6 (Construction materials) and 0 otherwise.
- χ_{15ij}^{m}: 1 if the commodity between any *i* and *j* is of type NSTR 7 (Fertilizers) and 0 otherwise.
- χ_{16ij}^{m}: 1 if the commodity between any *i* and *j* is of type NSTR 8 (Chemical) and 0 otherwise.
- χ_{17ij}^{m}: 1 if the commodity between any *i* and *j* is of type NSTR 9 (Machinery and others) and 0 otherwise.

Together, these features help to identify the preferences of planners in the routing and scheduling of different commodities between different logistics firms. Besides feature parameters, there might be some other necessary but unknown parameters in the model which can be estimated from the observed data. In this study, for example, the loading and unloading duration of commodities, s_i, is not reported. We, therefore, assume that the loading and unloading of commodities is a random process that its service time has a cumulative exponential probability distribution with an unknown average service time μ. Hence, the probability that the service time in location *i*, will be less than or equal to *t* is:

$$P(s_i \leq t) = 1 - e^{-\mu t} \tag{3.19}$$

The unknown average service time μ has to be estimated along with other parameters.

3.3.2 Parameter Estimation of the tour model

Having described the VRP model, we now describe how to estimate its parameters. In general, we begin with initializing the parameter of the models with some initial values. Then we use the ALNS search algorithm and Benders decomposition to generate route sequences and departure time schedules for the tours. We iteratively update the parameters of the model minimizing the deviation between generated and observed tour characteristics. This process requires repeatedly solving the VRP model after each update in parameters. Since VRP models are computationally expensive, we cannot easily adopt a complete optimization algorithm to search for a better set of parameters. Alternatives are local search algorithms, perturbation stochastic approximation methods, or gradient descent algorithms – all of which may stop in a local optimum and hence report inefficient results.

To deal with this problem, we adopt from artificial intelligence the model-based or Bayesian optimization technique that is successfully applied in hyper-parameter optimization of black-box functions like deep learning. Model-based algorithms are a class of global optimization methods that can be used to find the minimum of functions that are expensive to evaluate, do not have derivatives available, and can only be measured under noisy environments or simulations. Examples of such algorithms are LineBO (Kirschner et al., 2019), DONE (Bliek et al., 2016), and SMAC (Hutter et al., 2011). These algorithms are mainly based on two principal goals, the first of which is to explore the search space sufficiently for a global minimum and the second of which is to obtain a good solution in as few function evaluations as possible. These methods use a surrogate surface fitted to the parameter space in order to reduce the number of function evaluations while searching for the global optimum. In this paper, we propose the same approach to estimate the parameters of the tour model.

The parameter estimation algorithm has five steps as follows:

Step 1: Prepare instances of observed tours with all the matrices of features related to visiting locations in different time intervals. Since we provide instances of observed tours one at a time to the parameter estimation process, K in Equation 3.5 is set to be equal to 1 during the estimation phase. This converts the tour model from VRP to TSP as we only have one vehicle to visit all the firms/zones presented by each input instance. After parameter estimation, this parameter can be set back to the actual number of vehicles available at each firm, if the application of the model is at the firm level, and to the number of truck trips started from each zone if the application is at the zone level.

Step 2: Generate n initial parameters $\Theta=\{\theta_1,\theta_2,\ldots,\theta_n\}$ for the tour model using Latin hypercube sampling. Each θ_n includes all β parameters, bias terms b, and the μ parameter. we started with 15 initial samples of all parameters.

Step 3: For each parameter setting, we solve the tour model for all instances prepared in step 1. We employ ALNS using simulated annealing to provide approximate solutions for carriers having a large number of customers in the planning horizon, and branch-and-cut to provide the exact solution for carriers having fewer customers. Since our model is time-dependent and the tour-solving algorithm has to find actual departure time from each node considering traffic conditions, exact solutions for such problems require longer computation time. To help the algorithm search both time and space dimensions faster, we use a Benders decomposition method as presented in Castellucci et al. (2021). We thus decompose the time-dependent tour model into a master and sub-problems. In the master problem, we calculate the routes with constant travel time throughout the day. Once the optimum solution is found, we use a sub-problem to compute all the arrival and departure times using time-dependent travel time matrices. Then we apply feasibility and optimality cuts inferred from the sub-problem to the master problem. For more information about the method, we refer readers to Castellucci et al. (2021). For the implementation of the tour model, we use open-source Python packages, in particular the state-of-the-art Google OR-Tools.

Step 4: Given the observed tour sequences and generating tours based on initial zone feature weights Θ, we can evaluate the tour model by comparing observed and generated tours. Previous research introduced arc and route deviation metrics to measure similarity or dissimilarity between two sequences. However, we cannot use these metrics as our observed tours do not reveal such information. In our case where the tour information is partially available, we compare available tour information using the following loss function:

$$l(\theta) = \frac{1}{N}\sum_{n=1}^{N}(r_n - \tilde{r}_n)^2 \tag{3.20}$$

where N denotes the number of observed tours, r_n is the observed tour information, and $\tilde{r}_n$ is the tour information generated by the tour model. The observed or generated tour information presented by r_n or $\tilde{r}_n$ are two vectors containing observed and generated tour length, tour duration, tour start time and end time.

Step 5: In this step, we iteratively update the initial parameters θ minimizing the loss function defined in equation (3.20). We adopt the following steps of the Bayesian optimization algorithm to find parameters θ.

1- Surrogate construction: Fit a surrogate surface f to the evaluated Θ samples using a Multivariate Gaussian process with mean μ and covariance Σ (see Equations 3.21-3.23).

$$f(\Theta) \sim N(\mu, \Sigma) \tag{3.21}$$

$$\Sigma = K(\theta, \theta') = \begin{bmatrix} k(\theta_1,\theta_1) & \dots & k(\theta_1,\theta_n) \\ \vdots & \ddots & \vdots \\ k(\theta_n,\theta_1) & \dots & k(\theta_n,\theta_n) \end{bmatrix} \tag{3.22}$$

$$k(\theta, \theta') = \exp\left(\frac{1}{2}\frac{(\theta - \theta')^2}{d^2}\right) \tag{3.23}$$

The kernel function $k(\theta,\theta')$ determines how smooth the surrogate surface could be and the length parameter d scales the correlation between parameters θ and θ'.

2- Search for minimum: Find the minimum of the interpolated surface for the sample of parameters. This gives us the alternative point with the lowest mean sample values among the already evaluated points.

$$\theta_{\min} = \arg\min_{\theta \in \Theta} f(\theta) \tag{3.24}$$

3- Next evaluation point: the next evaluation point θ^* in parameter space is where the expected improvement (EI) measure is maximum. This EI measurement is introduced in Equations 3.25 and 3.26 to balance exploration and exploitation.

$$EI(\theta^*) = E[\max(\bar{f}(\theta_{\min}) - f^*(\theta), 0)] \tag{3.25}$$

$$f^*(\theta) \sim N(\bar{f}(x), \sigma(\theta)) \tag{3.26}$$

Where $\bar{f}(\theta)$ is the mean stochastic prediction at point θ and $\sigma^2(\theta)$ is the estimate of prediction error.

4- Update the surrogate surface: With the new parameter θ^*, we update the covariance matrix and consequently the surrogate surface (see Equations 3.27-3.31). The process iteratively continues until the stopping criteria apply.

$$J(\theta^*) \mid J(\theta_1), \cdots, J(\theta_n) \sim N(\mu(\theta^*), \bar{K}(\theta^*)) \tag{3.27}$$

$$\mu(\theta^*) = k_{\theta^*}(\theta^*)^T k_{\theta^*\theta^*}^{-1} g \tag{3.28}$$

$$g = (J(\theta_1), ..., J(\theta_n))^T \tag{3.29}$$

$$K_{\theta^*}(\theta^*)^T = (k(\theta^*, \theta_1), ..., k(\theta^*, \theta_n)) \tag{3.30}$$

$$\bar{K}(\theta^*) = k(\theta^*, \theta^*) + k_{\theta^*}(\theta^*)^T k_{\theta^*\theta^*}^{-1} k_{\theta^*}(\theta^*) \tag{3.31}$$

3.4 Empirical Validation

Having described the proposed tour model, we now apply it to a real-world set of shipment data and imperfect tour information in the Netherlands. We first describe the databases used and the data preprocessing that was applied, before estimating the parameters of the tour model.

3.4.1 Firms and carriers' tour databases.

For the development of this tour model, we use carriers' tour data collected by the Statistics Netherlands (CBS). This data is available only with the permission of CBS. The data includes over 2.7 million records of shipments across the country for the years 2013-2015. In this data collection, the largest third-party carriers are legally obliged to fill in a form that collects one week of their activities. In most cases, The data has been exported directly from the firm's transportation management system automatically through an XML data structure. The data includes the geographic locations of loading and unloading of shipments, their commodity type, and shipment size. It also provides information about the capacity of the vehicles and is enriched with other data sources (e.g. firm establishment data) to express locations of important activities including distribution centres, transshipment terminals, and producer/consumer of goods. The data provides aggregate information about each tour as well. This information is about the total tour costs and times, shipments that belong to one tour, the first and last visited locations, and the start and end time and location of the tour. Although the data includes very detailed information about the shipments, it does not provide us with the order of trips in a tour. The tour database only includes the total tour duration. However, the intermediate travel time between each visited customer is not provided. To impute these intermediate travel times and distances, we use the calibrated national Dutch regional traffic model. This model helped us to create skim matrixes of the morning, midday, and afternoon travel times between each traffic analysis zones in the Netherlands. For more information about the available data see de Bok and Tavasszy (2018) and Thoen et al. (2020).

After preprocessing the data, we selected a subset of 16171 tour activities (107398 shipments) of the 720 vehicles from the largest carriers are selected for this study. This subset does not include direct tours. We excluded direct tours from our sample. Because such tours do not reveal the routing preferences of the planners. The proposed method after training can still generate

direct tours if the capacity of the vehicle is reached or the VRP cannot bundle a shipment with others due to the costs of other tours. Table 3-1 reports the descriptive statistics of the selected tours.

Table 3-1: Descriptive statistics of tour data

Tours characteristics	Number of tours	Percentage of tours
Number of stops		
3-5	6734	44%
6-10	7378	48%
>10	1183	8%
Tour length (km)		
0-50	649	4%
50-100	1576	10%
100-200	4020	26%
200-400	7170	47%
>400	1880	12%
Tour Duration (hour)		
0-3	1651	11%
4-6	2925	19%
7-8	3171	21%
9-10	3095	20%
>10	4453	29%
Commodity types		
0	1530	10%
1	4457	29%
6	183	1%
7	41	0%
8	98	1%
9	8986	59%
Logistics activities Intermediate Loading locations	Number of shipments	Percentage of shipments
DC	17083	16%
TT	1861	2%
P/C	88454	82%
Intermediate Loading locations		
DC	5128	5%
TT	2277	2%
P/C	99993	93%

3.4.2 Results of the model estimation

We hold out 20% of the tour data for validation, and train the tour model based on the rest of the tour data, using the Bayesian optimization algorithm of section 3.3.2. We draw the index of the training set from a random uniform distribution to avoid any bias on the selected data. We set the number of initial parameter samples in Bayesian optimization to 15. Parameters associated with features are bounded between -10 and 10. The parameter of the service time generator is assumed to be an integer between 20 to 120 minutes. We specify a 0.05 optimality gap as the stopping criteria meaning that if the model does not improve the solutions by at least 5% in 5 consecutive iterations, the algorithm stops updating the parameters. We activate this criterion after 50 iterations. We also set the maximum number of iterations to 100. All the features are normalized between 0 and 1 using a min-max scaler. For tours with 10 or greater stops, the TSP solver automatically switches between exact and heuristic algorithms for solving the TSP models. In our experiment, the number of time intervals that the departure time of tours

may fall in is m=3. This indicate morning peak, afternoon peak, and off-peak periods. For feature selection, we started with features 1 to 6 from the list of features that are introduced in Section 3.3.1. These features can explain planners' preferences intuitively. Then we consecutively add/remove other features to the model only if they improve the final loss function value. Table 3.2 shows the estimated parameters of the tour model.

Table 3-2:Estimated parameters of the model

		Morning peak		Afternoon peak		Off-peak	
No	features	Estimates	std	Estimates	std	Estimates	std
1	χ_{1ij}^{m}	1.12	1.8e-4	2.1	1.52e-05	1.064	2.18e-4
2	χ_{2ij}^{m}	3.52	1.5e-3	4.8	0.014	4.15	2.87e-5
3	χ_{3ij}^{m}	-2.39	1.42e-5	-3.89	1.7e-4	-1.66	1.3e-5
4	χ_{4ij}^{m}	-0.251	1.4e-8	-0.312	3.29e-4	-0.107	1.3e-2
5	χ_{5ij}^{m}	-0.097	1.2e-3	-	-	-0.042	4.38e-2
6	χ_{6ij}^{m}	-2.379	1.4e-8	-3.289	1.3e-2	-1.263	3.29e-9
7	χ_{7ij}^{m}	-1.246	1.5e-6	-1.956	1.7e-4	-0.954	1.3e-5
8	χ_{8ij}^{m}	-2.074	2.1e-2	3.173	3.7e-3	1.063	1.73e-5
9	χ_{9ij}^{m}	-1.331	1.87e-3	1.212	2.7e-3	1.023	3.8e-5
10	χ_{10ij}^{m}	-0.471	6.3e-7	-0.423	3.9e-3	-0.529	2.2e-8
11	χ_{11ij}^{m}	0.516	2.5e-3	-	-	0.294	4,1e-3
12	χ_{12ij}^{m}	-0.170	1.6e-4	-0.289	1.7e-4	0.363	2.4e-6
13	χ_{13ij}^{m}	0.055	2.3e-4	-0.181	1.2e-3	-0.541	3.1e-5
14	χ_{14ij}^{m}	-0.134	1.7e-2	-	-	-	-
15	χ_{15ij}^{m}	-0.041	1.3e-2	-	-	-	-
16	χ_{16ij}^{m}	-0.314	1.83e-5	-0.155	3.96e-4	0.251	3.1e-3
17	χ_{17ij}^{m}	0.268	6.5e-4	0.164	2.8e-3	-0.512	3.8e-6
Service time estimate: 52 minutes							

In Table 2, each estimate is associated with the standard deviation of estimates around the surrogate surface (see Equation (26)). The columns ‘std’ shows how reliable estimates are in terms of the standard deviation. In general, we can see that the estimated parameters are less reliable for the afternoon as compared to morning and off-peak periods. This is probably because there are relatively lower trip legs in tours that fall in the afternoon period and therefore the algorithm has fewer observations that it can learn from.

The result shows that both transport costs and travel time costs between *i* and *j* (χ_{1ij}^{m} and χ_{1ij}^{m}), increase the cost c_{ij} and therefore lower the preference of planners to keep *j* after *i* is the sequence of trips. The transport cost has a higher impact on the planner preferences as compared to travel time cost. The higher weight for travel time in the afternoon indicates that carriers value afternoon travel times more than morning or off-peak periods. This means that carriers are more likely to avoid including a link in a tour that has a larger travel time during the afternoon.

Similarly, transport costs in the afternoon are more important than the morning and off-peak periods. This indicates that trips that are scheduled in the afternoon within a tour are more likely to be shorter than trips in the morning or noon. However, the standard deviations of estimates

for a time in the afternoon, although acceptable, are slightly higher than the morning and afternoon estimates. This is because of the lower number of tour samples in this category for parameter estimation.

The load of vehicles χ_{3ij}^{m} has significant negative impacts on the planners' cost c_{ij} and therefore, if the vehicle has a higher load on the link i-j then the cost of this link is less from the planners' perspective. The ratio of the number of commodities in *j* over *i* (χ_{4ij}^{m}) also has a significant negative impact on the planners' cost which implies that if there is a larger number of commodities in *j* as compared to *i* for pickup or delivery, the perceived cost of planners traversing the link i-j in a tour will become less. The same explanation emerges for the ratio of weights χ_{4ij}^{m}, however, the magnitude of the impact is relatively lower.

Parameters associated with features χ_{6ij}^{m} and χ_{7ij}^{m} indicate how travel time and cost of the depot to *i* and *j*, can influence the planners' inclination to keep link i-j in a tour. As the figures suggest, these features have negative impacts on the total cost (positive impact on planners' preferences). This intuitively shows that if *j* is further away from depot as compared to *i* or the travel time between depot and *j* is higher than the travel time between the depot and *i*, the planners would like to keep j after i in the tour.

Parameters associated with features χ_{8ij}^{m} and χ_{9ij}^{m} indicates how travel time and cost of *i* and *j* to the depot, can influence the planners' preferences to keep *j* after *i* in a tour. The sign of the parameters shows that, during off-peak and afternoon peak periods, these features have positive impacts on the total cost of traversing from *i* to *j*, from the planners' perspective. This means that if the depot is further away from *j* as compared to *i* or the travel time between *j* and depot is higher than the travel time between *i* and depot, the planners do not prefer to keep j after i in the tour. For the morning peak period, however, the sign is counter-intuitively negative. One possible reason for this could be that planners do not value trips back to the depot in the morning as during this period trips in tours are more outgoing flows. The lower std for these parameters also shows relatively low confidence which means probably there are fewer observations to support this parameter.

The negative sign of features χ_{10ij}^{m} indicates that the trips between distribution centers and producer-consumer have lower perceived costs from the planners' point of view. On contrary, if both i and j are distribution centersχ_{11ij}^{m}, then planners prefer not to put *j* after *i* in the trip sequence of a tour. Given our dataset, the model cannot estimate a significant parameter for the afternoon due to a lack of observations.

Feature χ_{12ij}^{m} to χ_{17ij}^{m} relates the planners' preferences for routing and scheduling of five different commodity types. Unlike other features that interdependently explain the preferences of planners towards routing and scheduling of trips inside a tour, these features mostly influence only the scheduling decisions. This is because the commodity type category does not change within a tour in 95% of our sample dataset. The positive sign of these commodity types within a particular time interval means that planners have less interest to schedule trips carrying the commodity type within that interval. For example the weights of χ_{12ij}^{m} is positive for the off-peak period and negative for morning and afternoon periods. This implies that planners prefer to schedule trips with agricultural products during off-peak periods.

The model also suggests generating on average 52 minutes of service time using exponential distribution for (un)loading locations. The service time generator has a minimum bound of 20

minutes. Drivers' break times are estimated collectively with the service time. Our dataset does not report on when and where they stop to rest but reports on the actual tour duration, the model can only capture these extra times along with the service time of each zone.

3.4.3 Model performance

To study the performance of the model, we report on the mean absolute percentage error (MAPE) and the coefficient of determination (R-squared) regarding estimated and observed tours' length and duration. To evaluate our model, we also compare the fit of the proposed data-driven TSP model with the benchmark TSP model.

Table 3-3: Performance of the model

KPI	Model	Train data		Test data	
		MAPE(%)	R^2	MAPE(%)	R^2
Tour length (km)	Data-driven TSP	5.1	0.92	9.3	0.9
	Benchmark TSP	18.6	0.81	17.9	0.82
Tour duration (min)	Data-driven TSP	12.3	0.88	12.7	0.86
	Benchmark TSP	23.2	0.68	21.8	0.71

The high R^2 measurement confirms the overall good performance of the data-driven TSP model. Table 3-3 shows that the purposed model can predict tour length with 5.1% error on the train and 9.3% error on the test datasets. This is higher than the benchmark TSP model with 18.6% and 17.9% errors on train and test data sets respectively. The data-driven model also outperforms the benchmark TSP model regarding tour duration. The lower performance of the model in predicting tour duration as compared to the tours' length is because of the stochasticity in the service time of the visiting locations.

3.4.4 Model evaluation

The tour distance and number of stop distributions are two important aggregate features that should be captured well by tour models. To evaluate the performance of our model to reproduce aggregate tour characteristics for the purpose of freight modelling, we simulate the model on the available shipment data. This time, however, we do not specify the link between shipments and tours. In other words, only shipment origins and destinations (OD) is available to the firms. In this experiment, firms do the routing and scheduling for multiple vehicles to meet all the shipments on the planning day. Therefore, we set back the K parameters in equation (3.5) to the maximum number of trips started from a zone. This will convert the data-driven TSP model to a data-driven VRP model. All the estimated parameters remain the same. Note that shipment OD is available at 6-digit postcode level. Multiple firms may be active in a 6-digit postcode. We assume, however, that all the shipments that belong to a postcode belong to one single hypothetical firm. This assumption can be refined using advanced population and shipment synthesis models, if available. We also add 11265 direct tours to our sample for model evaluation.

For each firm, we use benchmark VRP and the data-driven VRP model to route and schedule tours. We group tour length into different clusters and compared the estimated tour distance distribution with the observations in Figure 3-2. The results show that the model is able to

approximate this feature reasonably well. Unlike the benchmark VRP model, the proposed data-driven VRP model can re-produce tour length distribution.

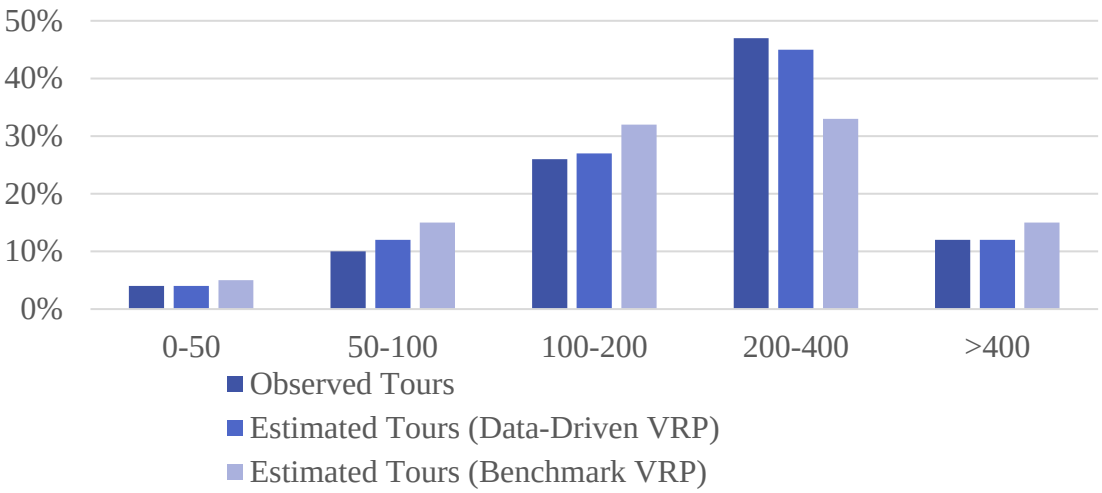


Figure 3-2: Estimated vs. observed tour distance distribution

Similarly, in Figure 3-3 we can see that our data-driven VRP model captures the general pattern in the frequency of tours with different numbers of stops. In general, it can be concluded that tours with a smaller number of stops (3 and 4) are slightly overestimated.

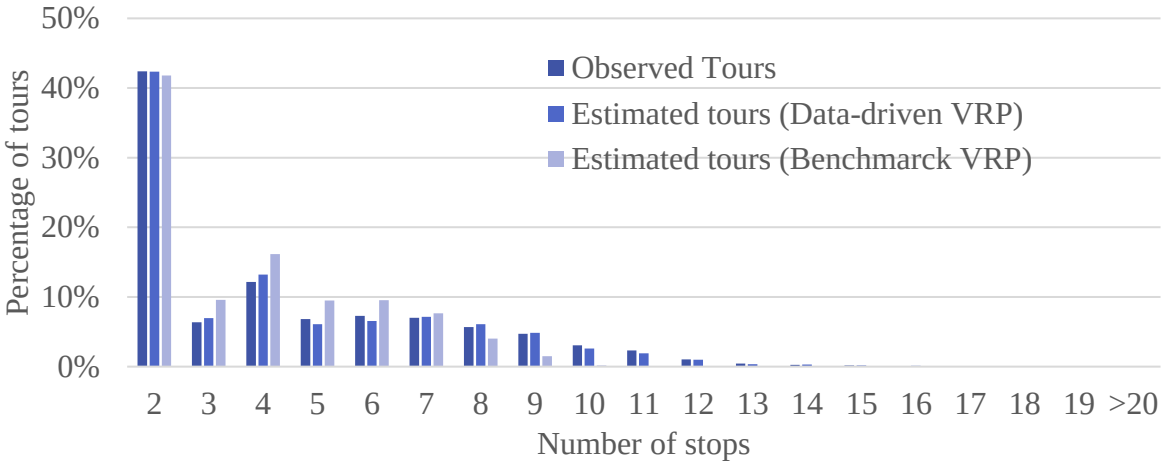


Figure 3-3: Estimated vs. observed tour number of stops distribution

We also compare the coincidence rate of estimated and observed tours' departure and end time in Figure 3-4. After estimating the departure time and the end time of tours using the proposed model, we count the number of tours that their departures and end times fall into the same time intervals as the observe tours. The figure shows that our model can correctly predict departure (up to 96%) and end (up to 95%) times of the tours.

a) Confusion Matrix for departure time

Observed \ Estimated	Off-peak	Morning	Afternoon
Off-peak	87%	11%	3%
Morning	15%	82%	4%
Afternoon	4%	0%	96%

b) Confusion Matrix for end time

Observed \ Estimated	Off-peak	Morning	Afternoon
Off-peak	82%	11%	9%
Morning	7%	81%	12%
Afternoon	2%	3%	95%

Figure 3-4: Observed and estimated departure time and end time intervals

3.4.5 Sensitivity analysis

In addition to the model validation, we test the sensitivity of our model towards the increase or decrease of travel time. From the data, we identify the most frequent link (i-j) in all tours. We increase and then decrease travel time on this link for both morning and afternoon by 10%, 20%, and 50%. The results in Figure 3-5 show that the model is sensitive to changes in travel time and assigns fewer trips to this link if we increase the travel time.

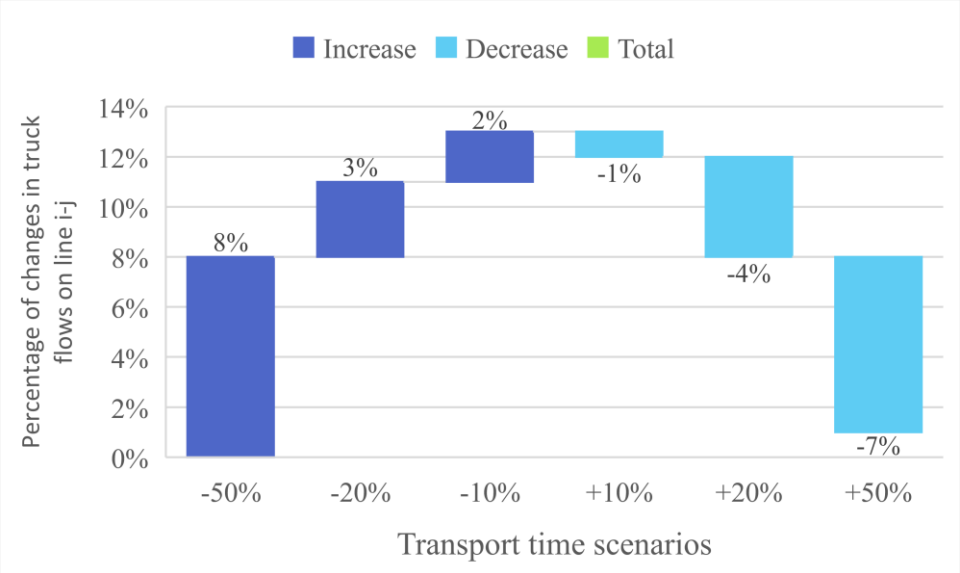


Figure 3-5: Sensitivity analysis for travel time increase/decrease on the most frequent link i-j

We also experiment with the sensitivity of the model to changes in travel costs per kilometre. Changes in transport costs can happen due to, for instance, tolling systems, increases in labour costs, fuel costs, or carriers' fixed costs. Figure 3-6 indicates that the proposed model responds logically to changes in travel costs.

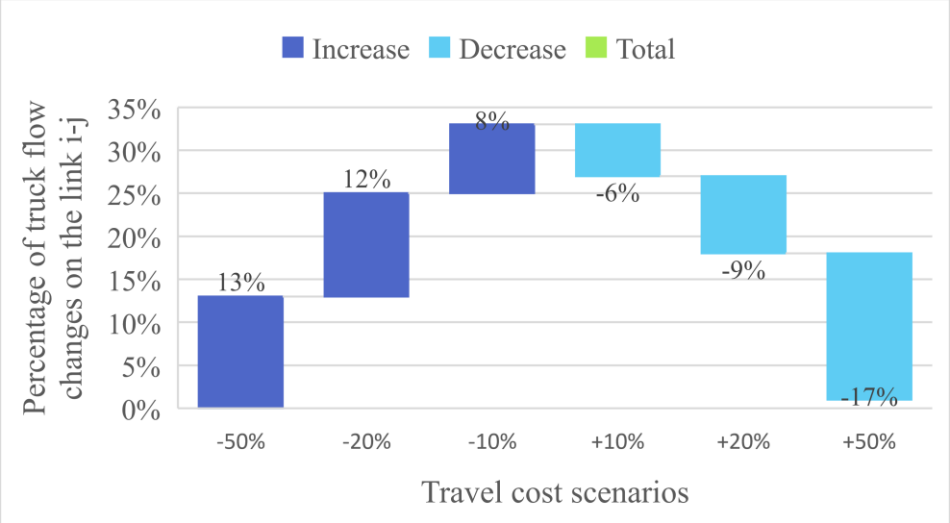


Figure 3-6: Sensitivity analysis of the model on transport cost variations

A comparison between the sensitivity of the tours to cost and time shows that carriers are more sensitive to transport costs than travel time. The sensitivity of this model towards time and cost, along with its ability to model tours spatially and temporally, makes this model a strong and more realistic tool, as compared to conventional models, for policy implications and impact assessment.

3.5 Discussion and conclusion

In this research, we investigated the possibility of developing a data-driven tour model to capture the activity of freight carriers from disaggregate and partially available tour data. We parametrized a time-dependent VRP model to capture the aggregate spatial and temporal correlations between trip legs in a tour. We used Bayesian optimization to estimate the parameters of this model. The results from the model show that the estimated model not only fits the individual tour flow data with high accuracy but also it captures aggregate tour features like tour distance distributions and the number of stops. The model is sensitive to time and cost changes and can be used as a tool in traffic simulations for integrated traffic and freight transport modelling.

The proposed model has practical capabilities, and also limitations. Estimating a general tour model which can represent the collective tour activities of large carriers can become an undetermined problem, especially when a limited imperfect amount of disaggregated data is available. Researchers have been dealing with these problems in transportation science for decades. The performance and accuracy of models in these cases depend largely on the amount of prior information used to help the model converge to the real behaviour of road users. Our proposed method shows that additional, although imperfect, tour information like tour duration, cost, start time, and end time can provide a robust and accurate estimation of a single tour model with meaningful parameters that can reproduce the preferences of carriers in planning tours.

Although the model validation shows that the proposed model can accurately predict the routing and scheduling of carriers, the high accuracy of the model on the test data is not a surprise as the test data also come from the same sample distribution (but unseen data) on which we build the model. The ideal way of evaluating the model performance would be to use secondary data

such as truck counts on road networks to validate and calibrate our model. This, however, requires integrating this model into a more comprehensive freight transport model.

To the best of our knowledge, there is no comparable study that we can compare our results with across all dimensions. De Jong et al. (2016) estimated models to explain the time-period preferences of receivers in road freight transport. They have applied this model to assess the impact of these preferences on the time-period choices of carriers. De Jong et al. show that road freight carriers are relatively insensitive to travel time changes. In contrast, carriers are more sensitive to peak hours if they realize an increase in transport costs. De Jong et al.'s model only considers the time dimension in freight transport without considering the impact of routing. Interestingly, our model, which treats both routing and scheduling of carriers, shows the same behaviour in tour planning. The spatial-temporal connection that trips have in our model makes this method a good candidate for spatial and temporal analysis. One can use this model, for instance, to study the impact of off-peak policies or to investigate the impact of an increase or decrease in transport cost (e.g., fuel cost, distance/shipment size-related costs) or cost of travel time (e.g., toll system), on the activities of carriers. These abilities make this an interesting tool for policy analysis with capabilities that go beyond current models.

Seen as a tool for analysts, our work can also be used to understand the impact of freight activities on the traffic system. The activities resulting from this model can be easily translated into a time-dependent truck OD table, which can be coupled with a traffic simulation model to investigate the interconnection between freight and traffic. In this study, the model used the travel time matrices from a traffic simulation model to estimate tours. In return, the time-dependent truck OD that resulted from this model can be also used as input to traffic simulation models. This could provide interesting insights into the inter-relation between freight transport operations and traffic systems. Auxiliary data such as truck counts on the road networks can also be used to enhance the accuracy of the VRP calibration and time-dependent truck OD matrix estimation.

Chapter 4

Traffic prediction based on freight trip schedules

After Modelling tours from trip data in the previous chapter, it is important to predict the impact of tour schedules on the dynamics of the traffic on roads. Short-term traffic prediction is a key component of traffic management systems. Around logistics hubs such as seaports, truck flows can have a major impact on the surrounding motorways. Hence, their prediction is important to help manage traffic operations. However, The link between short-term dynamics of logistics activities and the generation of truck traffic has not yet been properly explored. This chapter aims to develop a model that predicts short-term changes in truck volumes, generated from major container terminals in maritime ports. We develop, test, and demonstrate the model for the port of Rotterdam. Our input data are derived from exchanges of operational logistics messages between terminal operators, carriers and shippers, via the local Port Community System. We propose a feed-forward neural network to predict the next one hour of outbound truck traffic. To extract hidden features from the input data and select a model with appropriate features, we employ an evolutionary algorithm in accordance with the neural network model. Our model predicts outbound truck volumes with high accuracy. We formulate 2 scenarios to evaluate the forecasting abilities of the model. The model predicts lag and non-proportional responses of truck flows to changes in container turnover at terminals.

This chapter is based on the following paper:
Nadi, A., Sharma, S., Snelder, M., Bakri, T., van Lint, H., & Tavasszy, L. (2021). Short-term prediction of outbound truck traffic from the exchange of information in logistics hubs: A case study for the port of Rotterdam. *Transportation Research Part C: Emerging Technologies*, 127, 103111. DOI: https://doi.org/10.1016/j.trc.2021.103111

4.1 Introduction

Predicting Short-term traffic flow is indispensable for advanced traffic management systems. The truck flow around logistic hubs, which varies by time-of-day, has implications for the traffic on the surrounding motorways. Therefore, predicting the next hour of truck volumes is a precondition for traffic management systems for controlling the corresponding traffic. Nonetheless, short-term prediction of truck volumes has gained too little attention in contrast to short-term traffic flow prediction.

The literature mostly addresses the daily truck generated at logistic hubs or traffic analysis zones rather than addressing the short-term truck demand on the road network. This implies that no literature specifically describes methods to predict short-term truck flows. A comprehensive view of the existing methods for long-term freight generation can be found in Holguín-Veras et al. (2014). Additionally, only a limited number of data collection methods (surveys, demographics) have been tested for this problem. Furthermore, these methods mostly use site-related characteristics, i.e. the number of employees; area; and capacity, which results in significant errors in the case of a relatively high number of trucks e.g. seaport's terminal (Sarvareddy et al., 2005). Most importantly, however, these methods do not address the dynamics of within-day variations of truck flows. Currently, therefore, traffic managers lack the tools to make accurate short-term predictions of truck flows.

As truck demand is a consequence of the logistics operations of multiple actors, it is worthwhile to explore the possibility to predict flows based on the exchange of information between them. Since the early 2000s, several studies have emphasized the importance of information exchange in freight transport and logistics chains (Giannopoulos, 2004, Di Febbraro et al., 2016, Banister and Stead, 2004). Cooperation between different logistics actors is important in intermodal hubs, such as seaport terminals, where a prompt exchange of information is needed for cooperative planning and execution of intermodal freight transport (Di Febbraro et al., 2016). Due to the advances in information and communication technologies, many ports around the world now use different systems to benefit from information transmission. In particular, Port Community Systems (PCS) are there to ensure that everyone in the hinterland transport chain (e.g. Port authority, terminals, careers, depots) can easily exchange information. These systems are generally built upon an integrated central database where all the information from clients of the port and government agencies, like customs, come together. PCS could align the vessel arrival time and container discharge time to the truck's container pick up time. This information is mostly useful for optimizing terminals operations (Heilig and Voß, 2017). PCS also has the potential to provide other additional information services. For example, road carriers could be offered planning information services, to give terminals and empty depots a prior notification of the trucks' arrival times. Terminal operators, in return, could manage these arrivals and speed up the loading and unloading process of the containers at the terminal gates. One of these functions, that we will explore here, is the provision of timely warnings about truck flow dynamics to traffic managers.

The main objective of this paper is to propose an approach to predict the next hour's truck volume, based on PCS data and truck count data. The key contributions of this paper are as follows.

1. We predict short-term truck volume with high accuracy on motorways with high truck demand;
2. We propose an analytical framework making use of an artificial neural network (ANN) and NSGA-II to predict traffic, adding to existing approaches in two ways:
    a. It provides a robust feature extraction method

   b. It uses a novel method to generate different NN topologies and heuristically select the best model.
3. In this method, we combine PCS data, which comes from logistic activities, with class-specific loop detector data. To the best of our knowledge, this paper is the first to utilize PCS data for the prediction of short-term truck volume on the road network.

The remainder of this chapter is organized as follows: Section 4.2 provides a brief overview of the existing studies about the traffic flow prediction methods in general and truck volume prediction methods in particular. Section 4.3 presents the data collection outlines with which we describe the data characteristics and preparation challenges. Section 4.4 describes the ANN-based methodology we proposed to model the data. While section 4.5 denotes results and tests the prediction capability from a simulation of the model with some designed scenarios, section 4.5 at the end, offers concluding insights.

4.2 Literature review

In this section, we provide a brief overview of categories of the existing methods covering short-term prediction of the traffic volume. Then more specifically, we review studies that take truck volume prediction into account.

4.2.1 Short-term prediction of traffic volume

Short-term traffic prediction has been an important subject in transport research since the 1980s. The term short-term usually refers to predictions made from a few minutes to a few hours into the future using historical data. The application of accurate short-term traffic prediction is usually input to intelligent transport systems in order to optimize the traffic performance in the coming hours. Readers can find a comprehensive overview of the methods and challenges in Vlahogianni et al. (2014), Van Lint and Van Hinsbergen (2012), and Poonia et al. (2018). In general, state-of-the-art traffic prediction can be classified into three types of models (Van Hinsbergen et al., 2007). The first class consists of naïve methods, where the models use no specific intelligence to predict the target values. An example of these methods is the historical average. However, these methods have relatively low performance in predictions; The second class is known as parametric methods, where the models use traffic flow theory to predict the traffic state., in an approach based on the traffic flow model, we have inputs (demand, route choice), parameters (e.g. capacity, critical speed) and internal state variables (e.g. density). One can update/estimate all three using data.; Finally, nonparametric methods are another group of traffic prediction methods that use a flexible mathematical function approximation structure with adjustable parameters. These methods have the potential to learn a nonlinear model structure and parameters from the data. Examples of these methods are ARIMA (Kumar and Vanajakshi, 2015, Williams and Hoel, 2003), Spectral analysis (Nicholson and Swann, 1974), Kalman filter (Kumar, 2017), fuzzy logic (Zhang and Ye, 2008), support vector regression (Hong, 2011), stochastic differential equations (Rajabzadeh et al., 2017), and, shallow or deep Neural networks (Polson and Sokolov, 2017, Tian and Pan, 2015, Do et al., 2019). Among all the mentioned methods, state-of-the-art deep neural networks have gained ample attention in recent years. Since the last decade, various types of deep architecture of neural networks have been developed for short term traffic predictions. Lv et al. (2014) proposed a stacked autoencoder model to predict traffic flow 15 to 60 minutes ahead. Wu et al. (2018) developed a hybrid convolutional recurrent deep neural network based on a traffic flow model to predict traffic states. Deep networks like long-short term memory (LSTM) and gated recurrent units (GRU) networks have also been successfully used to predict traffic flow (Fu et

al., 2016, Zheng et al., 2020, Cui et al., 2020b). Moreover, graph-based deep networks and generative adversarial networks are other types of deep learning networks that were successfully designed and applied to this field (Xu et al., 2020, Cui et al., 2020a). More details about these methods and their differences can be found in Lana et al. (2018).

The difference between parametric (model-based) and nonparametric (black-box) approaches is that in black-box approaches such as e.g. a NN, there are typically many more parameters (with no physical interpretation) than in a model-based approach (in which most of the parameters have some physical interpretation).

4.2.2 Prediction of truck volume

Unlike traffic in general, the short-term truck volume prediction is not, at the present time, receiving sufficient attention. One of the reasons for the inattention to this transport sector is the absence of data. Besides, it is generally believed that truck volume is only a minor part of the traffic flow and its peak does not occur simultaneously with the passenger car peak. This is clearly not a proper assumption especially on motorways that connect logistic hubs (i.e. seaport terminals) and the hinterland. The existing truck volume prediction methods fall under the same categories as the traffic flow prediction. However, in most cases, researchers have used parametric methods. For example, Liedtke (2009) introduced the INTERLOG simulation model prototype. This model is a simulation framework that takes into account logistics choices to assign commodity flow between companies to truck tours on the road network. Holguín-Veras and Patil (2008) introduced a multi-commodity estimation model taking into account empty trips. This model is able to estimate truck OD table as well as link volumes. Other similar examples are Gonzalez-Calderon and Holguín-Veras (2017), Wisetjindawat et al. (2012), Sánchez-Díaz et al. (2015). Some studies also focused on the simulation of the traffic and congestion around cities with maritime container terminals. One of the earliest studies in this regard is Pope et al. (1995). In this study, they used a traffic simulation model to examine the impact of container traffic at terminals on traffic congestion around the port. More recent work on a similar topic however is from Li et al. (2016). They made use of a combination of discrete-event simulation and a traffic model. In this model, they used vessel arrival and terminal gate data to develop a queueing model. Then with making use of a traffic flow model, they evaluated traffic conditions for a number of what-if scenarios.

For the nonparametric methods, on the other hand, Dhingra et al. (1993) proposed a time series analysis using ARIMA to predict weekly truck volume attracted by the Bombay metropolitan region. In the year 2000, Al-Deek et al. (2000) developed a linear regression model to predict daily outbound truck traffic from freight container data of the port of Miami. Then they used this daily prediction to forecast hourly truck volume using average hourly distribution. Unlike the daily model, the hourly forecast is not validated in this model. In another study in 2002, Al-Deek (2002) used vessel data to predict daily outbound truck volume using Backpropagation neural networks (BPNN) with 92% accuracy. The model was reported to be sensitive to the variation of terminal activities. Klodzinski and Al-Deek (2004) also used the ANN method to model daily truck generation from the vessel data. They used a scaling factor that predicts hourly truck volumes during peak hours. Another study from Sarvareddy et al. (2005) compares the fully recurrent neural network (FRNN) with the backpropagation neural network (BPNN) model to predict daily truck volume using vessel data. In contrast to the BPNN, the FRNN was not sensitive to the variation of vessel data. Similarly, Xie and Huynh (2010) compared different machine learning algorithms (i.e. SVM, Gaussian process, and feed forward neural network) to predict daily truck volume from seaport terminal operation data. They find SVM and GP outperforming the feed forward neural network despite they need less effort for model

fitting. Despite all the great research mentioned above, the question about short-term truck volume prediction arises where traffic management interventions are required on motorways with high truck traffic. Moniruzzaman et al. (2016) addressed this question by developing a feed-forward neural network to predict short-term truck volume on a bridge crossing the border from the US into Canada.

A summary of the reviewed studies on truck volume prediction is shown in Table 4-1. All except one of the mentioned methods are directed at the aggregate level of traffic flows or intended for the planning of truck facilities over the long term: a week, a day, or years ahead. We conclude that an approach for short-term truck volume prediction for traffic management purposes is lacking. Therefore, our contribution focuses on the development of this approach. We employ logistics source data, that has not been used before for this purpose, and combine several available methods to achieve an accurate short term prediction of truck flows. The remainder of this paper sets out the modelling approach and the results.

Table 4-1: Examples of the existing studies on truck flow prediction

Author and Date	Area	Methodology		Horizon	Data
		Approach	Model		
Liedtke (2009)	Network-level	parametric	Micro-simulation	Long-term	Contract data
Holguín-Veras and Patil (2008)	Network-level	parametric	Trip simulation and OD synthesis model	Long-term	Survey data
Gonzalez-Calderon and Holguín-Veras (2017)	Network-level	parametric	Entropy maximization and tour synthesis model	Long-term	Truck counts
Wisetjindawat et al. (2012)	Network-level	parametric	simulation	Long-term	Survey data
Sánchez-Díaz et al. (2015)	Network level	parametric	Model-simulation	Long-term	GPS data
Pope et al. (1995)	Port area	parametric	Microscopic traffic model	Long-term	Survey-truck count
Li et al. (2016)	Port area	parametric	Discrete event	Long-term	Vessel data
Dhingra et al. (1993)	Port area	Non-parametric	ARIMA	weekly	Truck counts
Al-Deek et al. (2000)	Port area	Non-parametric	Linear regression	Daily	Vessel data
Al-Deek (2002)	Port area	Non-parametric	BPNN	Daily	Vessel data
Klodzinski and Al-Deek (2004)	Port area	Non-parametric	ANN	Daily	Vessel data
Sarvareddy et al. (2005)	Port area	Non-parametric	FRNN & BPNN	Daily	Vessel data
Xie and Huynh (2010)	Port area	Non-parametric	SVM & GP	Daily	Terminal data
Moniruzzaman et al. (2016)	Border	Non-parametric	ANN	Short-term	RTMS
Current study	Port area	Non-parametric	ANN	Hour (short-term)	PCS

4.3 Data collection

To develop and test our model, we use the case of the Port of Rotterdam in The Netherlands. The Port of Rotterdam is the largest in Europe. It has been growing at a yearly rate of 4.5% in terms of container throughput (Port-of-Rotterdam, 2019). Due to its growth expectations, it is likely to induce more truck traffic on the motorways in the Netherlands. We used container data of the port of Rotterdam to predict the next hour outbound truck traffic. Figure 4-1 shows the terminals located in the port area of Rotterdam and the location of the loop-detector that observes all relevant outbound truck traffic.

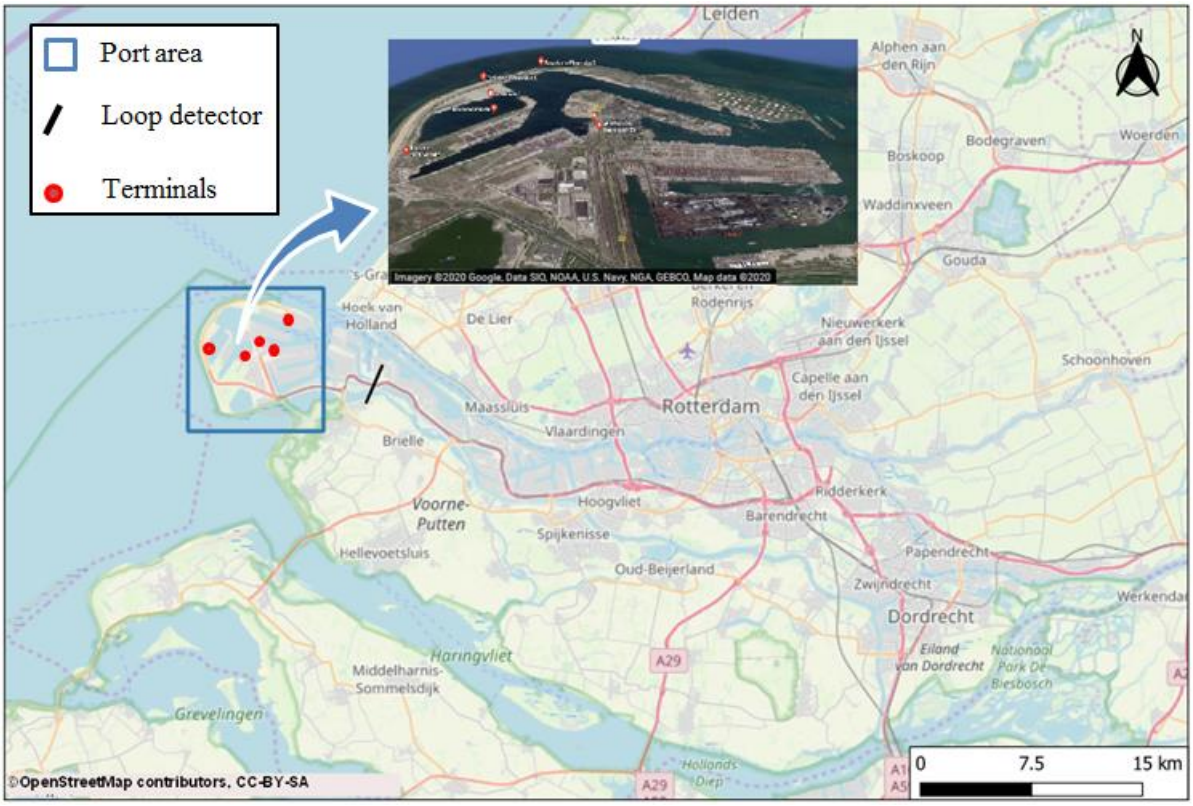


Figure 4-1: Terminal and main loop-detector location considered in the port of Rotterdam

4.3.1 Container schedule data

Container schedule data for five terminals operating in the PCS Port of Rotterdam are managed by the company Portbase. Data were provided for the year 2017. This dataset contained necessary and relevant information about sea-side (i.e. vessel arrival times), yard-side (i.e. container discharge times) handling of containers by terminal operators, and hinterland-side operation (i.e. estimated container pickup times by trucks). While the first two belong to the terminal operations, the third is linked to the trucking companies. In other words, the whole container process from unloading from vessels and moving to the container yards to loading them on trucks is recorded in chronological order. Previous analysis has shown that this data was highly important for the generation of truck volumes (Nadi Najafabadi et al., 2019). The main field which we used as the input layer for neural networks was the Estimated Pick-Up Time for containers. After aggregating container-level data, we obtained expected road-side outbound volumes of containers in a given hour.

4.3.2 Truck volume

In the Netherlands, a national repository named National Data Warehouse (NDW) provides a data stream of vehicle counts collected from loop-detectors installed on motorways, a subset of these loop-detectors can distinguish vehicle categories based on their lengths. Following the same classification, vehicles longer than 12 meters were labeled as trucks. To detect outbound truck traffic from the port area, we selected the first loop-detector located at the beginning of the A15 motorway, within the Port of Rotterdam area (see Figure 4-1). These data were available at a resolution level of 1 minute time periods. Truck traffic data was aggregated at an hourly resolution and was used for the output layer of our neural networks.

Since we apply neural networks for time-series forecasting, the continuity of the data should be maintained. Therefore, we used the first 6 months of data for the year 2017. It provided us with 4344 data points for both the input and output layers. Figure 4-2 shows hourly containers that are ready to be picked up and Figure 4-3 shows the truck traffic observed at the loop-detector on A15. Container data was complete; however, 3.84% of truck traffic data were not available. Missing traffic data is a common problem and can be caused by several factors such as malfunctioning of loop detectors or communication issues. In section 4.4, we provide our approach that we used to handle missing data before the neural networks were applied.

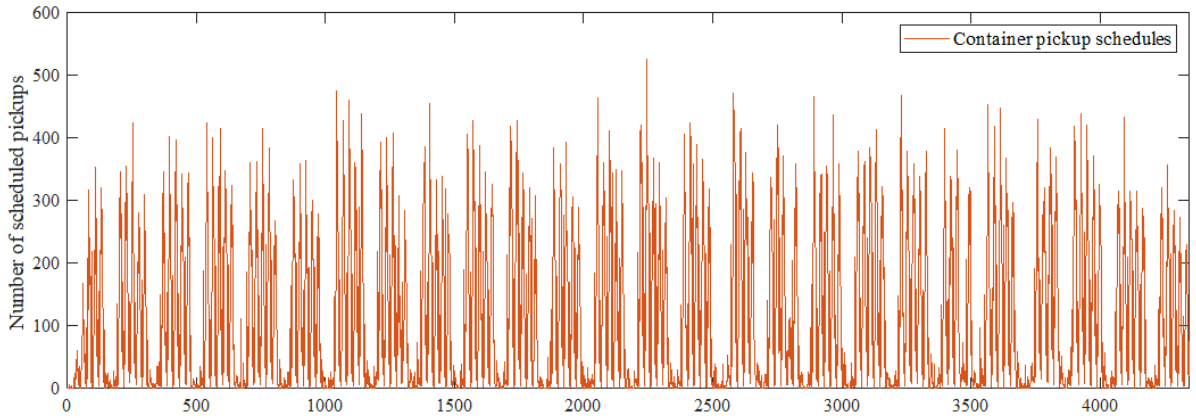


Figure 4-2: Hourly containers ready to be picked up at the port of Rotterdam

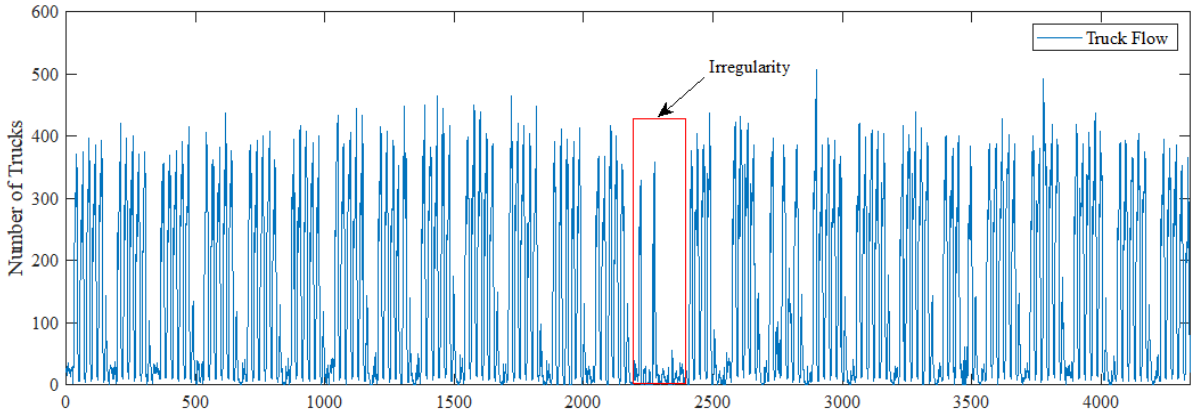


Figure 4-3: Hourly truck traffic counts observed at the loop-detector on the A15

4.3.3 Exploratory data analysis

As far as the time series data was concerned, we first set up a descriptive analysis to look into the input/output trends. Altogether 6 months of data were collected for this study. We examine the trends from an hourly, day of the week, and monthly perspective.

Figure 4-4 shows the average number of trucks versus the average number of containers at each time of day. It is obvious from the figure that, although they have a similar increasing and decreasing pattern in general, there are some differences, especially during peak periods. Truck flows peak from1:00 PM to 7:00 PM. By comparison, container pickup schedules show that containers are mostly scheduled to be picked up from 11 AM to 5 PM. This indicates that trucks likely face delays in picking up the containers, especially during peak periods. As we can see from the truck flow profile, unlike the container pickup schedules, the flow drops at 8:00 AM and 4:00 PM. This is due to the working shift change at terminals, which leads to fewer truck departures during these hours.

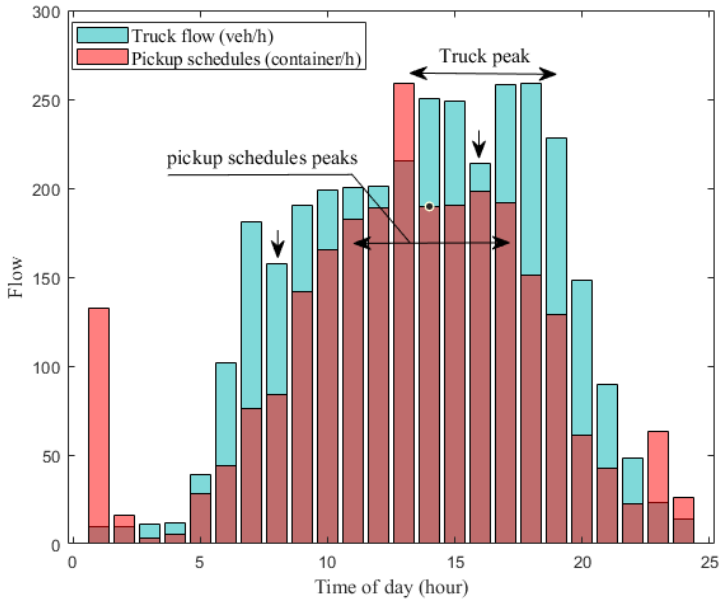


Figure 4-4: Number of containers and number of trucks throughout the hours of the day

Figure 4-5 compares the average truck flow and scheduled container pickups on every day of the week. This graph shows that the two trends are generally similar. As can be seen from the figure, the average truck volume on Wednesdays and Fridays is slightly higher than on the other days.

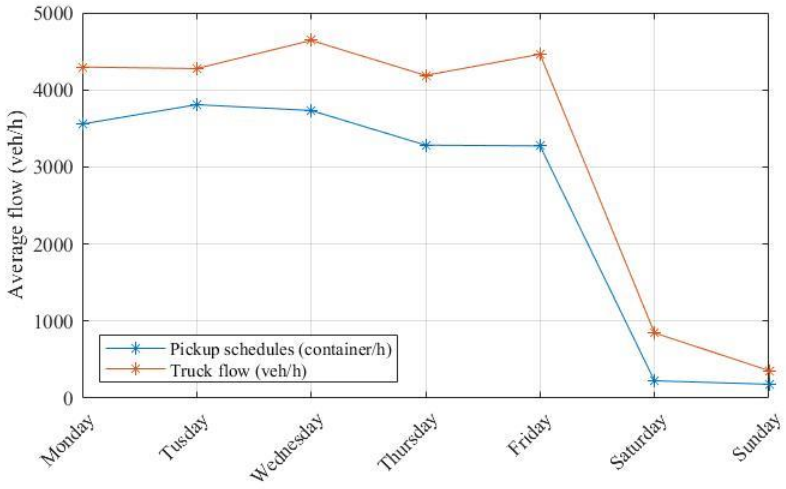


Figure 4-5: Comparison of truck flows and scheduled container pickups during the week.

Finally, Figure 4-6 compares the number of trucks vs. scheduled container pickups during the first six months of the year 2017. Although the two trends have a similar pattern in general, the average volume varies from one month to another. As we can see, there is an abrupt drop in the average truck volume in April. Of course, we would expect a slight drop due to the same pattern in the container scheduled flow; the decrease however is sharper than expected. This could be due to an irregularity in the truck flow data. This irregularity is also observable in Figure 4-3 for the first two weeks of April. We could impute these two weeks with the mean or median profile from historical data. However, we decided to keep the observed truck counts unchanged for two reasons. First, the container pickups may be scheduled days or weeks ahead and may not be updated for those two weeks. In other words, the abnormality could be in container schedules not in the truck count. Second, this abnormality is only for 8 percent of the data and not only may not have huge consequences on model prediction in general but also makes it more robust toward possible unexpected variations.

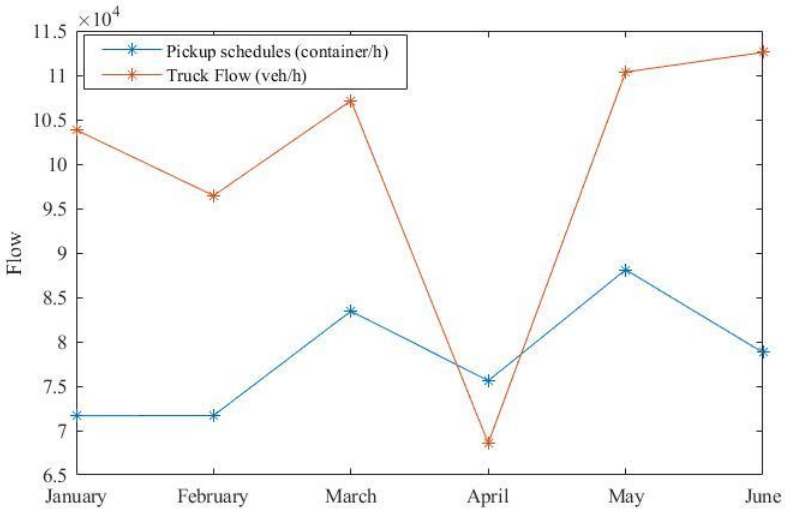


Figure 4-6: Comparison of the truck flows and scheduled container pickups in the first six months of 2017.

4.4 Methodology

We design and train an ANN model to predict hourly truck volumes on a specific section of motorway, directly connecting the container terminals with the main road networks. Figure 4-7 shows the building blocks of this model.

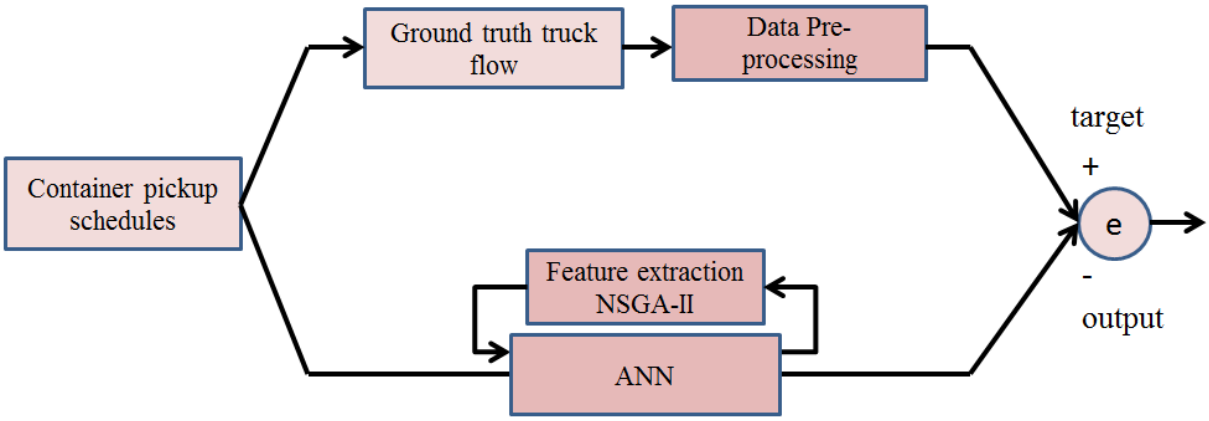


Figure 4-7: Building blocks of our truck traffic prediction model

The container pickup schedules contain hourly expected pickup times for each individual container in the port of Rotterdam. In this model C_t represents the number of containers that are expected to be picked up at time t and T_t denotes the ground truth outbound truck volumes on the selected section of the motorway. We formulate the problem as a way to map container demand time series to the supply truck flow time series, as follows:

$$T_t = f(C_t, C_{t-1}, C_{t-2}, ..., C_{t-d}) \tag{4.1}$$

where d is the number of time lags. In this model, we intend to predict truck flow at time t given expected container pickups at time t, and d previous time steps. The reason that we use the time series of container pickups is due to the problem that trucks do not always arrive at the expected pickup time, i.e. they could arrive later or earlier that day. We identified from our exploratory analysis that it is likely that trucks arrive slightly later. Therefore for prediction, the truck flow at time t might be attributed to the container pickup demand in current and previous time steps. The estimated weight of our trained ANN accounts for the contribution of each time lag in container schedules to predict the current truck flow. Moreover, the truck flow at time t also includes the empty trucks which are not represented in the estimated pickup time of containers. As the ratio is relatively low the nonlinear characteristics of our trained ANN can capture the pattern of all truck flows including the empty trips.

Missing data is a common problem to deal with in machine learning methods. Since we only have had 3.84% of missing truck volume data, we have used an imputation method; specifically, we have used historical hourly truck volume averages to fill missing data. As this concerned a very minor share of the counts, scattered across the time series, this had a negligible impact on the results.

4.4.1 Artificial neural network

An artificial neural network is a machine learning technique that uses a learning mechanism similar to the human brain. It has found very successful applications in many disciplines, especially in modelling complex dynamic features in data. This method is believed to be technically superior to classical statistical methods where there is a nonlinear relationship between dependent and independent variables, due to its ability to map input space to a nonlinear feature space. In transportation research, in particular, one can formulate the traffic dynamics of the road network as a time series of speeds and flows. A neural network often demonstrates high performance in capturing such spatiotemporal complexities.

As we have seen in the literature review section, there are many variants of neural networks that can deal with the prediction of traffic flow time series. Examples are feedforward networks, recurrent networks, and radial-based function (RBF) networks. In this study, we use a feedforward neural network to predict truck volumes. A typical feed-forward neural network consists of one input layer, one or more hidden layers, and one output layer. Every layer in an ANN contains a number of neurons. For the detail about the methodology of the artificial neural networks, we refer the interested readers to the fundamentals of neural networks textbook (Hassoun, 1995). The number of neurons in the output layer of an ANN depends on the designated problem. One, however, has to identify the number of neurons in the feature layers. In section 4.2.2, we outline an evolutionary optimization method to find the answer to this problem.

Error backpropagation is an approach that is used to estimate the weight of each connection between neurons of one layer and another layer. Initiating with random weights and biases, the training process continues with successive updating weights and biases in a way to minimize the total error of the model. The MSE (mean square error) is the most commonly used error function for estimating the model weights.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2 \tag{4.2}$$

Where N is the number of observations, $t_i$is the i^{th} observed target value and y_i is its corresponding predicted value.

In order to evaluate the performance of the model, we use root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and the probability of absolute percentage error (PAPE) as the indicators.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2} \tag{4.3}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left\|t_i - y_i\right\| \tag{4.4}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{t_i - y_i}{t_i}\right| \tag{4.5}$$

$$PAPE = \Pr\left(\left|\frac{t_i - y_i}{t_i}\right| < 10\%\right) \tag{4.6}$$

The RMSE represents the standard deviation of the residuals which indicates how the data are concentrated around the best fit. On the other hand, the MAE, tells us how big an error could be on average. We also use the correlation coefficient between targets and predicted values to evaluate the prediction power of the model. In this model, we use two hidden layers because adding more layers does not improve the performance of our prediction. However, this model requires three hyperparameters which we need to tune. That is the number of time lags as well as the number of neurons in each hidden layer of the ANN. In the next section, we outline an optimization method that deals with the model's hyperparameters.

For the optimization algorithm, we use the standard neural network toolbox of MATLAB which provides a flexible platform for training, validating, and testing. From a variety of optimization techniques, we chose to use the Levenberg-Marquardt algorithm, as it provided us with a high goodness of fit.

4.4.2 Feature extraction and model selection

In a neural network hidden layers generally extract a higher level of features from the input data. In fact, the number of neurons in each hidden layer represents the number of hidden features that we have to extract from input data to predict the target values accurately. Finding the required number of higher-level features in a model as well as identifying the activation function of hidden layers is a type of feature extraction problem. Besides, the number of time lag d in the model also identifies the features that have a contribution to the prediction model. In this study, the parameter d is even more important as it helps us to interpret our model better. Therefore, we also defined the problem of finding the minimum number of input features (e.g. d in this study) as a feature selection problem (an approach known as model selection).

4.4.3 Problem formulation

We adapt the same methodology for our problem as used by Huang et al. (2010): we formulate feature extraction and feature selection for this model as a multi-objective optimization problem. The goal of this formulation is to obtain the number of time lags as well as the number of neurons in each hidden layer so that the model's error is minimized. In other words, the aim is to find a tradeoff between the number of features, number of model parameters (i.e. weights and biases), and the sum of squared errors in the model. As the number of input features and the number of neurons increases, the number of the model's parameters increases as well. However, it is likely that the error of the model decreases. Therefore, the best subset of features is a point in the Pareto front of the solution space. The objective function of this formulation makes sure that the minimum number of features and neurons is used in a way that the performance of the model does not decrease significantly.

$$z_1 = \frac{1}{N}\sum_{i}^{N}(t_i - y_i)^2 \tag{4.7}$$

$$z_2 = \left|\vec{X}\right| \tag{4.8}$$

$$z_3 = \left|\vec{W}\right| \tag{4.9}$$

Where $|\vec{X}|$ and $|\overrightarrow{W}|$ are cardinality of the input feature vector and cardinality of the model's parameters respectively. The aim is to minimize all these three objectives. Decomposition methods and direct methods are two types of algorithms that can solve a multi-objective optimization. From the decomposition point of view, one way is to use a weighted sum of all objectives and solve the problem as a single-objective optimization.

$$z = \min(z_1 + \alpha z_2 + \beta z_3) \tag{4.10}$$

However one has to tune α and β correctly to get accurate results. To cope with this problem, the Akaike information criterion (AIC) is used as defined by Hurvich and Tsai (1993). Bayesian Information Criterion (BIC) is also another indicator which is widely used in model selection problem. Both these methods use information theory to trade off model parsimony against goodness-of-fit.

$$AIC_c = n\log(MSE) + \frac{n+m}{1-\frac{(m+2)}{n}} \tag{4.11}$$

$$BIC = n\log(MSE) + m\log(n) \tag{4.12}$$

where n is the number of observations and m is the number of model parameters.

In order to use AIC_c and BIC for the feature selection, there must be a set of candidate models to choose from. In our case, we have three decision variables: a time delay and the number of neurons in the first and the second hidden layer. In this study, we have no prior knowledge about the range of the effective time delay. Therefore, any combination of these decision variables can be a candidate model. This results in a huge set of candidate models. To cope with this problem, we can choose from either weighted sum or direct methods to solve the proposed multi-objective problem. However, there are two main drawbacks to using the weighted sum approach as discussed in several studies, e.g. Kim and De Weck (2006). The first drawback is that the obtained solutions on the Pareto front are not evenly distributed; the second is that no solutions can be found in non-convex regions. Therefore, for our case, we choose to use an NSGA-II setting which can directly trade-off model performance, time delay, and the feature extraction parameters. NSGA-II is a population-based evolutionary algorithm that uses a non-dominated sorting technique and is able to heuristically find all the models located on the Pareto front (Deb et al. (2002). The algorithm starts with a random setting for the model and heuristically searches the solution space to find those models that cannot be dominated by other models. This provides us a small list of the best candidate models among which we can use the BIC or AIC to choose the best model. Accordingly, we call this method a multi-layer neural network with automatic feature extraction (MLP-AFE).

4.4.4 Ensemble models

The NSGA-II trains ANN with different settings (topology) at each iteration. However, ANNs with similar settings may produce multiple predictions. To reduce the uncertainty of the model's error and thus have the most accurate model, we use the ensemble learning approach at each

iteration of the NSGA-II. There are different ensemble techniques such as averaging, bagging, boosting, stacking, and blending. We use the averaging technique as it works well for our problem. We trained the ANN 10 times with the same setting at each NSGA-II iteration and then used the average MSE as the relative objective of that setting.

4.5 Results

In this section, we describe the results of our analysis for the prediction of truck flows, given the container pickup schedules in the Port of Rotterdam.

4.5.1 Model selection

We used MATLAB R2018b to train the ANN for our forecasting model. The model uses a feedforward network architecture with two hidden layers (as extra layers did not improve the fitness of the model significantly). For the activation functions, we use *logsig* in each hidden layer and a linear regression function (i.e. *purelin* in MATLAB) for the output layer.

In the previous section, we proposed a multi-objective optimization technique to find the optimum setting of our model. In this technique, we use NSGA-II to find models that have: the minimum number of delays; the minimum number of neurons in the first hidden layer (NH1); the minimum number of neurons in the second hidden layer; and the minimum MSE (see Figure 4-8). As the MSE is in disagreement with the model parameters, the result of the NSGA-II will be a set of non-dominated models that cannot dominate each other as well. On the other hand, the NSGA-II selects a set of candidate models that are dominated by the other models. These models belong to the Pareto front of the objective space. Figure 4-8 shows the approximation of the Pareto front using the NSGA-II algorithm. The dark blue points on the 3D graph are those on the Pareto frontier.

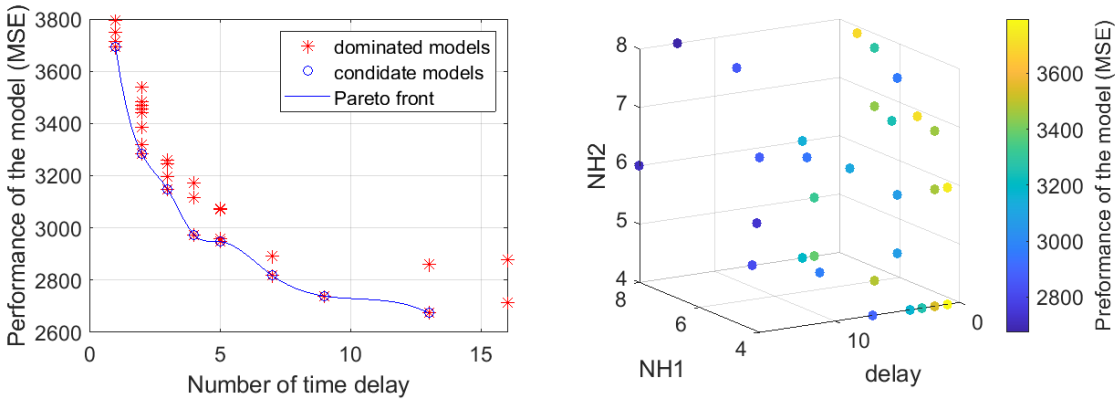


Figure 4-8: Approximation of the models' Pareto frontier using the NSGA-II algorithm.

Table 4-2: Comparison of candidate models using their $\boldsymbol{BIC}$ and $\boldsymbol{AIC_c}$.

Models	delay	NH1	NH2	MSE	AIC_c	BIC
Model 1	1	7	8	3694.127	28217.49	25734.17
Model 2	2	6	8	3285.286	27841.07	25329.26
Model 3	3	8	6	3147.437	27727.31	25287.1
Model 4	4	7	4	2973.562	27498.84	24929.95
Model 5	5	7	6	2948.125	27516.96	25096.3
Model 6	7	8	4	2817.519	27380.84	25014.89
Model 7	9	7	5	2739.393	27317.17	25009.73
Model 8	13	8	8	2674.029	27382.66	25530

From the AIC point of view, Table 4-2 says that model 7 with delay 9, has the lowest AIC. This model has 7 and 5 neurons in the first and the second hidden layer respectively. On the other hand, model 4 has a lower BIC. This model has 7 and 4 neutrons in its first and second hidden layers respectively. In general, the BIC uses a higher penalty for model parameters than AIC. This usually leads to a model with fewer parameters. However, in this study we choose model 7 owing to the lower accuracy of model 4 in predicting the peaks in truck flow. With this model, we can use the previous 9 hours and the current schedules of the container pickups to predict outbound truck volumes on the A15 motorway.

To prevent overfitting of the ANN, we used 70% of the input data for training, 15% for validation, and 15% for testing the performance of the model. We defined 100 epochs to train a feedforward neural network. MATLAB uses the validation set to test the model performance during training. Training will be stopped if the model fails in a number of successive iterations to improve the prediction accuracy for the validation set; this prevents overfitting of the model. In this paper, we set the number of maximum fails in the validation process to 10.

4.5.2 In-sample validation

Figure 4-9 compares the linear fit between the actual and the predicted truck flows for the three data splits and the whole data. The correlation coefficient between the actual and predicted truck flow is just around 0.92 (with P-value close to zero) for Train, Validation, Test, and All data. This indicates a high predictive capability of the model. As we can see from the figure, the model output fits the target values for both validation and test data with a slope close to 1 (0.85 and 0.87 respectively) and a small intercept (25.81 and 22.77 respectively). We conclude that the model has a high generalization power to predict out-of-sample data.

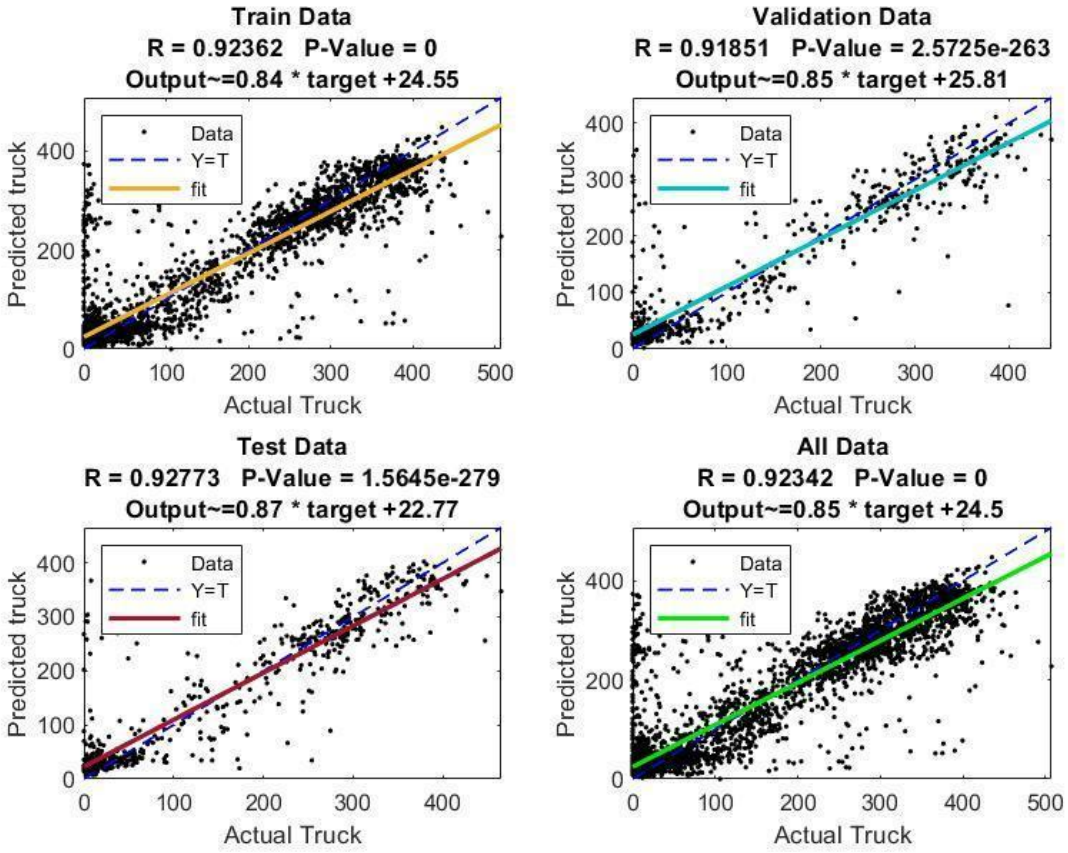


Figure 4-9: Correlation coefficient and linear fit between modeled and observed volumes.

Figure 4-10 compares the error histogram of the model for the three splits of the data. The closer the model error distribution lies to a normal distribution with zero means, the better the model performs. This figure indicates that the error for all partitions has a normal distribution with a mean close to 0 (range between -5.7 and - 2.9) and a standard deviation of around 55.

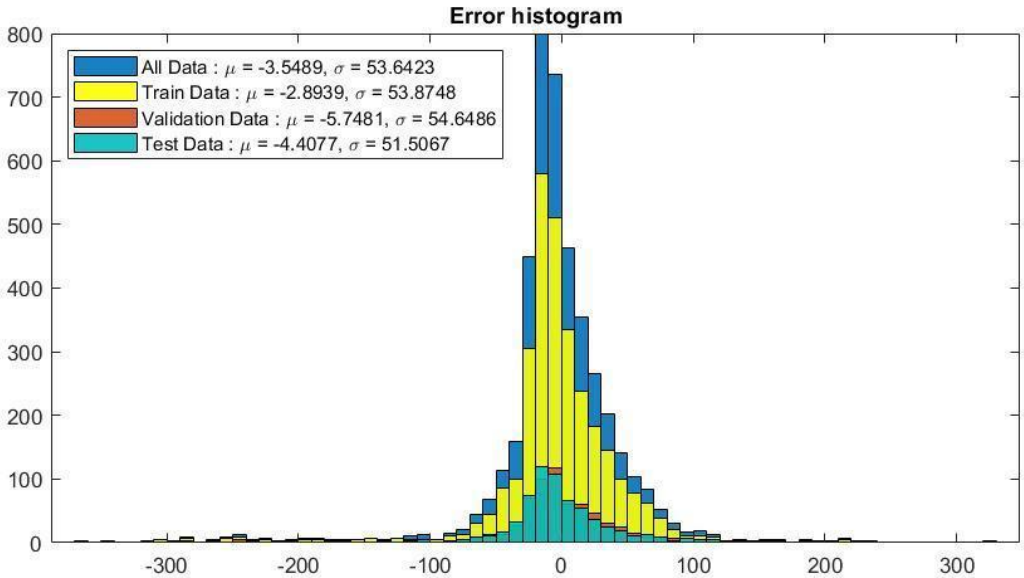


Figure 4-10: Error histogram of the model for the three splits of the data

To examine the error associated with the model prediction, we use three performance indicators (MSE, RMSE, and MAE) to compare the results for all the data splits. Figure 4-11 shows the model residuals as well as the results of the model performance indicators for every data split. The RMSE for the test data remains in the range of 50 to 55 for all the data splits. This means that, from the two-sigma empirical rule, approximately 68 % of the predictions have less than about 55 absolute error $\Pr(\mu - RMSE \le e \le \mu + RMSE) \approx 0.68$. On the other hand, the MAE of all splits shows that the error does not exceed 31.52 on average. The mean absolute percentage error is 3.2 % for the test data and not more than 4.9 % for other splits, which is close to other similar studies. This shows that the model has high accuracy in predicting truck flows. Unlike the MSE, RMSE, and MAE, a model with higher PAPE has better performance. This indicator shows 0.92 and 0.94 probability of error less than 10% error for validation and test data respectively.

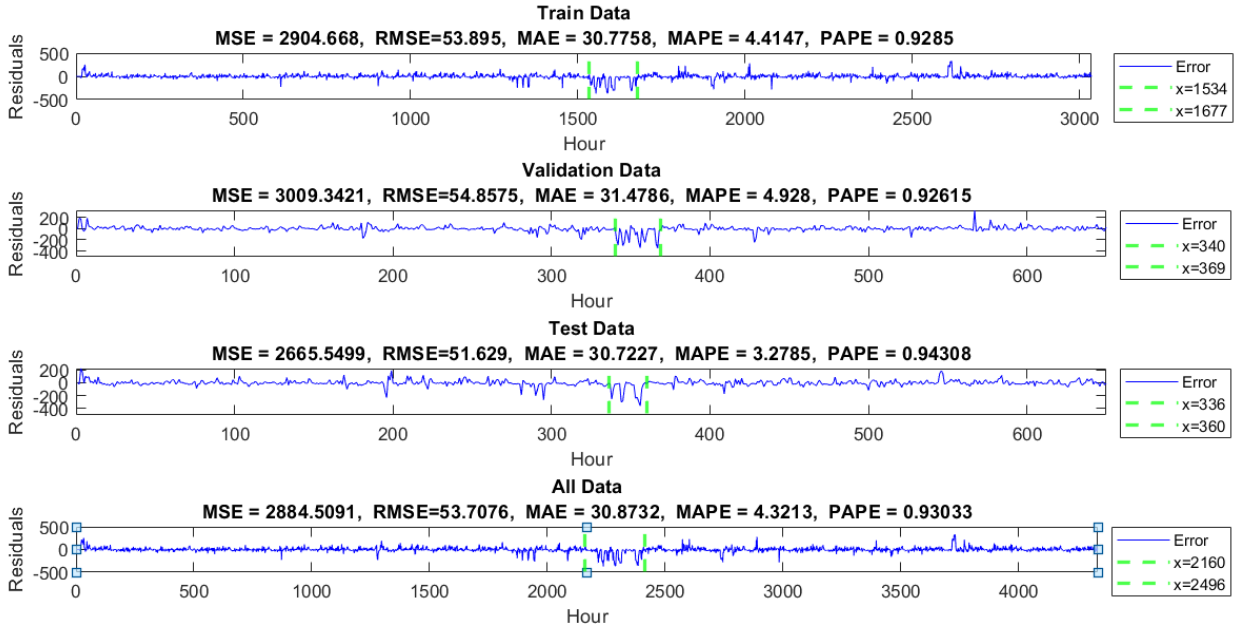


Figure 4-11: Models residual and performance indicators for three splits of the data.

All indicators show the high accuracy of the model despite the irregularities in the truck flow data. Figure 4-11 and Figure 4-12 show that most of the relatively large errors happen in one specific section of the time series. This period belongs to the first two weeks of April, where we indeed have irregularities in the truck counts (see Figure 4-3 - this could be because of a malfunctioning detector, for example). However, the model tries to generalize in that sense and that is why we get a higher error in that period. Having said that, the error profile of this model might also help us to detect irregularities in truck volumes.

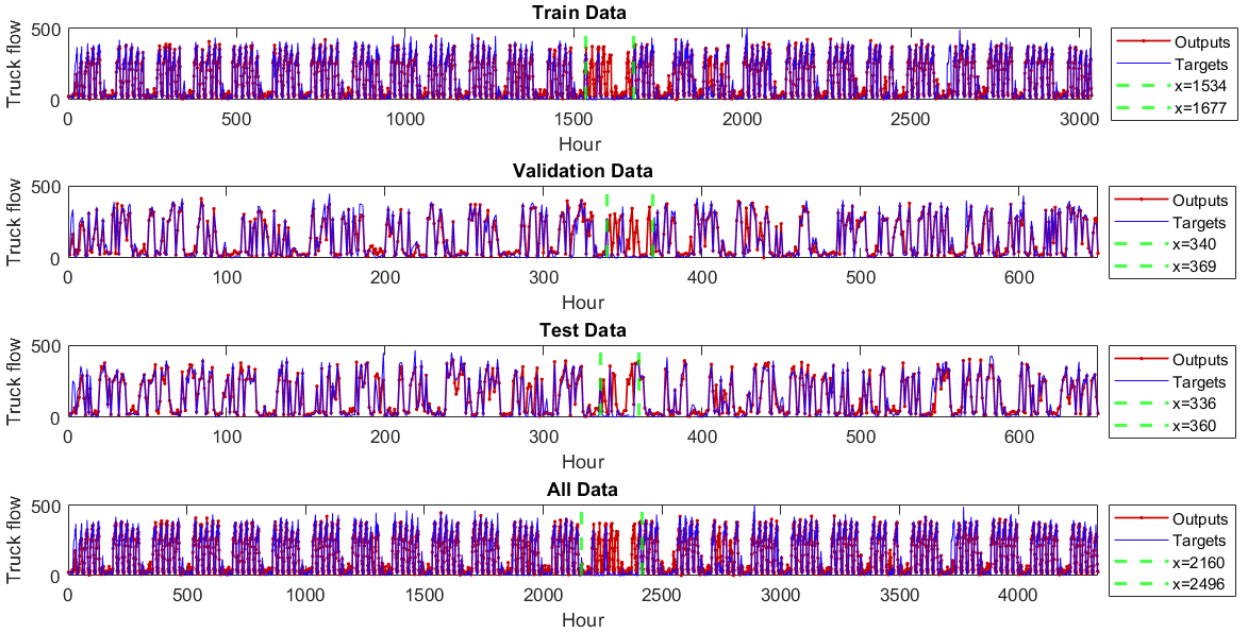


Figure 4-12: Comparison of model fit with observed values for three splits of the data

4.5.3 Out-of-sample validation

To evaluate the capability of our model to extrapolate, we used the month of November for out-of-sample validation as November is the month with the most complete data available in our out of sample data set. Figure 4-13 shows that even though the model is trained for the months of January to July, it can accurately predict truck volumes in November. The correlation between target and output is relatively high with R=0.96. The indicator $PAPE \cong 0.94$ indicates that only 6% of the predictions have more than 10% error. Finally, the absolute errors are 1.9 percent on average. All these indicators prove the strong capability of our model for forecasting out-of-sample data.

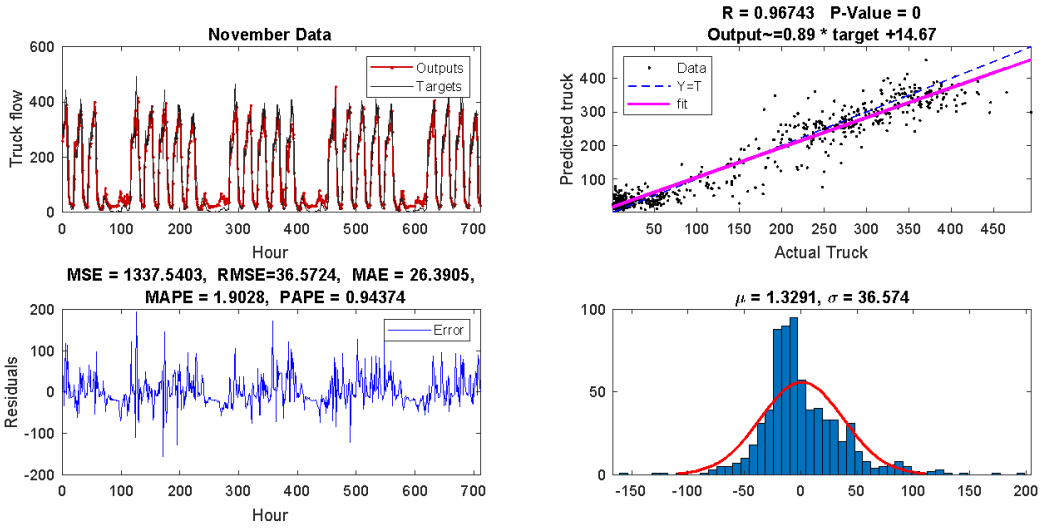


Figure 4-13: Out-of-sample truck traffic forecast for November 2017.

4.5.4 Temporal resolutions and methods comparison

In this section, we compare our proposed MLP-AFE model with the naïve historical average (HA) model, Least square Error (LSE), BPNN, and state-of-the-art LSTM network. We also compare different temporal resolutions used to predict from 5-min to 60-min ahead of truck volume. Among these methods, HA is a baseline method that only considers the average of the number of scheduled trucks in previous timesteps to predict truck volumes. For the LSTM network, we used the 'adam' algorithm for training 100 hidden units (by trial and error). This comparison is based on the predictions on the out-of-sample validation set.

Table 4-3: Prediction temporal resolutions and methods comparison

method	HA			LSE			BPNN			MLP-AFE			LSTM		
Horizon	RMSE	MAE	R	RMSE	MAE	R	RMSE	MAE	R	RMSE	MAE	R	RMSE	MAE	R
1-hour	116.5	90.4	0.58	153.5	110.7	0.26	65.2	40.9	0.88	36.6	26.4	0.96	36	23.1	0.96
45-min	76.74	57.8	0.71	58.21	40.1	0.83	49.4	30.2	0.88	37.3	21.9	0.96	21.3	14.9	0.98
30-min	46.27	32.8	0.78	39.55	27.42	0.82	31.2	18.3	0.89	28.2	15.8	0.92	19.4	11.8	0.96
15-min	21.88	14.7	0.81	20.67	14.40	0.81	18.7	11.8	0.85	15.5	9.2	0.90	9.1	6.23	0.96
5-min	8.57	5.5	0.73	8.16	5.78	0.74	7.4	4.90	0.80	6.1	3.7	0.86	5.4	4.04	0.87

The results in Table 4-3 shows that MLP-AFE has relatively higher goodness-of-fit as compared to the HA, LSE, and BPNN in all temporal resolutions. However, the goodness-of-fit drops when the prediction time horizon decreases. This is because the level of non-linearity increases in finer resolutions for both truck volume profile and container schedules. Moreover, in finer aggregation, the differences between the observed and predicted volumes cause larger relative errors which result in a drop in goodness of fit (Lv et al., 2014). Our model also proved to be as accurate as the state-of-the-art deep LSTM network, especially in 1-hour resolution. For finer temporal resolutions, however, LSTM performs slightly better which comes at the cost of a very complex structure and relatively long training phase. Nevertheless, MLP-AFE's goodness-of-fit and errors for finer resolutions are also promising, robust, and comparable with the LSTM network.

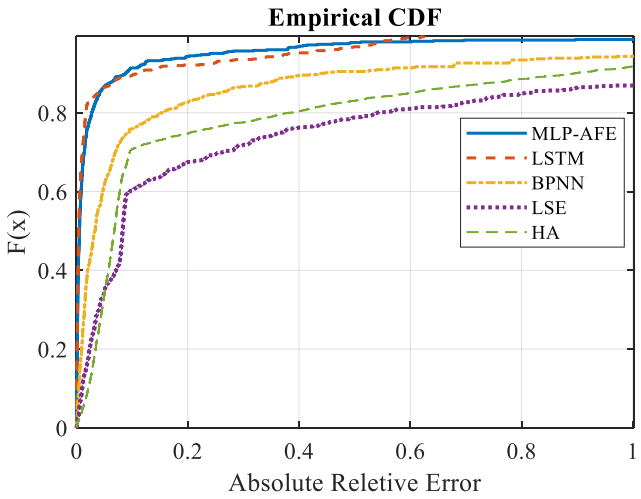


Figure 4-14: methods comparison in 1-hour resolution with CDF of absolute relative errors.

To compare for accuracy, Figure 4-14 visualizes the cumulative distribution function of the absolute relative errors for each model. It shows that, in our case, learning long-term dependencies with LSTM does not result in a measurable improvement in the predictive performance of the model.

4.5.5 Feature importance

As we saw in Table 4-2, our method proposes a model with 9 delays. This means that to predict truck volume at time t the best model requires container schedules with 0 to 9 delays (see Equation (4.1)). To see the relative importance of each of these delays in truck volume prediction, we used the permutation feature importance approach. In this technique, we used the trained model to predict on the dataset while one of the delays in the input layer is scrambled. We used the mean square error to calculate the relative score of the model with the scrambled delay. This process is repeated 10 times for each delay and then we used the average score of each delay as its final score. Figure 4-15 shows the average score of each delay.

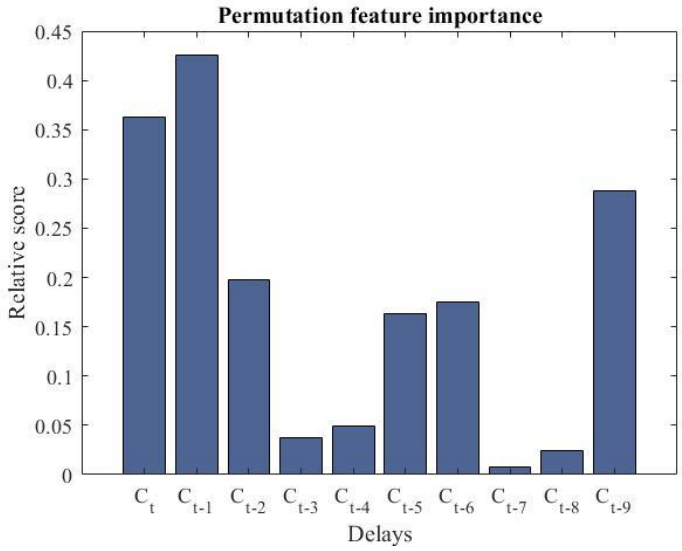


Figure 4-15: Global explanation of the relative importance of each delay on truck volume predictions

We can see, from Figure 4-15, that C_t, C_{t-1}, C_{t-2}, C_{t-9} are the most important features. Although Ct is in the same timestamp as T_t (see Equation (4.1)), C_{t-1} has more contribution to the prediction of T_t. This also confirms the observed delays between container schedules and truck actual pickups explained in section 4.5.1.

4.5.6 Scenarios

We further examine the model's ability to project changes in outbound truck volume by considering an increase in container throughput at the port of Rotterdam. An increase in container throughput can occur due to the arrival of large vessels. In addition, we have also analyzed the response of our model to inactivity at the port of Rotterdam, due to e.g. strikes or bad weather. As a consequence, we should observe different than normal truck volumes generated from the port area. Because of the strong underlying dynamics of the transshipment system, we cannot tell beforehand how the increase of truck flows will follow the growth of container volumes. We use the dataset for November 2017.

Effect of increase in container throughput on truck flow patterns

We consider five scenarios to forecast outbound truck volume, corresponding to a 5%, 10%, 15%, 20%, and 25% increase in the container throughput. In these scenarios, the hourly container demand is increased uniformly and we apply our model to retrieve the outbound truck counts. We report both the mean and the median hourly increase in outbound truck traffic as a result of this growth. The base case refers to the truck counts and container traffic observed in November 2017. In Table 4-4 we present the results. We notice a monotonic but non-linear increase in the outbound truck volume with respect to an increase in the container throughput (see Figure 4-16). Interestingly, both mean and median hourly increases are less than proportional until a 20% increase. For changes in container throughput above 20%, the mean flow increases more than proportionally. For forecasting, the median hourly increase provides us with a more stable indication of an increase in the generation of truck traffic, where the median growth is again less than proportional, but this time across the entire range of changes. Knowing that the terminals have some inventory facilities and that the container volume increase may be buffered for a while and then released over an extended period, this may explain the observed pattern. As the model takes the 9-hour delay, we believe it has the ability to predict this effect. In any case, the application of this model provides an estimate of the impact on flows of a sudden increase in the container throughput, which is non-trivial and original.

Table 4-4: Forecasted outbound truck in a growth scenario

Mean hourly increase in container throughput (%)	Mean hourly increase in outbound truck volume prediction (%)	Median hourly increase in truck volume prediction(%)
5	0.8	1.8
10	1.9	4.3
15	5.4	6.7
20	17.9	11.4
25	38.9	17.5

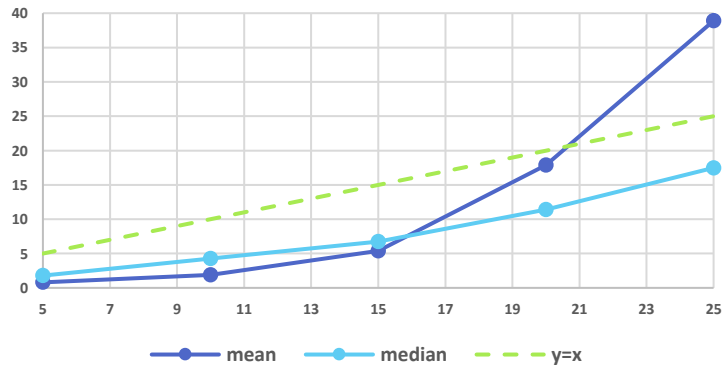


Figure 4-16: Mean and median trends in the forecasted outbound truck in a growth scenario

Effect of a period of inactivity at seaport terminals

Here we test the capabilities of our model to respond to random downward shocks in container throughput. Inactive container handling periods can occur due to unexpected circumstances such as strikes or bad weather. This may be relevant to spot traffic management opportunities in case of little expected freight activity, e.g. to free up road capacity that would otherwise be dedicated for trucks or to lift truck-specific bans on roads. We have considered two inactivity scenarios: weeklong and daylong events. For the former, we have assumed zero container throughput for the second week of November 2017. In the latter, we have assumed zero container throughput for the 8th day of November 2017. Figure 4-17 shows the performance of the model in terms of its sensitivity to respond to sudden stoppage of the container throughput. Despite the inactivity in terminals of 1 week, our model will still predict a movement of on average 21 trucks per hour. In the case when terminal activities are suspended for a day, the model has predicted 18 trucks on average generated by the port area during the inactive period. It implies that the model can identify some other minor necessary activities (i.e. existing distribution centers in port area) even if the terminals are inactive.

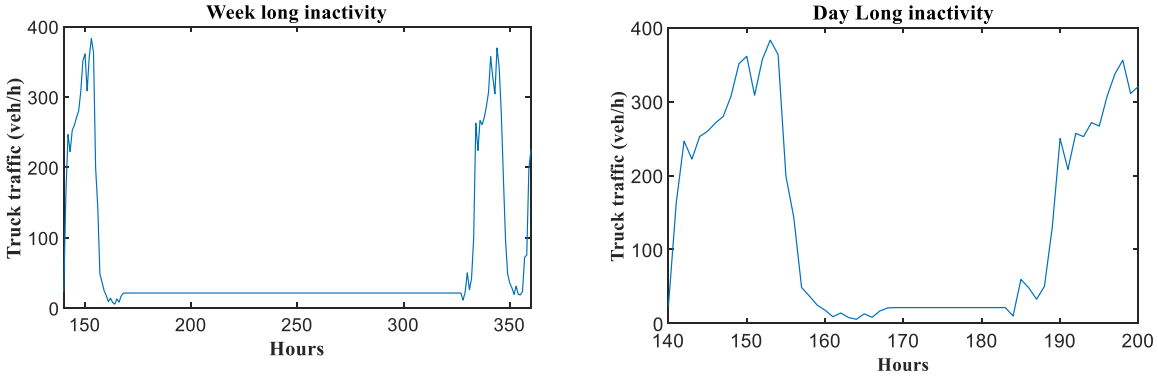


Figure 4-17: Forecasted truck volume (a) weeklong inactivity and (b) day-long inactivity at seaport terminals

In sum, the results of the two scenarios confirm the sensitivity of the model to different scales of variations in container schedules. Therefore, we believe that the model can be used for other seaports as well, as long as the schedule of containers for pick-up is available, for example via the port community system. We do recommend updating the parameters of the model through a transfer learning procedure for other ports. This will maintain the predictive ability of the model and reduce the risk of high errors.

4.6 Conclusion

In this paper, we explore the link between logistics activities at a seaport terminal and the truck traffic volume being generated by those activities. We use PCS data from five of the largest terminals in the port of Rotterdam for the first six months of 2017 to predict the next hour of outbound truck traffic volume, using a feed-forward neural network model. Our main conclusions are:

- We developed an analytical framework that enables us to (a) identify relevant feature vectors from a range of input data sources (pickup schedules, truck counts); (b) rank the resulting (in our case neural network) models in terms of how well they predict the outbound truck traffic volume.

- The final (best) model from our experiments can be used for short-term predictions of truck volume with reasonable accuracy on out-of-sample data this model achieved a 96% accuracy.
- We found that the predictive models are sensitive to the logistics activities at the port of Rotterdam. It can predict changes in the generation of truck traffic volumes as an effect of the dynamics of the handling containers at seaport terminals.

We believe these results are relevant for both science and practice. The main innovation of methodological nature is that we formulate the problem of feature selection and feature extraction as an optimization problem, and subsequently use NSGA-II -- not only to find the most important features, as commonly employed, but also to design the most appropriate topology for MLP. The model can be used to understand the consequences of logistics activities at the seaport terminal on the traffic system. The predictions can be useful for traffic management agencies to better manage traffic around major freight hubs.

Given our findings, a promising future direction of research is the development of the model within a simulation-based framework, to evaluate the impacts of departure time shifts (i.e., by changing the container pick-up schedules) on the traffic system. Future work could investigate the extensions of the model for network-wide prediction of truck traffic. This will require a model that also includes spatial correlation in combination with loop-detector data. Fast and accurate network-wide prediction of truck traffic can be useful to design more advanced traffic management strategies to further improve network reliability.

Chapter 5

Time-shifted operation for freight transport

In the previous chapter, We identified that the schedules of carriers have a strong impact on truck volume on roads. This chapter proposes a data-driven transport modelling framework to assess the impact of freight departure time shift policies on road networks. We develop and apply the framework around the case of the port of Rotterdam. Container transport demand data and traffic data from the surrounding network are used as inputs. The model is based on a graph convolutional deep neural network that predicts traffic volume, speed, and vehicle loss hours in the system with high accuracy. The model allows us to quantify the benefits of different degrees of adjustment of truck departure times towards the off-peak hours. In our case, travel time reductions over the network are possible up to 10%. Freight demand management can build on the model to design departure time advisory schemes or incentive schemes for peak avoidance by freight traffic. These measures may improve the reliability of road freight operations as well as overall traffic conditions on the network.

This chapter is based on the following paper:

Nadi, A., Sharma, S., van Lint, H., & Tavasszy, L., Snelder, M. (2020). A data-driven traffic modelling framework for analyzing the impacts of a freight departure time shift policy. *Transportation Research Part A: Policy and Practice*, 161, 130-150, DOI: https://doi.org/10.1016/j.tra.2022.05.008

5.1 Introduction

As heavy goods vehicles contribute significantly to congestion problems on inter-urban motorways, it is important to understand the interrelations between traffic and logistics systems. Also, predicting the impact of freight activities on road traffic conditions is a necessary ingredient of the design of peak avoidance traffic management policies for freight transport. Previous research has shown the effectiveness of departure time shift policies for passenger traffic (Thorhauge et al., 2016a, Thorhauge et al., 2016b) and urban freight deliveries (Sánchez-Díaz et al., 2017). Until now, however, optimization of freight departure time shift (FDTS) policies for inter-urban traffic impacts has not yet been presented in the literature. The main contribution of this paper is to help fill this gap, by developing a modelling framework that evaluates and optimizes the FDTS strategy using the prediction of short-term road traffic dynamics (i.e. flow, speed, and losses). To demonstrate the empirical working of the model we develop the case of container pickup in the port of Rotterdam.

The main questions addressed are :
1. Does FDTS lead to a significant gain for the traffic system?
2. How large is the monetary gain/loss for all network users (i.e. passengers and trucks)?
3. How large is the average contribution of every individual truck to this social gain?
4. what is optimal FDTS's from the perspective of overall traffic conditions?

Candidate modelling approaches that can provide answers fall into two main categories: firstly, models that use traffic theory and are known as model-driven approaches; secondly, the data-driven models, which use historical data to predict traffic conditions. In general, the methods each have arguments for and against and cannot outperform each other under all conditions (Calvert et al., 2015). The main difference between them is that the data-driven approach typically has parameters that have no physical interpretation, while a model-driven approach has fewer parameters where most do have a physical interpretation. For a comprehensive overview of these methods and their challenges for short-term traffic prediction, we refer to Vlahogianni et al. (2014), Van Lint and Van Hinsbergen (2012), and Poonia et al. (2018). In this paper, we propose a data-driven traffic model to predict short-term traffic states, taking the impact of freight demand into account. We choose data-driven models over model-driven approaches to reduce the calibration effort and to allow operation with short calculation times as well as lower computational complexity. The estimation of data-driven models can be performed offline and is easier than O-D estimation or calibration of road capacities. Data-driven models can follow the actual traffic state correctly, whereas simulation model predictions typically deviate from the actual traffic state (Calvert et al., 2015).

We build up the model around the case of port-related container transport, which marks the second contribution of the study to the literature. We use pickup schedules of containers in a seaport to predict short-term traffic dynamics on the surrounding road network, predicting volumes, speeds, and multi-class monetary losses in the traffic system. Besides the data-driven traffic model, we propose an approach to design optimal departure time shifts of container movements from the peak hours to off-peak periods. The shifts take place within the framework of a freight peak avoidance policy and concern the actual pickup time of containers from the Port of Rotterdam. We use the predictions from the data-driven model to optimize the scheme of shifts for the FDTS policy implication. A practical application of this framework can be a departure time advice system for road freight transport which can be useful both as a pre-trip

travel guidance tool for carriers and/or a policy assessment and decision support system for traffic managers. The main function that we will explore in this paper, is the social benefits that departure time shift as a freight peak avoidance policy may bring to the traffic system.

In summary, the original contributions of this paper are as follows:

1. We study the departure time shift for road freight and its consequences for an inter-urban freight corridor, complementing existing work on passenger traffic departure time shifts and on delivery time shifts for city logistics. Also, the empirical context of a maritime port is original.
2. This is the first paper that uses data about logistics activities, in our case container pick-up times, for a network-wide, short-term prediction of truck-intensive motorway traffic.
3. We introduce a graph-based, modular neural network, using novel message passing and neighborhood aggregation rules to capture spatial and temporal patterns in traffic.
4. This is the first modelling study of an optimized FDTS scheme linked to overall road network traffic conditions.

The remainder of this chapter is organized as follows: Section 5.2 provides a general overview of the existing studies about departure time shift policies and methods. Section 5.4 presents the model for short-term traffic prediction and the data-driven decision support system for FDTS. Section 5.5 denotes results, tests the predictive capabilities of the model, presents the designed scenarios, and discusses policy implications. Finally, section 5.6 offers the conclusions and recommendations of the paper.

5.2 Literature review

The departure time of commuters is believed to be one of the most important travel dimensions that play a significant role in reducing peak-hour congestion on road networks (Thorhauge et al., 2016a, Thorhauge et al., 2016b). Mahmassani and Jayakrishnan (1991) used a simulation model to assess the effects on the level of congestion in urban traffic under real-time in-vehicle information. Using this simulation, they implemented a choice of departure time. Based on their findings, peak spreading achieved by shifting vehicles' departure time, has considerable potential to reduce travel times under peak avoidance policies. Similarly, Yoshii et al. (1998) examined an application for a part of Tokyo and reported that "shifting departure times is more effective to reduce traffic congestion than switching routes, and the short degree of shifting time is enough to eliminate heavy traffic congestion". Based on these findings and similar studies, it became trivial that temporal demand spreading can improve traffic conditions. Therefore, researchers continued to offer different approaches to mitigate congestion on road networks concerning the departure time of travelers. Examples of such solutions are road pricing, parking pricing, departure time advice, and various incentives for behavioral change. Researchers mostly focused on passenger cars to assess impacts, rather than on trucks, probably because of their higher penetration rate. However, the impact of trucks can be major near logistic hubs where the percentage of trucks is high. For example, studies like Watling et al. (2019) show that fuel emissions and travel times of truck drivers are sensitive to their choice of departure time. Their findings indicate that certain departure time choices incur significantly longer travel times. Below, we review both passenger and freight transport DTS policies to cover the existing approaches.

Generally, peak spreading policies range between charging-based schemes, incentive-based schemes, and intelligent transportation systems (ITS) in the literature (Sánchez-Díaz et al. (2017)). Peak avoidance strategies like road pricing have long been proposed for internalizing

the external cost of road network congestion. These schemes typically apply charges to penalize travelers who pass through a congested area. Studies generally found that road pricing can produce considerable economic benefits (Eliasson, 2008). For example, Holguín-Veras et al. (2006) studied the impact of road pricing specifically on the behavior of freight carriers and commercial vehicles. The results showed that except for 9% of carriers, who pass on the charges to their customers, the majority of carriers change their behavior because of the pricing initiatives. In subsequent research (Holguín-Veras, 2008), the finding was that receivers respond weakly to price signals, as the charges transferred by the minority of carriers are too small to force a change. Although this condition could hold for shifting urban deliveries off-peak using prices, the case may be different for moving departures off-peak. From the departure time perspective, Zou et al. (2016) proposed an agent-based model for joint travel mode and departure time choice to evaluate congestion charging policies. They used a utility maximization approach along with a Bayesian learning process where agents can update their spatial and temporal knowledge and decide whether to search for alternative departure time and mode. They used this framework to assess the impact of congestion charging on travelers' mode and departure time decisions. Their simulation results showed that travelers switch mode and departure time under various congestion charging schemes when demand increases. The findings indicated that congestion charging is an effective way to mitigate traffic congestion.

Despite its reported benefits, there is public opposition to road pricing policies as taxes are an unpopular measure. Using rewards as a positive incentive to avoid peaks has been the subject of several experiments in the Netherlands (Ettema et al., 2010). In this experiment ("SpitsMijden", or peak avoidance in Dutch), participants received 3 to 7 euros per day if they could avoid traveling by car during peak hours. To assess the potential of these peak avoidance policies on traffic congestion, Bliemer and van Amelsfort (2010) developed a departure time discrete-choice model and traffic simulations to investigate travel time savings for different reward and participant levels. Their findings indicated that the travel time gain was largest for a 3 euro reward level when 50% of the drivers participated in the experiment. Other studies also investigated the flexibility or acceptance of different participants and explored the significance of different attributes that can explain the departure time choice of commuters under such an incentive-based departure time shift policy (Arian et al., 2018, Knockaert et al., 2012, Thorhauge et al., 2016a, Ben-Elia and Ettema, 2011). One of the challenges in the incentive-based departure time shift policies is the source of the funding. To cope with this problem, Holguín-Veras and Aros-Vera (2015) propose a freight demand management system in the context of urban freight delivery that uses pricing schemes to generates an incentive budget for receivers. These incentives influence the behavior of receivers choosing off-peak delivery times which, in turn, affects the carriers. In this study, they used a microsimulation approach to simulate behavioral interaction between carriers and receivers under an urban off-peak hour delivery policy. Further, they investigate the impact of this policy on the traffic system using a regional travel demand model and a mesoscopic traffic simulation model (Ukkusuri et al., 2016). Their results show significant improvements in congestion levels and overall network conditions. Although microsimulation and traffic models are promising tools to assess the impacts of given policies, their long runtimes make these tools impractical for optimizing policies, or for applications in a real-time traffic management context. In addition, simulation models are often calibrated on average traffic patterns and, therefore, predictions typically deviate from the actual day-to-day dynamics. Therefore, data-driven traffic models that are fast and accurate in capturing congestion patterns are often more practical for developing and applying ITS-based congestion alleviation policies.

De Boer et al. (2017) propose two data-driven methods to show the impact of having variable departure times on travel time reliability. A study for Amsterdam and surroundings in the Netherlands showed that large reliability improvements are possible after introducing variable departure time advice within the peak hours. Another example of an ITS-driven approach is a study by Calvert et al. (2015) that proposes a data-driven real-time travel time prediction framework for a departure time advice and route guidance system. Ma et al. (2009) propose a concept of departure time slot allocation to redistribute the demand over time slots at on-ramps to reduce congestion in the network. Minimizing system travel time and maximizing network utilization. They used non-linear programming and a genetic algorithm for this optimization problem. Findings from the simulation indicate that spreading demand optimally over time slots reduces the total travel time by 7.0% and reduces congestion by 8.9%.

All the mentioned studies have shown the importance of the departure time shift of passenger cars in congestion reduction policies. However, these investigations are relatively rare for inter-urban freight transport. Among the few existing studies that took truck activities into account, they consider the impact of travel time on the departure time choice of trucks - and not the reverse, which has our interest. Kleff et al. (2017) propose time-dependent route planning for truck drivers. They also consider the effects of departure time choice by incorporating time-dependent travel times. De Jong et al. (2016) estimated models to explain the time-period choices of receivers (e.g., producers, retailers, and wholesalers) in road freight transport. They have applied this model to assess the changes in the time-period choices of carriers. Their findings show that road freight carriers are relatively insensitive to changes in travel time. In contrast, they are more prone to avoid peak-hours if they sense an increase in transport costs (i.e., fuel cost, wage cost, loading, or unloading cost) in peak hours. They also note that "One reason that, even with heavy congestion in the peaks, not all goods transport take place off-peak is that (road) haulage companies want to use their trucks all times of the day". Beyond travel time and costs, Kourounioti and Polydoropoulou (2018) show that container characteristics and the receiver of goods are among the important factors affecting the time of day choice of trucks to pick up containers at terminals. Carriers, therefore, have multiple incentives for their choice of departure time. We will not go into this topic further but focus on the assessment of the impacts of changes in departure time choice.

In summary, while departure time shifts in passenger transport have been well studied, the case is different for inter-urban freight transport. Research on freight transport has highlighted city logistics off-peak deliveries or has focused on the reverse impacts, of congestion on departure time choice. In addition, traffic was not modeled explicitly or in a way that would allow us to seek optimal schemes from a traffic perspective. Ultimately, it is not clear which shifts in the departure times of trucks contribute most to improve traffic conditions. To help fill these research gaps we develop a data-driven approach to study the impact of departure time shifts in container transport on the traffic system. The next sections describe the data that we start from, for the development of the model.

5.3 Data

We use data from a seaport that indicates departure schedules of individual containers, as well as the network traffic data surrounding the port. Container schedule data for five major terminals operating in the Maasvlakte 2 region in the Port of Rotterdam are managed by the company Portbase. Data were provided for the year 2017. This dataset contains information about the seaside (i.e. vessel arrival time) and landside (i.e. container discharge time) handling of containers by terminal operators. The main field which we used in this analysis is the

estimated container Pick-Up Time for trucks. We aggregated data provided for each of these five terminals as one demand node or so-called centroid. For more information about these data, we refer you to a comprehensive exploratory analysis that has been done by the same authors on these data (Nadi et al., 2021).

In the Netherlands, the National Data Warehouse (NDW) provides a data stream of vehicle counts and traffic speeds collected from loop-detectors installed on motorways. The average distance between these loop detectors is 200 meters. A subset of these loop detectors can distinguish vehicle categories based on their lengths. These data are available at a resolution level of 1 minute time periods.

We collected time series of volumes and speeds for five motorways near the port of Rotterdam. Table 5-1 and Figure 5-1 show the characteristics of this network. Data were collected for 6 months (181 days) from January 1th to June 30th in the year 2017.

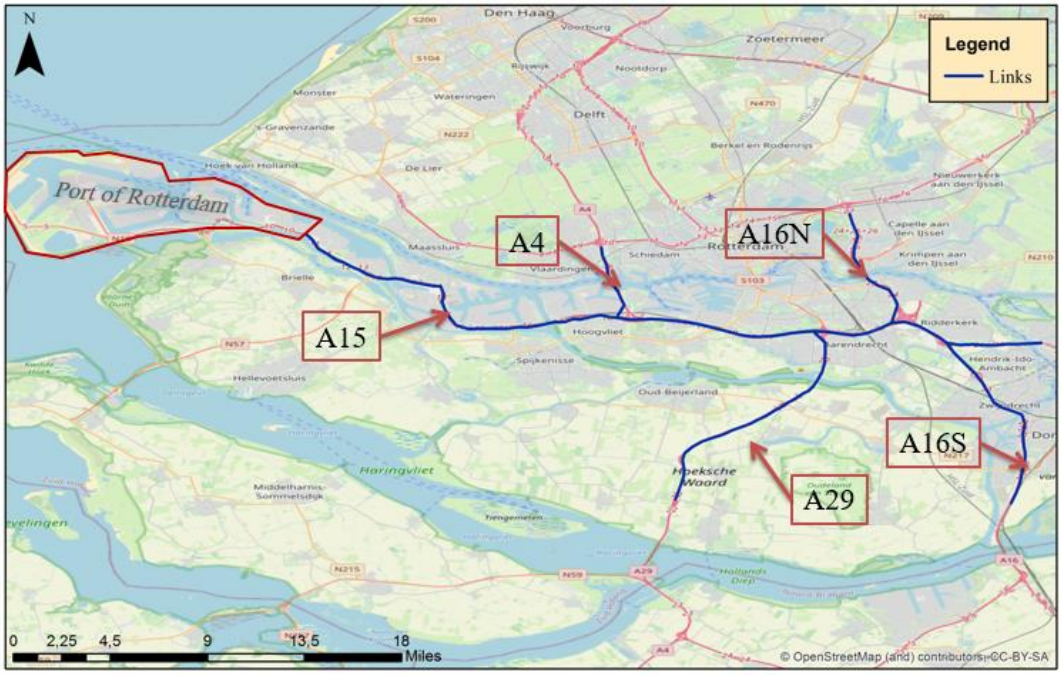


Figure 5-1: The road network that provides accessibility for the Port of Rotterdam and Rotterdam city

Table 5-1: Characteristics of the port-hinterland motorway network

Link	Direction	Sensors	Nodes	Truck sensors	Section length (m)	Start point (km)	End point (km)	Length (km)
A15	East	270	90	24	600	25	80	55
A4	North	24	8	3	600	75.4	70.6	4.8
A29	South	60	20	9	600	10.8	22.2	11.4
A16	North	39	13	3	600	23.6	16.5	7.1
A16	South	105	35	5	600	28.2	49	20.1

To reduce the computational burden, we aggregated data somewhat in the space dimension from 3 consecutive loop detectors in such a way that the law of conservation of vehicles holds (i.e. number of nodes in Table 5-1 equals the number of sensors divided by 3). We also

aggregated in time to 5 minutes intervals (i.e. $\Delta t = 5$). Besides loop detector nodes, we also have one truck demand node which is the port of Rotterdam. Altogether our data allows a model structure consisting of a graph of 166+1 nodes, to form a delayed feed-forward neural network, to jointly predict 3 characteristics of traffic: speed, flow, and monetary loss of all vehicles. To deal with the noisy speed and flow data, we used the adaptive smoothing method (ASM), as developed by Treiber and Helbing (2003) and further improved by Schreiter et al. (2010). After these measures, 35 of the 181 days were left with a high number of missing values; these were removed from the analysis to avoid risking bias from imputation.

5.4 Methodology

In this section, we describe the data-driven forecasting model and its application within the context of a freight peak avoidance policy, to help set the right parameters for a peak avoidance scheme. This decision support framework is pictured in Figure 5-2, and consists of two main modules: (1) a data-driven traffic forecasting model and (2) a predictive departure time control model.

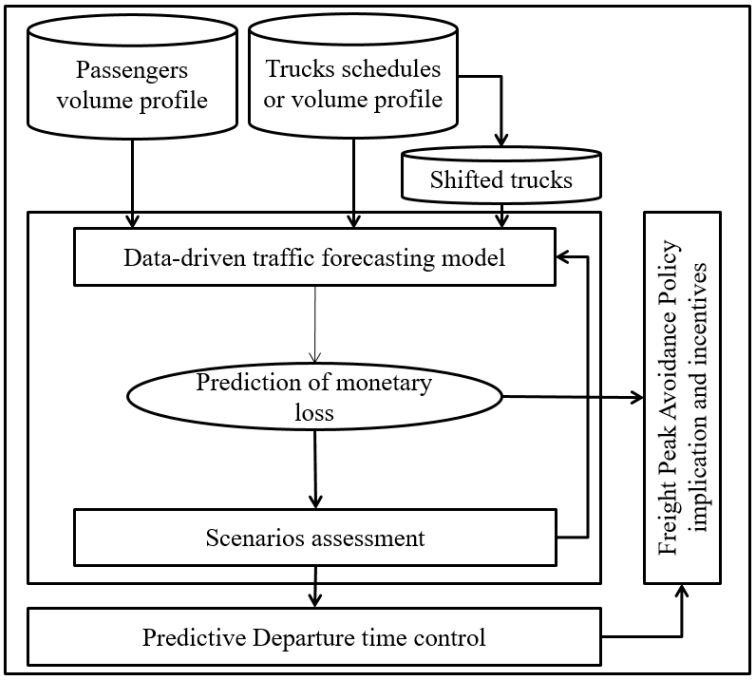


Figure 5-2: Data-driven decision support framework for freight departure time and peak avoidance policy

In the following subsections, we formalize the data-driven traffic forecasting problem and describe in detail how to model the spatial-temporal dependency structure using a graph-based modular recurrent neural network. Then, we elaborate on the departure time shift control model which uses a search heuristic to solve the combinatorial assignment problem.

5.4.1 Data-driven traffic forecasting model

Short-term data-driven traffic models aim to predict traffic dynamics (e.g. speed, volume, delays) a few minutes into the future given a set of past observed traffic features. The traffic nature is spatially correlated and highly depends on the structure of the road network.

Additionally, temporal interaction between various locations as well as the contribution of demand nodes (trip generation centroids) on a road network is, to a certain extent, very complex. To cope with this complexity and to make the model as close as possible to the physical properties and composition of a real traffic system, we use an artificial neural network with a graph representation. We represent the loop detectors as weighted directed graphs *G(V, E, A)* where *V* is a set of nodes ($V \in \mathbb{N}$) representing sensors on motorways, *E* is a set of edges representing road segments and $A \in \mathbb{R}^{N\times N}$ is a weighted adjacency matrix that represents node connectivity as well as proximity. Let $X \in \mathbb{R}^{N\times M}$ be the signal input of the graph *G*, where *M* is the number of features of each node (e.g. volume, speed) and *N* denotes the number of nodes. Also, let $C \in \mathbb{R}^{n}$ represent the signal input of the truck demand generation centroids connected to $n \in N$ nodes in the graph *G*. Having X^t and C^t representing observed signals at time *t*, this model aims to learn a generalized non-linear approximation function $f(.)$ for each node in a given graph *G* that maps $d \in \mathbb{N}^{N\times(N+n)}$ historical input signals to the future signal $X^{t+\Delta t}$.

$$X^{t+\Delta t} \cong f(X^{t}, X^{t-\Delta t}, X^{t-2\Delta t}, \dots, X^{t-d\Delta t}, C^{t}, C^{t-\Delta t}, \dots, C^{t-d\Delta t} | G) \qquad (5.1)$$

Figure 5-3 illustrates the graphical representation of a sample road network in this formulation. In a modular graph-based neural network scheme, each node in this graph can be a single layer or a fully connected feed-forward neural network where nodes can pass input and output to each other and predict the desired target variable.

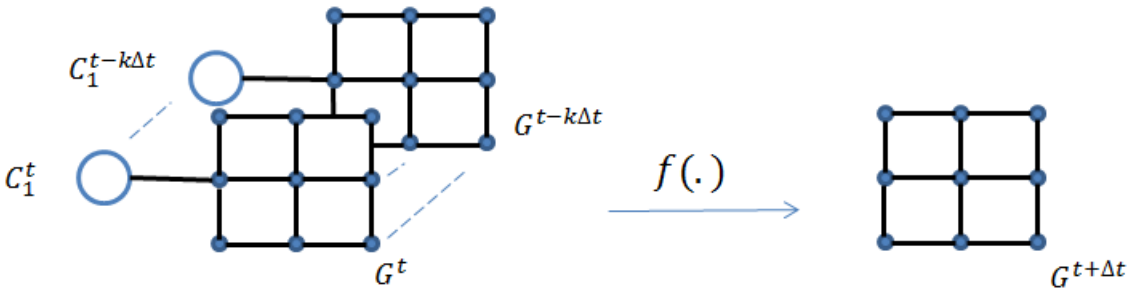


Figure 5-3: An example road network with one centroid in a graph representation.

It's important to mention that using the graph to model the structure of the road network helps to prevent violating the law of conservation. First, we made sure that all the three aggregated consecutive loop detectors are either before or after the on-ramps and off-ramps. Second, the off-ramps and on-ramps are all considered as a single node in the graph and are well connected to the other nodes to ensure that the inflow and outflow to the road network match well. Third, as Figure 5-4, shows, information and flow data from one node in a graph are aggregated and passed to its neighbor nodes via the adjacency matrix. The adjacency matrix keeps the conservation law valid in the graph because the output, i.e. the prediction of flow from one node, is an input to its neighbors.

Spatial and Temporal dependencies

We model the spatial dependencies by using graph convolution, message passing, and neighborhood aggregation techniques. Graph convolution filters have recently gained special attention for short-term traffic prediction (Li et al., 2021). These techniques compute the input

feature of each node as an aggregate of the features of its neighbors before passing it to a hidden layer of a fully connected neural network (Zhou et al., 2018). One of the best propagation rules for aggregating each node's feature is proposed by Kipf and Welling (2016).

$$conv^* = \bar{D}^{-0.5}\tilde{A}\bar{D}^{-0.5}\bar{X}W \tag{5.2}$$

$$D_{ii} = \sum_j \tilde{A}_{ij} \tag{5.3}$$

$$\tilde{A} = A + I \tag{5.4}$$

where D is the diagonal degree matrix, I is the identity matrix and $\tilde{A}$ is the adjacency matrix with added self-loops. These are added because we also need to consider the features of the node itself as well as its neighbors. Finally, W is a matrix of trainable weights. In other words, the convolution operator $conv^*$ in Equation (5.2) computes the aggregate (i.e. normalized weighted sum) of features for all nodes. With this symmetric normalization, we not only take into account the degree of the i^{th} node, but also the degree of the j^{th} node. We have added to this convolution rule in Equation (5.2) by introducing a new attention mechanism that produces a dynamic k-order weighted adjacency matrix. In this new formulation, some nodes can benefit from the information coming from their second or third-order neighbors. Therefore, we initially define a k-order adjacency for all sensors weighted by their distance. In other words, the closer nodes to the i^{th} node get higher weight in the adjacency matrix. We assume that these distance-related weights in the adjacency matrix are random variables with Gaussian distributions.

$$S_{ij}^n = \exp\left(-\frac{1}{2}\left(\frac{p_{ij}}{\sigma_i^n}\right)^2\right) \tag{5.5}$$

where p is the node proximity matrix and, for each node i, the parameter σ_i^n needs to be estimated globally along with the local training process. We name this parameter σ^n as a spatial memory that helps the neural network remember the spatial dependency of various locations on the road network. This parameter also identifies the order of adjacency for each node. When σ_i^n for the i^{th} node is small, the node i gets information from its close-by nodes. In Equation (5.5), S_{ij}^n are elements of the matrix S^n that we name as the node attention matrix. This process is a new type of attention mechanism that can adaptively capture spatial correlations on the road network. This operator will be multiplied to the adjacency matrix $\tilde{A}$ (element-wise) and then will be fed to the convolution operator for aggregation of input features. Equation (5.6) shows graph convolution operator with node attention on $\bar{X}$ input features.

$$conv_X^* = \bar{D}^{-0.5}(S^n \otimes \tilde{A})\bar{D}^{-0.5}\bar{X}W_X \tag{5.6}$$

Where W_x is learnable parameter for convolution operator on x inputs. Since we use a directed multivariate graph, we can apply the aggregation rule on all or some specific features and in any desired direction. For instance, this operator can aggregate information of previous nodes (i.e. upstream) for the volume and aggregate information of the next nodes (i.e. downstream) for the speed features. This bidirectional aggregation helps us to capture spillback phenomena once we have congestion on a link (see Figure 5-4).

In addition to the input features $\bar{X}$ in above equations, the truck demand input signal $C \in \mathbb{R}^n$ generated from each centroid should also be provided as input to the network. We introduce a similar Attention mechanism as in Equation (5.6) with spatial memory σ^c to force all the demand nodes to be the k^{th} neighbor of all sensors. We also named this mechanism the centroid attention S^c with which we can adaptively capture correlations between centroids and nodes on the road network.

In the case of $m \in\{1,2,..M\}$ demand nodes, We also added a softmax decision gate E_{mi} that can decide to what extend adding inputs c_m from a demand node m can improve predictions in a specific node i on the road network. Then, we apply the estimated weights, explained in the previous paragraph, to the selected demand nodes.

$$E_{mi} = \frac{e^{c_m}}{\sum_{j=1}^{M} e^{c_j}} \qquad \forall i \in N \tag{5.7}$$

Equation (5.8) shows the graph convolution operator with centroid attention on $\bar{C}$ input features.

$$conv_C^* = (\bar{E}S^c \otimes \bar{O})\bar{C}W_c \tag{5.8}$$

Where $\bar{E}$ is a matrix of decision gates with elements of E_{mi}, $\bar{O}$ is a matrix of all ones (because centroids are initially considered as the neighbor of all nodes in the network), S^c is centroid attention matrix, $\bar{C}$ is the demand generated from centroids, and finally W_c is learnable parameter for convolution on centroids inputs. Please note that although the centroids are considered to be the neighbor of all the nodes in the road network, the decision gate matrix E will automatically decide the normalized contribution rate of input data from a specific centroid on the predictions of traffic at a given node on the road network. In other words, the centroid attention layer helps the network to predict the destinations of the demand generated at each specific centroid.

Finally, the adjusted inputs are passed to g with is an activation function of a hidden layer or a fully connected feed-forward network. In this formulation, both attention mechanisms let the network pay relatively more attention to the most valuable information coming from different nodes and centroids.

$$f(\bar{X}, \bar{C} \mid G) = g(conv_X^* \times \theta_1 + b_1, conv_C^* \times \theta_2 + b_2) \tag{5.9}$$

Equation (5.9) indicates the final structure of a trainable block including all layers and a fully connected feedforward network where θ_1 and θ_2 are learnable parameters, b_1 and b_2 are bias terms, and $f(\bar{X}, \bar{C}|G)$ is the output of the layer given input $\bar{X}$ and $\bar{C}$ on graph G.

For the temporal dependency, we considered a delay matrix $d \in \mathbb{N}^{N\times(N+n)}$ (see Equation (5.1)). For each node in the graph (i.e. row of the matrix d), we initially fill each element with a random temporal delay related to each of the centroids and k-order neighbors. This delay needs to be estimated globally along with the internal training process of the neural network. The next section explains how these delays are considered to adjust the inputs before being fed to the attentions and then convolution layers.

Graph Neural network structure design and model specification

The eventual structure of the graph-based modular neural network model depends on the topology of the road network. For this study, we aim to predict the impact of truck demand generated from the port of Rotterdam on the surrounding 5 motorways to a certain threshold distance (see Figure 5-4). Every loop detector in this network can measure the speed and flow of all vehicle types at every minute. We also have loop detectors that measure specifically truck volumes and speeds.

Figure 5-4 shows how nodes in the convolutional graph are stacked together and how the output of one layer is incorporated into the inputs of the next layer.

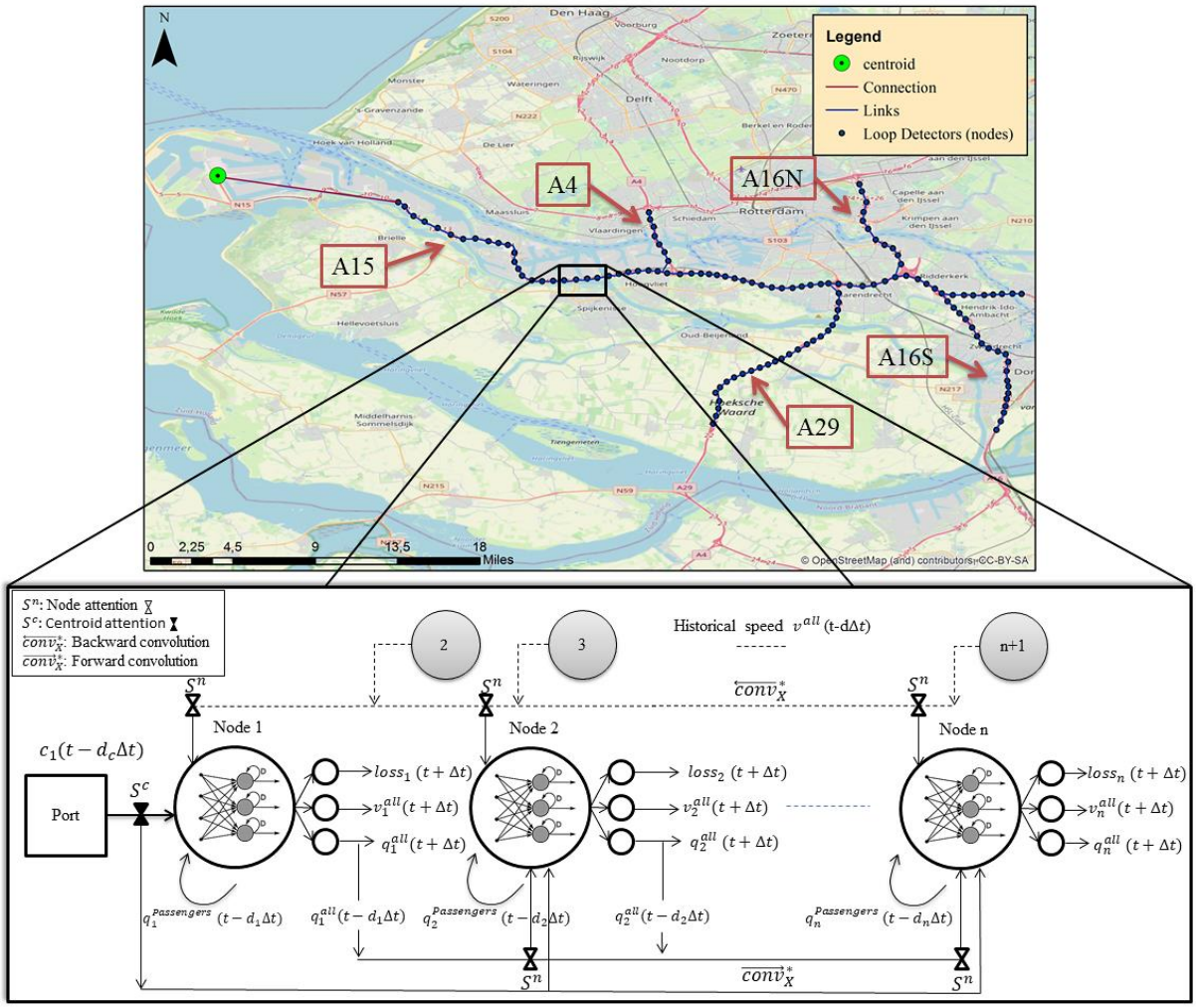


Figure 5-4: Graphical representation of the structure of the proposed graph-based modular convolutional neural network model

The prediction in each node in this graph-based model is based on the following input signals:

- X^o: Adjusted demand input which is delayed container pickup demand generated in centroids. In the case of more than one centroid, all centroid input signals have to pass first through the centroid attention and then the convolution layers. Otherwise, it can be directly connected to other nodes in the graph.

$$X^0 = conv_c^*(C(t-d_c\Delta t)) \tag{5.10}$$

Where C is the container pickup demand generated in centroids and d_c is the delay in the input signal.

- X^1: Adjusted volume data which is a delayed weighted aggregation of predicted volumes in previous nodes calculated through an upstream forward spectral convolution layer $\overrightarrow{conv}_X^*$,

$$X_j^1 = \overrightarrow{conv}_X^*(q_j^{all}(t-k\Delta t)) \qquad \forall j \to i \tag{5.11}$$

- X^2: Adjusted volumes of non-truck vehicles which is delayed and passed through a forward spectral convolution layer,

$$X_i^2 = conv_X^*(q_i^p(t-d_i\Delta t)) = \overrightarrow{conv}_X^*(q_i^{all}(t-d_i\Delta t) - q_i^{truck}(t-d_i\Delta t)) \tag{5.12}$$

- X^4: the adjusted speed data which is a delayed weighted aggregation of speeds in the next nodes in downstream passing through the backward spectral convolution layer $\overleftarrow{conv}_X^*$.

$$X_j^4 = \overleftarrow{conv}_X^*(v_j(t-k\Delta t)) \qquad \forall i \to j \tag{5.13}$$

As far as monetary losses in the system are concerned, we design this model to be able to predict the monetary value of vehicle loss hours regarding variations in speed and flow. As we can see in Figure 5-4, this model is designed in such a way as to learn multiple tasks at the same time. Therefore, the prediction of monetary losses in the system takes place regarding the joint features extracted to learn the dynamics of speeds and flows. In a conventional neural network scheme, the concatenated inputs in each hidden layer pass through a transition function g (e.g. *logsig*) that projects these inputs onto an m-dimensional space. This m-dimensional space is a representation of features learned from the input data to predict outputs. These dimensions depend on the size of the hidden layers (m=number of neurons in a layer). In our multi-task learning framework, we transfer these learned features of one target variable to predict two other target variables. In this study, we used a heuristic trial and error to set the number of hidden layers (i.e. two layers for each node) and neurons (7 for the first layer and 6 for the second layer). The activation function for hidden layers is *logsig* and for the output layer is *purelin*.

To calculate real-world vehicle loss hours, we use real-world speed and volume as is presented in Equations (5.14) to 15.22. Then the calculated monetary loss hours are added to the prediction model. We could predict speed and volume and then calculate the vehicle loss hours using the predicted speed and volume. However, we argue that prediction of speed and volume, even in the case of high accuracy, each comes with some errors. Calculating the vehicle loss hours using the predicted speeds and volumes will, in some cases, incorporate and consequently amplify the errors in loss hour prediction. That's the reason why we decided to calculate the losses a priori from the real-world data and then provide it to the model so that we can have a better prediction.

Trucks and passengers are two classes of vehicles in this system. Loop detectors (i.e. nodes in the road network Graph) can measure traffic either for trucks or a mixture of vehicles (i.e. trucks and passengers together). Each node *i* in the road network graph measures traffic on a section of road with length *l* at each timestamp T={1,2,…,t}. To calculate vehicle loss hours, we need to compare the current situation of a section with that of its free-flow speed. Please note that vehicles, especially passengers, can drive at a wide range of speeds in free flow situations. Based on the Highway Capacity Manual, chapter 12, (Reilly, 1997) the practical definition of free-flow speed for a multilane highway is the average speed of vehicles in a free flow situation. The free flow situation is where a section is uncongested (low density) and when the flow rate is low to moderate (between 0 and 1400 passenger car/hour/lane). Based on the field measurements in our study area, we estimate the free-flow speed for trucks and passengers equal to 80.983 and 101.9414 km/h respectively. These free-flow speeds are calculated based on the space mean speed of vehicles of each class in the free flow situation. As you see, these figures are slightly above the speed limits (80 km/h for trucks and 100km/h for passenger cars) of the study area. In this paper, we use free-flow speeds to calculate the costs and benefits of the time-shifted policy in the system. However, It is not acceptable for policymakers to use free-flow speeds above the legal speed limits as it's against the strict enforcement of police regulations. We, therefore, adjust free-flow speeds to the maximum legal speed to be used as the reference for the gain calculation. Please note that using this model in other applications may require the use of the free-flow speeds calculated from field measurements without this adjustment. Equations (5.14) to (5.22) show how we calculate the monetary loss for each node *i* at time *t* in this multi-class system.

$$tp_{i,t} = \frac{q_{i,t}^{Trucks}}{q_{i,t}^{all}} \qquad \forall i \in V \quad , \quad \forall t \in T \tag{5.14}$$

$$q_{i,t}^{Passengers} = (1 - tp_{i,t}) q_{i,t}^{all} \tag{5.15}$$

$$Vh_{i,t}^{Truck} = \frac{q_{i,t}^{Trucks} l}{v_{i,t}^{Trucks}} \tag{5.16}$$

$$Vh_{i,t}^{Passengers} = \frac{q_{i,t}^{Passengers} l}{v_{i,t}^{all}} \tag{5.17}$$

$$Vh_{free}^{Trucks} = \frac{q_{i,t}^{Trucks} l}{v_{free}^{Trucks}} \tag{5.18}$$

$$Vh_{free}^{Passengers} = \frac{q_{i,t}^{Passengers} l}{v_{free}^{Passengers}} \tag{5.19}$$

$$VLH_{i,t}^{Trucks} = \begin{cases} Vh_{i,t}^{Trucks} - Vh_{free}^{Trucks} & Vh_{i,t}^{Trucks} > Vh_{free}^{Trucks} \\ 0 \quad otherwise \end{cases} \tag{5.20}$$

$$VLH_{i,t}^{Passengers} = \begin{cases} Vh_{i,t}^{Passengers} - Vh_{free}^{Passengers} & Vh_{i,t}^{Passengers} > Vh_{free}^{Passengers} \\ 0 \quad otherwise \end{cases} \tag{5.21}$$

$$Loss_{i,t} = VoT^{Passengers} \cdot VLH_{i,t}^{Passengers} + VoT^{Trucks} \cdot VLH_{i,t}^{Trucks} \tag{5.22}$$

Where $q_{i,t}^{all}$ are $q_{i,t}^{Trucks}$ are the number of vehicles per hour passing the section (node) *i* at time *t* collected by loop detectors for each category of vehicles i.e. all vehicle types and trucks respectively. Please note that we calculate the number of vehicles per hour for passengers i.e. $q_{i,t}^{Passengers}$ using $tp_{i,t}$ which is the percentage of trucks in node *i* at time t (see Equations (5.14) and (5.15)). For each timestamp and node in the road network graph, Equations 5.16 and 5.17 compute vehicle hours for trucks $Vh_{i,t}^{Trucks}$ and passengers $Vh_{i,t}^{Passengers}$ respectively. Vehicle-hours is a flow-weighted travel time or, in other words, the number of vehicles multiplied by the number of hours they have driven to pass a section of a road. Consequently, Equations 5.18 and 5.19 calculate the vehicle hours for each class of vehicles under the free flow conditions. Then class-specific vehicle-loss-hours is the deviation between vehicle-hours in current and free flow conditions (see Equations (5.20) and (5.21)). Finally, the *Loss* indicates the space-time monetary loss matrix. To calculate the monetary value of losses in the system, we use the value of time for each of the classes (i.e. $VoT^{Truck} = 45$ Euro for trucks and $VoT^{Passengers} = 10$ Euro for passengers). The loss is then the summation of vehicle loss hours for passengers ($VLH^{Passengers}$) multiplied by the value of time for passengers and vehicle loss hours for trucks (VLH^{Truck}) multiplied by the value of time for trucks.

Model estimation

In the previous section, we presented a novel graph-based modular recurrent convolutional neural network for our data-driven traffic model. We used MATLAB 2019b academic to implement and configure the model. This model contains two types of internal and external parameters. The internal parameters θ, b are the weights and biases of each connection between neurons of one layer and another layer in fully connected feedforward networks. These parameters represent the features learned to map inputs to output space. The external parameters are those which are introduced in Equations (5.5) to (5.8) to control the temporal and spatial dependencies. Error backpropagation is an approach that is often used to estimate the internal parameter of neural networks. Initiating with random weights and biases, the training process continues with successive updating weights and biases in a way to minimize the total error of the model. To estimate the external parameters, one approach is to design a neural network layer to simultaneously predict external parameters and control the spatial and temporal gates along with estimating internal parameters. This approach is the core idea behind the design of deep LSTM neural networks. The advantage of this approach is that a built-in optimization algorithm can be utilized to estimate both external and internal parameters. However, this approach immensely increases the number of parameters in the model which brings new computational challenges. In this paper, we propose a bi-level model estimation approach. In the lower level, we used the Levenberg-Marquardt (LM) algorithm to estimate internal parameters. This algorithm uses the Jacobian matrix, the first derivatives of the network error, to approximate the Hessian matrix. This algorithm is known to be the fastest method to train networks with not more than a few hundred weights (Hagan and Menhaj, 1994). It makes this algorithm a good candidate for our bi-level network parameter estimation. Additionally, LM uses a validation set to avoid overfitting. Validation sets are used to test the network during

training. The training process stops if the performance of the network fails to improve for a predefined number of successive tests.

In the upper level, we estimate external parameters of the model e.g. spatial memory and delays. We used a GA algorithm to tune the external spatial and temporal memory gates. To see how the structure of a neural network can be optimized by the genetic algorithm we refer readers to Vlahogianni et al. (2007). Figure 5-5 shows how the model's parameters are estimated in this bi-level parameter estimation.

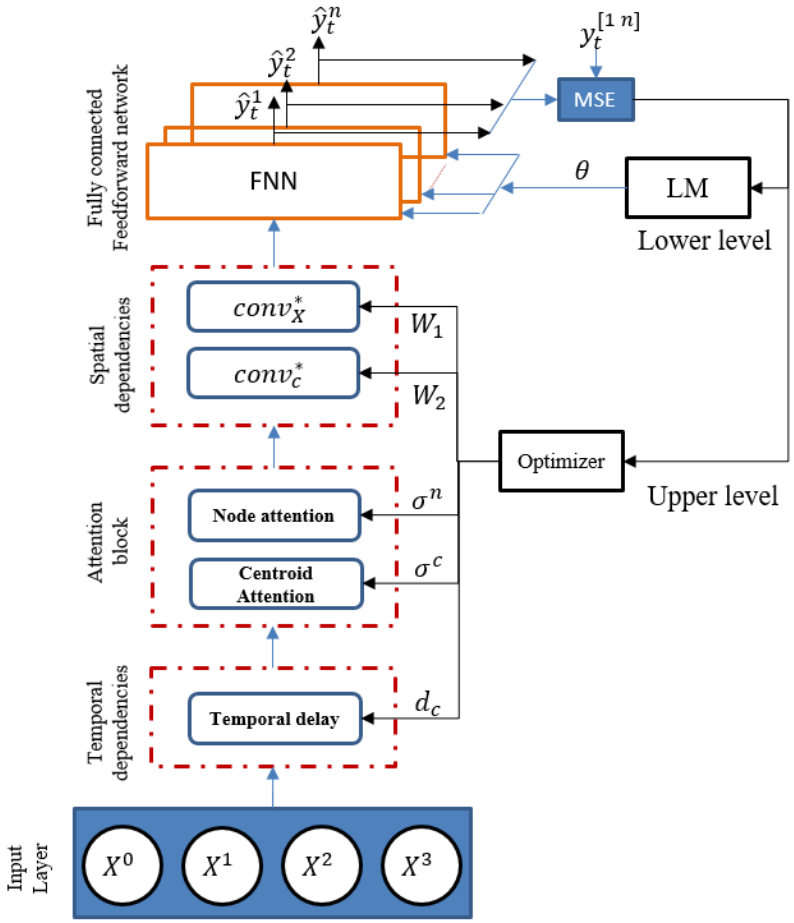


Figure 5-5: High-level presentation of network structure and parameter estimation procedure

We used the mean square error (MSE), the most commonly used error function, for estimating the model weights and parameters.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{5.23}$$

Where N is the number of observations, y_i is the i^{th} observed value and $\hat{y}_i$ is its corresponding predicted value.

Model evaluation

Here we use a set of indicators like root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and probability of absolute percentage error (PAPE) to evaluate the performance of the model.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2} \tag{5.24}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\|t_i - y_i\| \tag{5.25}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{t_i - y_i}{t_i}\right| \tag{5.26}$$

$$PAPE = \Pr\left(\left|\frac{t_i - y_i}{t_i}\right| < 5\%\right) \tag{5.27}$$

The RMSE demonstrates the standard deviation of the residuals which indicates how the data are concentrated around the best fit. The MAE represents how big an error could be on average. We also use the correlation coefficient between observed and predicted values to evaluate the predictability of the model.

5.4.2 Departure time control

After training the model, we need a procedure to control departure time shifts considering the most likely scenarios. This procedure helps us to systematically design FDTS scenarios. In this section, we formulate this procedure as an assignment problem to reschedule the departure time of candidate containers in such a way as to minimize the total loss in the system, while also minimizing the deviation between the scheduled departure time and the recommended departure time.

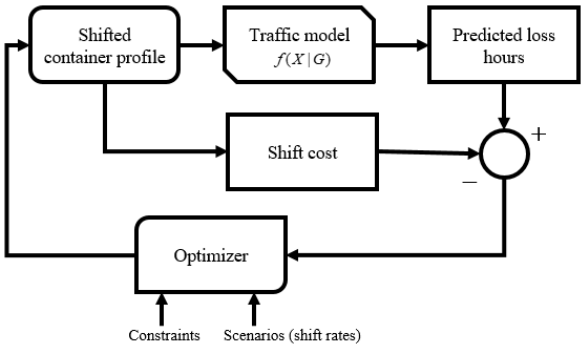


Figure 5-6: A Graphical representation of departure time control procedure.

Through this formulation, we place a terminal capacity constraint on this FDTS shift to prevent possible queues at gates. This also prevents unexpected sharp increases in container flow which the model has never seen during the training process. The algorithm that we propose to solve this problem has the following steps:

1. Split the time horizon into $T = \{1,2,\dots,t\}$ intervals each has 30 minutes of span.
2. Assume $C_{K\times T}$ be a matrix of container flow generated by K={1,2,..,k} truck generating centroids.

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1t} \\ \vdots & \ddots & \vdots \\ c_{k1} & \cdots & c_{kt} \end{bmatrix}_{K \times T} \tag{5.28}$$

Where c_{kt} is container volume generated by centroid *k* at time *t*. Please note that this matrix is used as an input signal in the GCN (see Equation (5.8)) for the loss hour predictions.

3. Based on the predefined scenarios, λ is a matrix of shift rates where λ_{kt} denotes the shift rate for centroid *k* at time *t*.
4. Determine the average capacity of each time interval

$$C_k^{average} = \frac{1}{|T|} \sum_{t=1}^{T} (C_k^{\max} - C_t) \tag{5.29}$$

where C_{max} denotes the maximum number of containers departed from terminals within a time interval in historical data and C_t is the number of containers scheduled for interval *t*.

5. Split the number of containers computed in step 2 into $M = \{1,2,\dots,m\}$ groups of containers in which the number of containers in each group does not exceed the average capacity per interval of terminals. In other words, if the number of containers in an interval is greater than the $C_{average}$, we divide them into two groups of containers with equal cardinality, each of which has less than the average capacity. This way, we can assign more than or equal to one group of containers to an interval in the off-peak period.
6. Given that $Y_{K\times T}^m$ is the shift matrix where its elements y_{ij}^m is 1 if a group of container m is shifted from time slot *i* to *j* and 0 otherwise, we calculate the shifted profile for each centroid.

$$\hat{C}_{K\times T} = C_{K\times T} \otimes (O_{K\times T} - \lambda_{K\times T}) + C_{K\times T} \otimes \lambda_{K\times T} \times Y_{K\times T}^m \tag{5.30}$$

Where O is an all-ones matrix, $\hat{C}_{K\times T}$ is the shifted container flow matrix and the signs $\otimes$ and $\times$ are element-wise and matrix multiplication respectively. Substituting the calculated shifted container flow matrix in Equation (5.10), we have

$$\hat{X}^0 = conv_c^*(\hat{C}(t - d_c\Delta t)) \tag{5.31}$$

7. To find the optimum shift matrix $Y_{K\times T}^m$, we solve the following optimization problem

$$\min \sum_{n=1}^{N} \sum_{t=1}^{T} f_{nt}(\hat{X}^0 \mid G) \tag{5.32}$$

$$\min \sum_{m=1}^{M} \sum_{i=1}^{T} \sum_{j=1}^{T} d_{ij} y_{ij}^m \tag{5.33}$$

s.t.

$$\sum_{i=1}^{T}\sum_{j}^{T} Y_{ij}^{m} = 1 \qquad \forall m \in M \tag{5.34}$$

$$\hat{C}_{kt} \leq C_{k}^{\max} \qquad \forall t \in T \tag{5.35}$$

$$d_{ij} \leq D \qquad \forall i, j \in T \tag{5.36}$$

$$Y_{ij}^{m} \in \{0,1\} \tag{5.37}$$

In this problem, we want to assign m groups of containers selected from peak period i to j interval. This formulation lets the candidate containers take any of the time slots earlier or later than that is scheduled for them. The schedule of the containers can move back and forth in time while trying to stay as close as possible to the scheduled departure time. We want to minimize two objectives in this problem which minimizing one increases the value of the other one. The first objective in the cost function sums over all N locations (nodes) in the network to calculate the total losses predicted in the traffic system for all j intervals after the shift is applied. The second objective, on the other hand, is the cost of shift which keeps the shifts as close as possible to their initial schedules. The more we deviate from the peak period, i.e. the value of the second objective becomes larger, the less the predicted loss hours on the road network will be. This is while we want to keep the deviation small as the large deviation from the planned departure time is not applicable in practice. Therefore, this problem is multi-objective optimization in essence. The d_{ij} is the distance from the center of interval j to the center of the interval i in the peak period, that group m initially belonged to. Equation (5.35) ensures that all the groups of containers are assigned to one interval. Equation (5.36) indicates that the number of containers generated at centroid k at time t i.e $\hat{C}_{kt}$ does not exceed the capacity of the interval t after applying the shift. Finally, Equation (5.37) is the time shift constraint that controls the maximum shift (D=2h) that can happen for all groups of containers. We used the NSGA-II algorithm to solve this multi-objective assignment problem.

5.5 Results

In this section, we describe the results of our data-driven decision support model for the departure time shift. This model aims to predict 5 minutes ahead of traffic dynamics (i.e., speed, volume, vehicle-loss hours) on the given network concerning changes in container demand in the port of Rotterdam. The model considers the surrounding passenger traffic as well as trucks.

5.5.1 Model performance

We divided the collected data into three subsets of training (70 %), validation (15%), and test (15%) data. We use the validation set to test the model performance during training. Training will be stopped if the model fails in a certain number of successive iterations to improve the prediction accuracy for the validation set; this prevents the overfitting of the model. Please note that all the results reported in this paper are based on the test data which is dropped out of the model in the training phase so that the results also show the high generalization power of the model to predict unseen data with high accuracy. As mentioned in the methodology section, this model works in a graph structure in which each node in the graph can learn and predict observed variables related to a desired location on a transportation network. Therefore, the performance of the model varies from node to node. We report on the performance of 5 decisive

nodes which are located in the congestion part of the network. We have already seen that this model can simultaneously predict volumes, speed, and monetary social losses for a desired section on a given network. Figure 5-7 shows the performance of the model regarding the prediction of volumes of all vehicles for 5 selected locations on each of the motorways. The subfigures on the left demonstrate the linear fit between observed and predicted volumes and the figures on the right side compares predicted and observed time series.

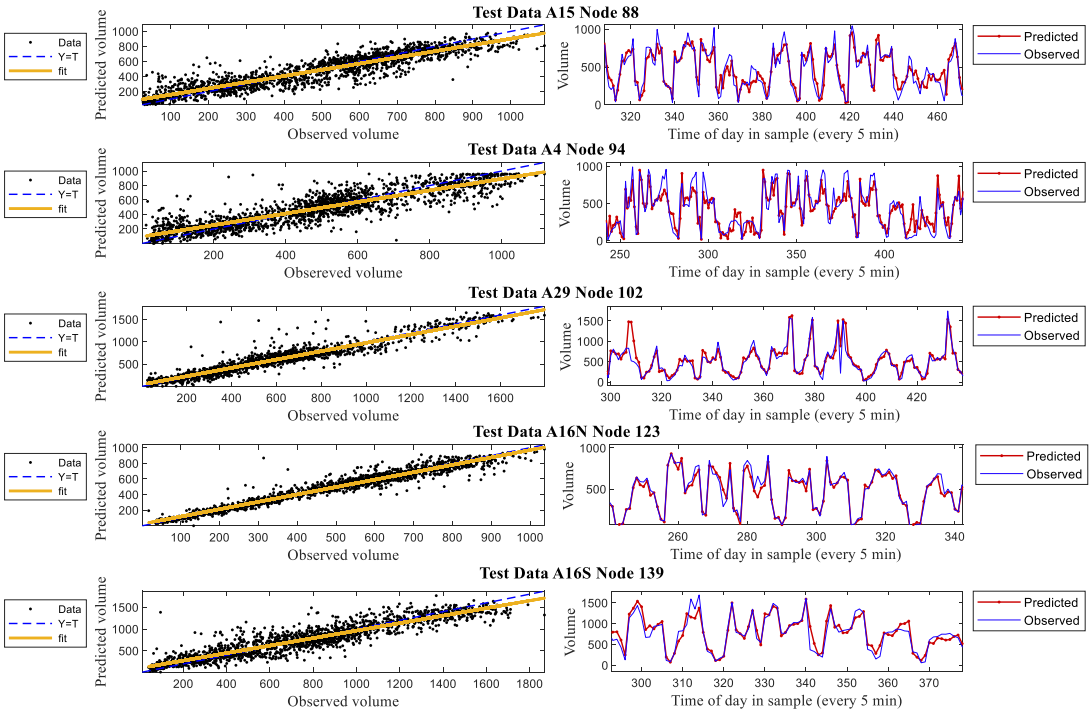


Figure 5-7: Linear fit between predicted and observed volumes

Prediction of the speed for the congested location is crucial as we may see sharp drops in speed profile during peak periods. As we can see from Figure 5-8, relatively larger errors happen within congested times of the day where the speeds are low. The largest errors however belong to the nonrecurring traffic condition on the network. We can see predictions with a relatively large error (i.e. in nonrecurring congestion) are more likely to happen on A15 and A16S motorways.

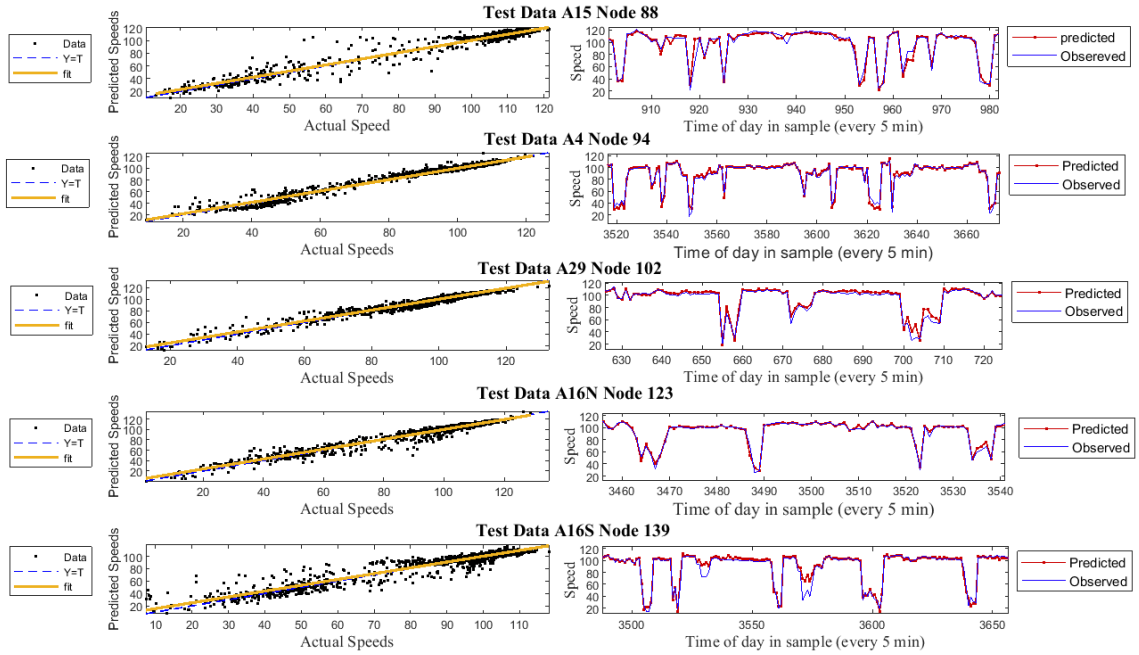


Figure 5-8: Linear fit between observed and predicted speed

This model takes into account the interaction of speed and flow from neighboring nodes and therefore can capture shockwaves moving upstream. To illustrate the performance of the model in capturing these kinds of traffic dynamics, we compare the space-time speed profiles for the observed and predicted values as well as their speed-flow fundamental diagrams in Figure 5-9. Also, visual inspection at this level of detail confirms a very satisfactory result.

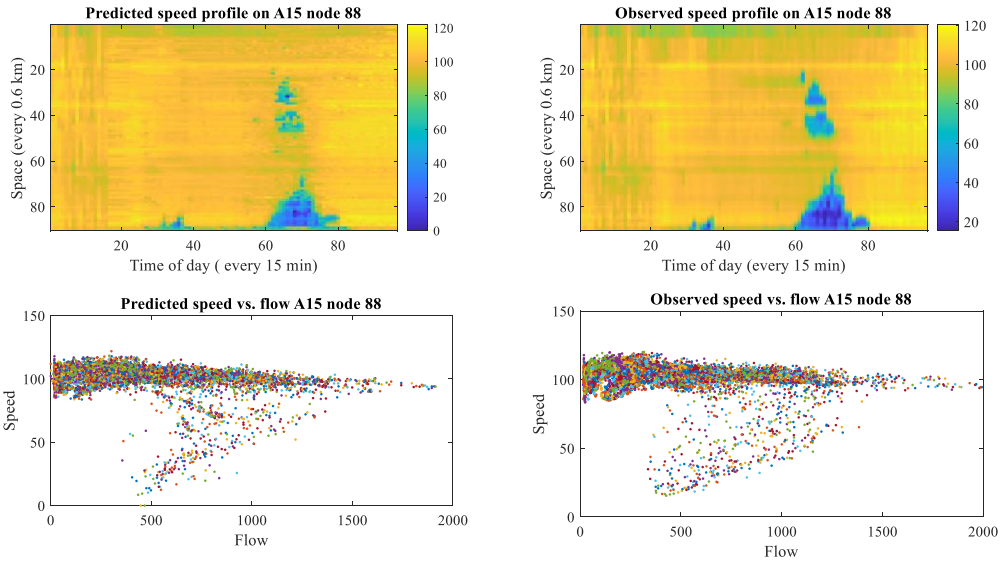


Figure 5-9: Observed versus predicted space-time speed profile and speed-flow fundamental diagram

Besides volumes and speeds, our model can predict the monetary value of loss-hours for each node. Figure 5-10 depicts the linear fit as well as time series for observed and predicted losses.

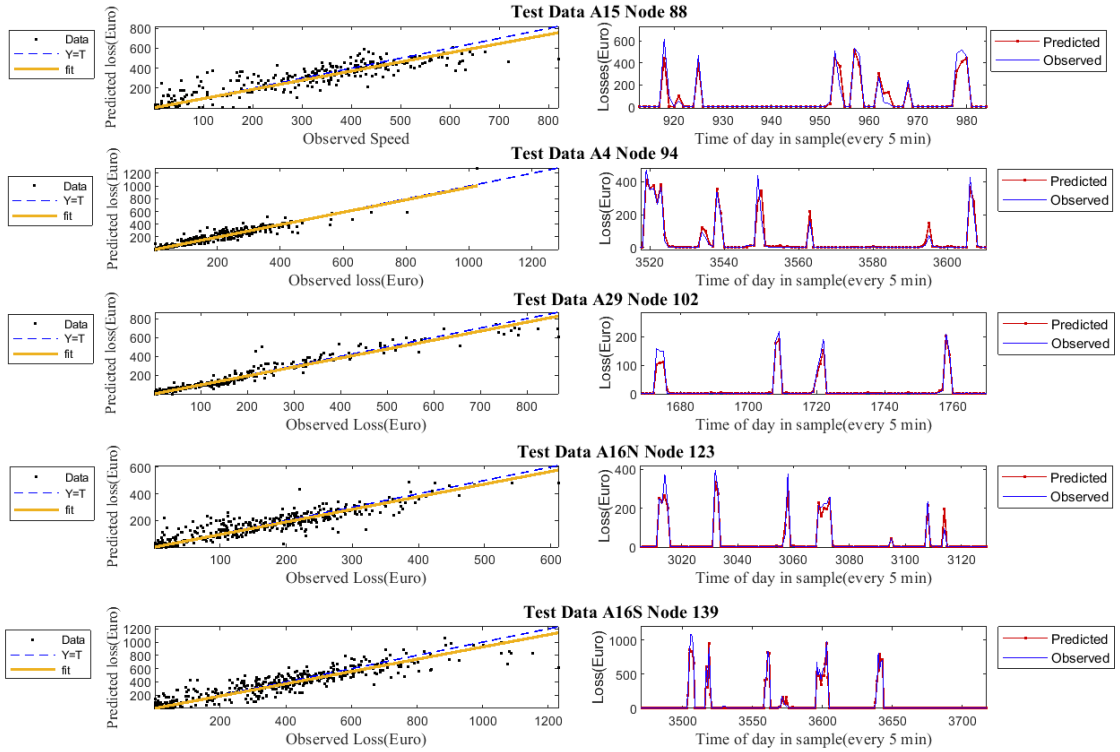


Figure 5-10: Correlation coefficient for loss prediction

Losses in the system are often a result of lower speed and higher volume where we could see predictions are with larger errors. As a result, the errors are also relatively larger when the losses peak.

Overall, the model provides satisfactory predictions for speed, volume, and losses. The network is generalizing well on predicting unseen test data. Table 5-2 illustrates the facts about the performance of the model under several indicators introduced in section 4.1.3. Table 5-2 shows the high accuracy of the model (i.e. in the worst case, $R^2 = 0.91$ for the test data on A16N) in terms of the correlation coefficient between observed and predicted speed. The model can predict volumes of all vehicles with $R^2 = 0.91$ for the A15 motorway, $R^2 = 0.89$ for A4, $R^2 = 0.95$ for A29, $R^2 = 0.91$ for A16N, and $R^2 = 0.95$ for A16S. The prediction power of the model for losses also ranges from $R^2 = 0.95$ for A16N to $R^2 = 0.98$ for A29.

Table 5-2: Evaluation results of the trained model on test data

link	Node	Target	AE	MSE	RMSE	MAE	MAPE	PAPE	R^2	Observed vs. Predicted linear fit
A15	88	volume	6.9	10296.67	101.47	75.47	2.8%	95%	0.91	$y \approx 0.83x + 77$
		speed	-0.19	25.29	5.02	2.90	0.4%	98%	0.97	$y \approx 0.96x + 3.7$
		losses	0.16	1209.81	34.78	10.90	3.2%	99%	0.96	$y \approx 0.92x + 3.7$
A4	94	volume	-0.2	15719.67	125.37	93.60	4.4%	99%	0.89	$y \approx 0.8x + 94.25$
		speed	0.03	7.98	2.82	1.80	0.2%	96%	0.97	$y \approx 0.96x + 4.33$
		losses	0.2	191.24	13.82	2.00	3.0%	98%	0.96	$y \approx 0.88x + 1.1$
A29	102	volume	-4.1	11565.83	107.54	71.16	2.0%	99%	0.95	$y \approx 0.93x + 41.67$
		speed	0	5.69	2.39	1.57	0.2%	97%	0.96	$y \approx 0.94x + 5.83$
		losses	0.34	110.79	10.53	2.15	4.2%	98%	0.98	$y \approx 0.95x + 1.1$
A16N	123	volume	-0.72	6837.60	82.69	57.14	1.4%	99%	0.94	$y \approx 0.92x + 44.2$
		speed	0.01	3.67	1.92	1.02	0.8%	99%	0.91	$y \approx 0.93x + 6.18$

Table 5-2 (continued)

link	Node	Target	AE	MSE	RMSE	MAE	MAPE	PAPE	R^2	Observed vs. Predicted linear fit
		losses	0.081	40.90	6.40	1.77	1.5%	99%	0.95	$y \approx 0.95x + 0.61$
		volume	5.66	14179.95	119.08	86.33	1.6%	99%	0.95	$y \approx 0.91x + 66.13$
A16S	139	speed	-0.06	14.06	3.75	2.09	0.3%	99%	0.96	$y \approx 0.94x + 6.34$
		losses	0.42	771.11	27.77	6.07	3.9%	98%	0.97	$y \approx 0.92x + 1.85$

The size of the error in percentage terms (i.e. MAPE) is relatively low in all cases. This means that if the model makes predictions with errors, the average size of the error would be between 0.2% (for speed prediction on A4) and 4.4% (for losses on A29) of the actual values. The PAPE also indicates that from 95% to 99% of the predictions have less than 5% error.

5.5.2 Sensitivity for truck volumes

To evaluate the sensitivity of the model, we gradually increase/decrease the demand from 10 to 40 percent during the peak period (i.e. between 15:00 to 18:00). The decrease in demand could mean a strike in the terminals and the increase in demand could mean growth in import/export of containers in the future. From the predicted speed profile for each link, we calculate the average percentage of travel time savings for the system. We use the filtered speed based (FSB) trajectory method proposed by Van Lint (2010) to calculate the travel time of a synthesized trajectory (i.e. heading toward each of the 5 motorways) of a vehicle departing from the port of Rotterdam in the second half of the afternoon peak period. Figure 5-11 shows that the model is sensitive to changes in demand, with the expected signs of impacts.

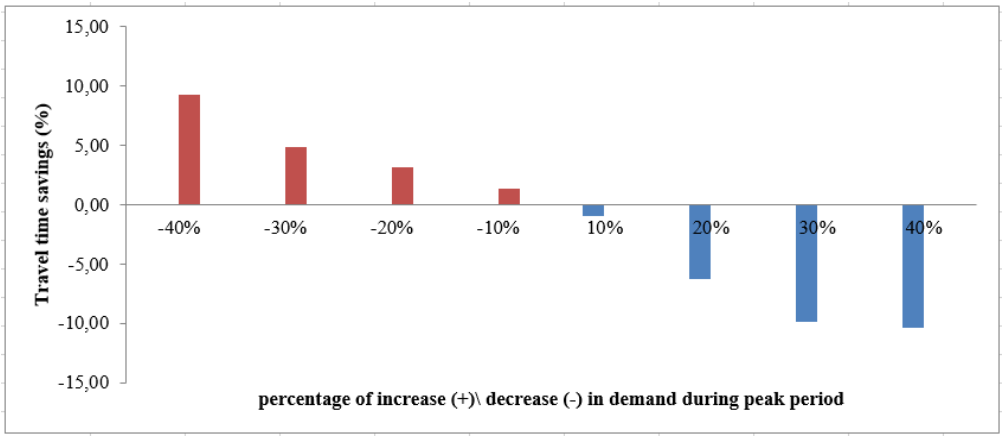


Figure 5-11: Sensitivity of the model towards increasing or decreasing demand during the peak period.

In the next section, we utilize this model to predict the impact of changes in freight transport departure times on the surrounding traffic.

5.5.3 Departure time scenarios

To assess the sensitivity of the model concerning changes in container departure time, we designed 4 scenarios for different magnitudes of shifts during the afternoon peak period. These scenarios shift 10 to 40 percent of the containers' departure time to an earlier or later time. The time shift operator ranges from 30 minutes to 3 hours across all scenarios. The gain is calculated based on the differences between the base case and the scenario losses. Figure 5-12 shows the histogram of gains for A15 in all scenarios for the selected 146 days, indicating the diverse

results of FDTS schemes, depending on the unique circumstances of every single day. This figure is just based on an unoptimized shift process where we shift trucks just to the previous time intervals regardless of what could happen if we shift them to the next intervals. The aim is to get initial insight into the application of this policy and find the possibilities for optimizing the process.

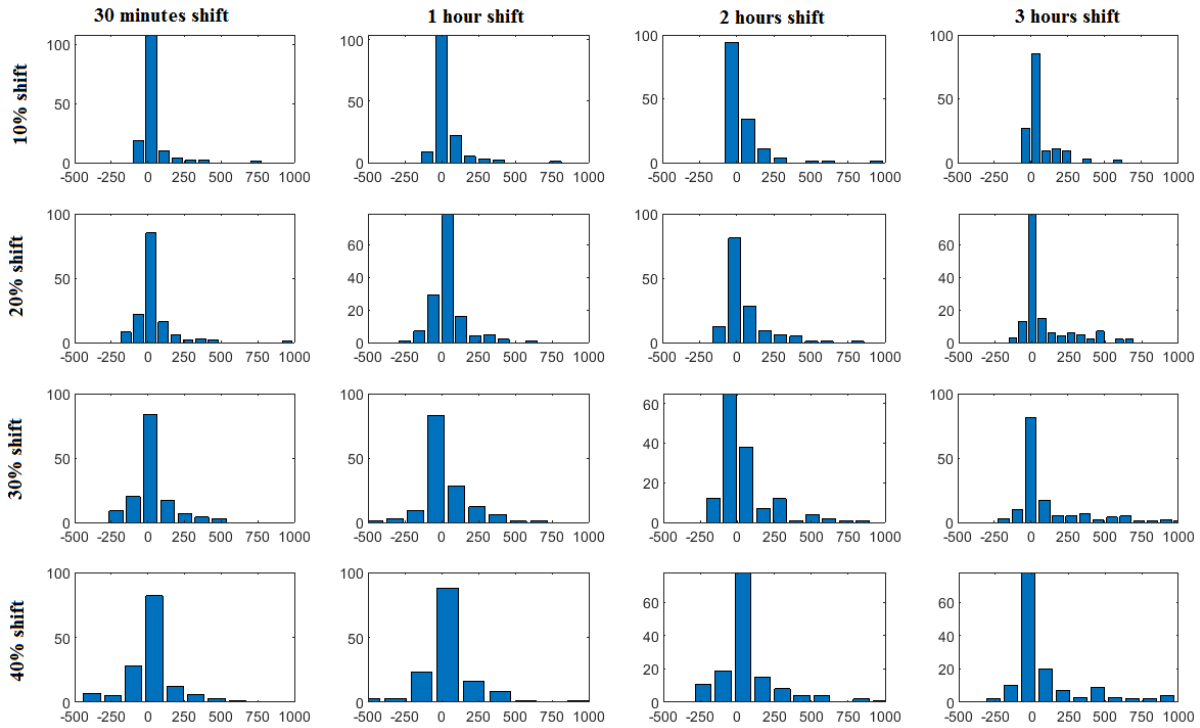


Figure 5-12: Histogram of gains for different scenarios on A15

We can see from Figure 5-12 that these time shift scenarios may lead to negative or positive gains for various days. The distribution of gains is differing from 30 minutes to 3 hours and from 10% to 40% shift. All in all, it is likely to have positive gain across all days in the experiment for all ranges of shifts when the shift rate is 10%. The probability of positive gain, however, decreases in higher shift rates and shift hours. This roughly shows the potentials for this policy which can lead to the highest gain with the lowest payoff. We should note that these distributions come from aggregated gains across all motorways in the network. While the gain for one motorway is negative for a given day, it could be positive on other motorways. For instance, Figure 5-13 points to a particular day on A15 with nonrecurring congestion. This congestion lasted almost 5 hours and spread over 30 km. Therefore, shifting departure times of trucks may, in this case, worsen the condition on this motorway. Although shifting departure times of containers may not lead to a positive gain on this motorway for this particular day, it still may lead to some gains on other motorways and thus a total positive gain on the system.

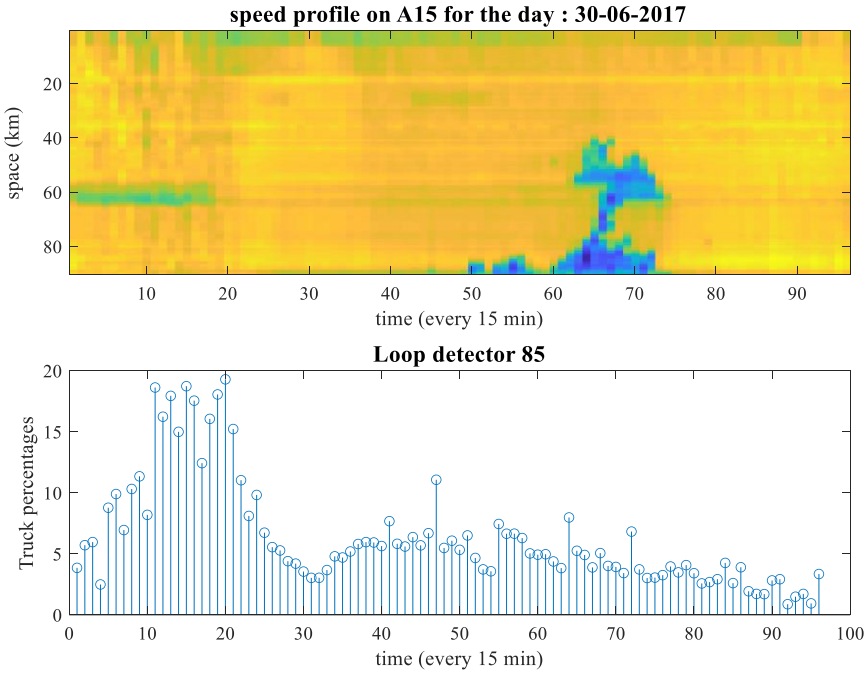


Figure 5-13: changes in truck percentages due to nonrecurring congestion on A15

5.5.4 Social Benefits of an optimized FDTS policy

In this section, we assess the possible monetary social gain from implementing the optimized FDTS policy. Our findings evaluate the potential of freight reward-based peak avoidance policy as an alternative congestion management strategy on motorways with high truck percentages. The incentives can be estimated using the proposed predictive model. Based on the number of trucks shifted and their associated social gains, we estimate the amount of monetary gain each truck contributes to the social benefits. We should note that the departure time shifts may put more pressure on terminals at one particular time slot. Slot management using the proposed predictive departure time control system could relax this burden. For the 10% participants scenario, as an example, Figure 5-14 shows the Pareto frontier and the optimum solution regarding the time shift constraint.

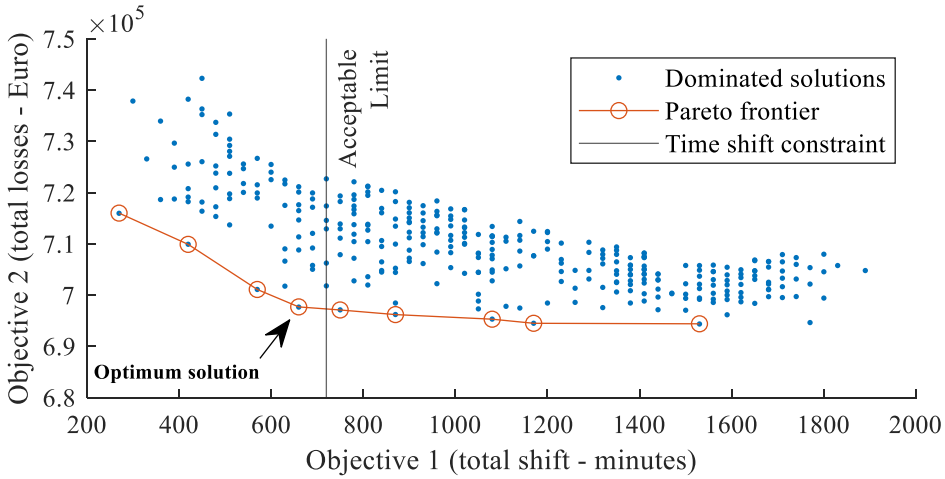


Figure 5-14: Pareto frontier of the optimized FDTS for the 10% participation rate

Table 5-3 shows how the departure times of participants are distributed across different time slots. To assess the impact of the departure time shift on traffic dynamics, first, we used the filtered speed based (FSB) trajectory method to calculate the travel time of a synthesized trajectory (for all paths) of a vehicle departed during the afternoon peak period from the port of Rotterdam. Table 5-3 also indicates the average improvement of travel time during peak period on every 5 alternative paths after applying departure time shift policy.

Table 5-3: Travel time improvement during peak hours for optimized departure time shift

Scenario	**Recommended departure time modification in**				**Links travel time improvement**				
	±30 min	**±1 h**	**±1.5 h**	**±2 h**	**A15**	**A4**	**A29**	**A16N**	**A16S**
10 %	31%	18%	24%	27%	1.8%	0.35 %	2.0 %	2.2 %	1.6 %
20 %	27%	17%	28%	28%	3.4%	0.38 %	2.2%	6.8 %	6.6 %
30 %	21%	12%	33%	34%	5.8%	1.36 %	5.8%	7.2 %	8.3 %
40 %	18%	8%	46%	28%	9.7%	2.12 %	10%	7.5 %	9.1 %

We can see that the average improvement in travel time of one synthesized vehicle varies from 0.35% to 10.00% depending on the path it chooses and the time-shift scenarios.

One of the outputs of the predictive departure time advice system is the prediction of the monetary social benefit in the system. These predictions are based on the changes in the departure time of the containers. Figure 5-15 shows the predicted social benefits based on the recommended departure time for any of the shifting scenarios.

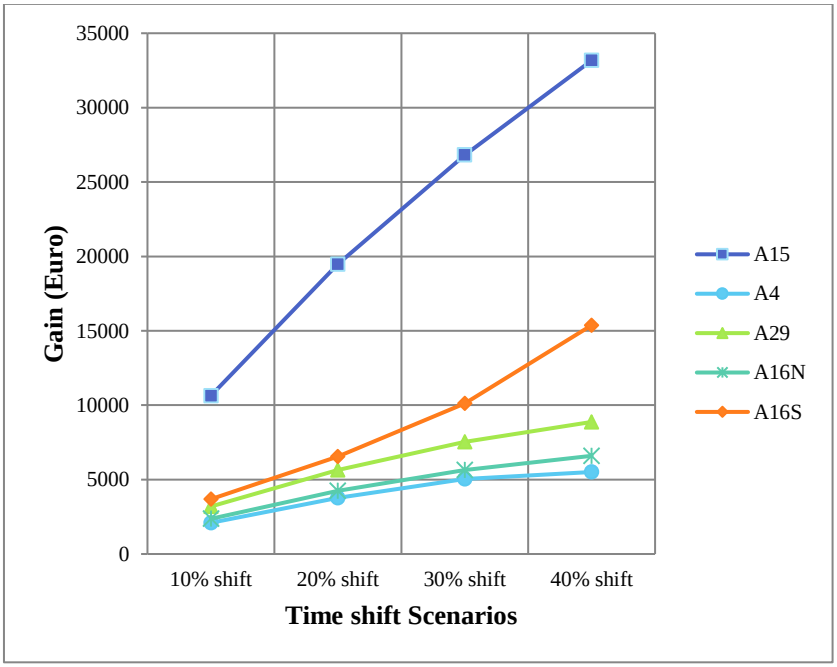


Figure 5-15: Predicted social benefits after application of DTS policy.

We calculate the social benefits for the afternoon peak and off-peak period (from 12:00 to 21:00). Table 5-4 shows that between 23144 to 69528 Euro can return to the carriers, depending on the percentage of participants, where the government can pay between 5.50 and 7.30 Euro per container to ask carriers to shift 10% to 40% of their pick up schedules to off-peak hours. Interestingly, total gains scale less than proportionally with the numbers of trucks shifted, and

thus the gain per container is largest for a 10% scenario. By adding more demand to previous time slots, we may worsen the situation for the off-peak period and thus vehicles may experience higher loss hours during those time slots.

Table 5-4: Estimated incentives based on recommended departure time modification

Share shifted	**Trucks shifted**	**Social gain per corridor (Euro)**					**Total Gain (Euro)***	**Gain/container (Euro)**
		A15	A4	A29	A16N	A16S		
10 %	3187	10644	2258	3440	2549	4253	23144 (3.2%)	7.3
20 %	6375	19485	3776	5642	4238	6537	39678 (5.5%)	6.2
30 %	9562	26816	5048	7529	5637	10112	55141 (7.6%)	5.8
40 %	12750	33170	5511	8876	6598	15373	69528 (9.6%)	5.5

* Total monetary loss for the base-case scenario is 720842 Euro.

We note that these incentives would be in addition to the travel time savings that participants experience directly by traveling during the off-peak hours. These travel time gains are calculated based on the differences between the estimated travel time during the peak and the estimated travel time during the off-peak hours. As we can see from Figure 5-16, participants approximately experience travel time savings ranging from 4 to 10 minutes depending on every 5 alternative paths after applying departure time shift policy. This can accordingly add to the FDTS benefits from 3 to 7.5 Euro per truck.

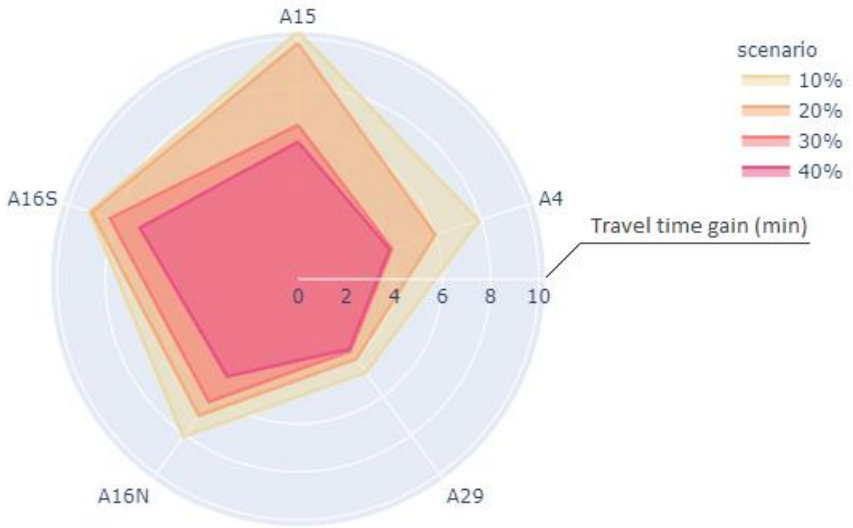


Figure 5-16: Travel time savings for shifted trucks traveling during off-peak.

We should underline the fact that our research is limited to a small local network around the port of Rotterdam because of the lack of origin and destination data. For this small network and with current traffic conditions on it, we believe that around 15 Euros of total gain per shifted truck is quite large and can suggest even a larger gain for a larger network. We should note that the relatively small network considered for this study is barely considered as highly congested.

Therefore, this gain can be considered large for the area of study. This means that the social gain could get to millions of Euros instead of thousands considering a larger and more congested network with full information of origins and destinations.

In addition, other gains can be expected from the application of FDTS. One example is the gain related to the emission reduction during the peak period caused by less congestion on the network. Truck shift policy can also shave the peak at terminals and hence less waiting times at gates.

All in all, this research emphasizes more on the method to design, implement, and evaluate FDTS policy. We believe that the results clearly show that there is potential for the application of such a policy from a traffic perspective. However, for a successful implementation of FDTS we recommend consideration of the gains and costs associated with a larger network and other aspects of the system i.e. logistics and emissions.

5.6 Conclusions and recommendations

In this section, we summarize our main findings and discuss the potential of the proposed framework regarding its use for practical FDTS policy, including the limitations of the model. Finally, we consider opportunities for further research.

Our analysis shows that changing the departure time of road freight transport can lead to an overall measurable social benefit. These gains can get back to the logistic system by offering financial incentives to the carriers if they shift their departure time to the recommended time slot before or after the peak period – in a similar way as the Dutch Spitsmijden project. The results show that the gains on this network are significant. In addition, if one takes into account the potential additional benefits from a reduction in CO_2 emissions or improved travel time reliability, total gains increase and may make shifting more attractive. Although there is no empirical evidence about FDTS prior to this study, the numbers in terms of total vehicle loss hours are quite in line with a recent report on the topic of the Dutch association of carriers (TLN, 2019). They report the economic cost of vehicle loss hours for trucks per road section. The road sections considered in our study are not in the top 50 list where the cost of vehicle loss hours for trucks is the highest. This indicates that the benefits may be higher if a larger network is considered and/or congestion will be larger in the future.

We applied the model for one truck generation centroid in the port of Rotterdam, while a network-wide application would involve many more freight generation nodes. Adding more centroids to the model may lead to higher social benefits and consequently a more attractive compensation scheme for carriers. Although our model takes into account two user groups, only a part of the freights user group comes from a demand (seaport) zone. For the rest of the traffic, we only used the information coming from the loop detectors, assuming that this traffic is equally representative of demand. Moreover, companies that implement FTDS may need to change certain operation aspects, which may incur extra costs for them. These costs are not included in this framework and, therefore, further empirical work including carriers' departure time choice would be needed to design a successful policy.

In sum, this research provides a novel data-driven traffic modelling framework using a graph-based modular convolutional neural network and an optimization model for departure time shifts. We utilized the complete framework of models to assess the effect of the truck's

departure time shifts on the state of the traffic on motorways. From the modelling perspective, our main conclusions are :

- The proposed data-driven decision support multi-class traffic model enables us to have an accurate prediction of short-term traffic dynamics (i.e. speed and flow) based on the variation of the container demand in freight hubs.
- In a multi-task learning scheme, features learned to predict flow and speed are useful and should be transferred to predict vehicle loss hours if needed.
- Using traffic flow knowledge to design the graph-based structure of neural networks makes AI traffic models not only more accurate but also more interpretable.

In terms of implications for policy and management, the relevant findings are:
- Optimized peak avoidance schemes for freight transport can lead to significantly reduced congestion on surrounding motorways and thus important social benefits. These benefits can be used to compensate the affected freight carriers or to incentivize autonomous departure time shifts.
- The identified gains of shifts in our study probably underestimate the real gains, as we have only included one freight generating hub for the network and we haven't accounted for reliability and emission benefits.
- The modelling framework can be utilized to understand the consequences of changing the departure time of trucks at the seaport terminal on the traffic system. Also, it allows optimizing the shift schemes, such that they are adapted best to local congestion patterns. The network-wide predictions can be useful for traffic management agencies to better manage congestion on roads around major logistics hubs, in a real-time context.

The above also leads to a number of research opportunities.
- Firstly, one can explore the performance of (long-short term memory) LSTM networks to adaptively capture temporal dependencies not only from the previous time steps in the current day but also from similar days in previous weeks or months.
- Secondly, integrating this model more strongly with traffic flow theory can turn the model from a correlation machine to a causality model, which adds to the model the advantages of physical interpretation.
- Thirdly, this research provides a good starting point for further work on reward-based peak avoidance policies. Arguably the most urgent need is to understand the internal costs that logistics firms or carriers may meet in case of changes in their departure time schedules.

Chapter 6

A decision support system for time slot management

In the previous chapter, we established the importance of freight demand management on both the traffic and the logistics system, by controlling departure time of the trucks from a logistics hub. Controlling truck departure time, however, requires an intelligent decision support system which can control and manage truck arrival times at terminal gates. This chapter introduces an integrated model that can be used to understand, predict, and control logistics and traffic interactions in the port-hinterland ecosystem. This approach is context-aware and makes use of big historical data to predict system states and apply control policies accordingly, on truck inflow and outflow. The control policies ensure multiple stakeholders' satisfaction including those of trucking companies, terminal operators, and road traffic agencies. The proposed method consists of five integrated modules orchestrated to systematically steer truckers toward choosing those time slots that are expected to result in lower gate waiting times and more cost-effective schedules. The simulation is supported by real-world data and shows that significant gains can be obtained in the system.

This chapter is based on the following papers:

1. Nadi, A., Snelder, M., van Lint, J.W.C., Tavasszy, L. (2022). A Data-driven and multi-actor decision support system for time slot management at container terminals: A case study for the Port of Rotterdam - Submitted to a journal.
2. Nadi, A., Nugteren, A., Snelder, M., van Lint, J.W.C., Rezaei, J. (2022). Truck Arrival Shift: An Advisory-Based Time Slot Management System to Mitigate Waiting Time at Container Terminal gates. *Transportation Research Record*, 1-14. DOI: https://doi.org/10.1177/03611981221090940

6.1 Introduction

The problem of high waiting times for trucks at seaport terminals is receiving increasing attention from practitioners and researchers. The growing volume of international trade has put ports and their ecosystem under pressure. Long queues of idling trucks at terminal gates waiting to pick up or deliver a container lead to congestion further upstream, and induce emissions, costs, and delays (van Asperen et al., 2013, Sharif et al., 2011). Therefore, effective traffic management policies at terminal gates and the surrounding areas are becoming an imperative task for most large container ports. The problem of congestion, high waiting time, and therefore non-optimal turnaround time for trucks at the terminal is often due to a lack of port-hinterland alignment (Notteboom, 2009). Establishing such alignment concerns various stakeholders, and is highly related to the connectivity between port and hinterland (Notteboom, 2009, Wan et al., 2018). In general, the port-hinterland connection can be viewed from two perspectives. The first is the physical port-to-hinterland connectivity that can be improved through the expansion of physical infrastructure. Extending physical capacity takes considerable time and hence requires more long-term strategies.

The second perspective is digital connectivity, where multiple stakeholders can communicate and exchange information for better cooperation and coordination. Digital connectivity facilitates short-term as well as medium-term policies to control the demand patterns. Various studies (Sharif et al., 2011, Wibowo and Fransoo, 2021, Chen and Yang, 2010) found that there is potential for digital-connectivity-based solutions to control truck traffic demand patterns. This form of connectivity is often relatively cheap and fast to implement. Despite its advantages, there are also some barriers against improving digital connectivity. For example, exchange of data and information has always been problematic due to privacy issues and fear of losing competitive advantages to other stakeholders. Recently, large ports around the world are developing safe and reliable data-sharing i.e. port community system (PCS) platforms to ease communication and facilitate digital connectivity. Even in the case of available safe data sharing platforms like PCS, the port community has not, in many cases, utilized these data properly due to the cumbersome process needed to transform big raw data into valuable information. These difficulties have led to relatively limited research towards digital connectivity as compared to physical connectivity. In this research, we contribute to the literature of enhancing digital connectivity by using shared data in port community systems and explore shorter-term solutions to solve day-to-day truck traffic issues at the terminals.

To reduce congestion at terminal gates, terminals have to balance the arrival time of the demand inflow with the available terminal processing capacity. This can happen through accurate time slot management at terminals gates. There are roughly three means of time slot management controlling demand inflows in which digital information plays a role. The first one is to provide real-time traffic information to facilitate more self-organized (user) optimal scheduling behavior of truck drivers and companies (Sharif et al., 2011). The challenge is, however, that the situations at terminals' gates may change rapidly due to the volatility of the demand. Therefore, providing real-time information may in some cases be counter-effective, even leading to trust deterioration in the system. The second approach is an incentive-based or charging-based scheme to spread demand across the day by providing monetary incentives to nudge scheduling behavior towards more (system) optimal decisions. Although charging-based policies like peak hour charges (Chen et al., 2011) can be, in many cases, effective for traffic mitigation, they may raise social objections. Incentive-based approaches also require sufficient funding sources for successful application.

The third approach towards more optimal scheduling is truck appointment systems (TAS) (Wibowo and Fransoo, 2021, Chen and Yang, 2010). A TAS typically uses a reservation system that allocates trucks to different time slots across the day based on the terminals' capacity to improve terminal efficiency. In the past decade, several results have been published around designing TAS, increasing its importance for future port development (Huynh et al., 2016, Abdelmagid et al., 2021). However, its design intricacies justify a deeper analysis into several aspects of the port-hinterland system (see Figure 6-1). This system includes multiple stakeholders with conflicting interests in terms of costs and benefits. For example, transport companies aim to decrease the truck waiting times and the container rehandling time. However, for terminal operation, it is necessary to balance the workload of the yard cranes (Im et al., 2021) and serve more customers. Also, the arrival of a large vessel may force the terminal operator to assign more cranes to the seaside and hence reduce the service rate in the hinterland side.

In order to portray the direct impacts of TAS on different stakeholders, the system can be divided spatially into the port and hinterland areas of concern and functionally into the logistics and traffic subsystems. Their relations are summarized in Figure 6-1.

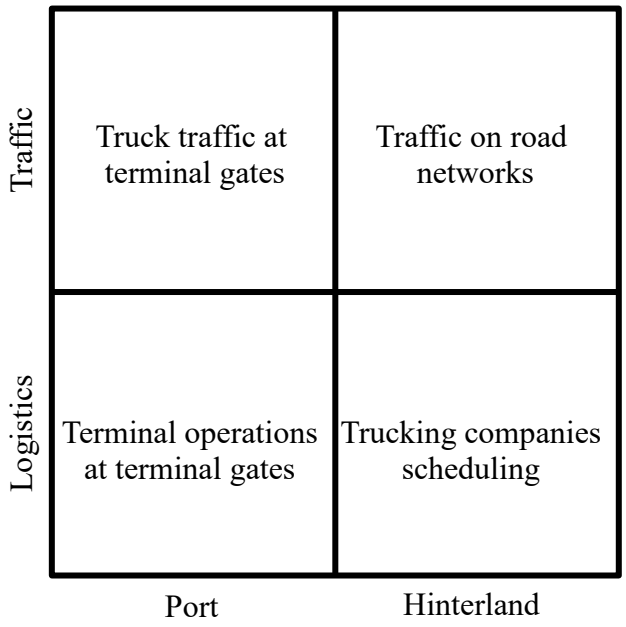


Figure 6-1 Connections between port-hinterland and logistics-traffic ecosystems

On the port side, the application of a TAS allows terminal operators to improve their operational efficiency. Because they can assign more cranes to the seaside operations due to reduced truck traffic at terminal gates (Chen and Yang, 2010, Zhang et al., 2013, Phan and Kim, 2015). On the hinterland side, it affects the operations of trucking companies as they may have to change their container pickup and delivery times. Additionally, changes in trucks inflow and outflow could be influenced by traffic on road networks. In some cases, the peak in truck arrivals happens in off-peak periods on the road network. This could happen because truckers would like to avoid facing congestion on road networks. Although a TAS can reduce congestion at the port by shifting a portion of trucks to off-peak time slots in the port, it may add more trucks to the peak periods in the hinterland side. This not only increases the costs of truckers as they have to drive in traffic jams but also adds to the cost of other road users by increasing the percentage of trucks on road networks during congested periods. However, much

of the research up to now has not treated these important traffic effects in the design of TAS. Therefore, designing a more effective TAS requires an integrated logistics and traffic modelling framework that coordinates between multiple stakeholders in port and hinterland. Previous studies predominantly ignore the hinterland side while designing time slot management systems, i.e. have neglected the roadside conditions from the users' perspective. Here, we argue that multiple stakeholders in the hinterland including, trucking companies and roadside traffic agencies can also benefit or lose from the application of TAS. In this paper, we address this knowledge gap by introducing a new decision-support approach for TAS that takes all these aspects of the system into account.

The main scientific contribution of this research is the design of a comprehensive modelling framework for a TAS that integrates the different perspectives discussed above, using appropriate techniques for each problem:

1. deep neural networks that predict truck demand at port and traffic states on road networks;
2. discrete event simulation of the truck handling process at terminals' gates;
3. behavioral modelling of trucking companies' preferences for container pick-up times;
4. mathematical optimization modelling of the time slot scheduling problem.

The model provides users of the system with the following advanced capabilities;

1. it allows priority to trucks to approach terminals' gates, based on their market preferences;
2. it provides accurate predictions of operation time at the planning horizon allowing truckers and terminals to make better decisions;
3. it suggests appropriate time slots to drivers;
4. it assures accurate coordination between multiple stakeholders in the port and the hinterland, and gives a comprehensive assessment of the potential gains/costs for the application of the proposed TAS.

This chapter is structured as follows: First, we review the literature for TAS. Next, we explain the modelling framework and methodology. Afterward, we present the results, and finally, we conclude the paper by discussing the findings.

6.2 Literature review

The truck appointment system (TAS) is a solution to control truck arrivals to the container terminal and improve terminal efficiency. In this system, trucks are appointed to specific time slots to load and unload their cargo considering the constraints of terminals and trucking companies. A TAS system has two main components i.e. the quota and the appointment mechanism. The quota mechanism relates to the optimum number of time slots during working hours, the duration of time slots, and the maximum number of trucks are allowed per time slot (Lange et al., 2020). The appointment mechanism, on the other hand, relates to the process with which a time window is assigned to an import/export container to be picked up/delivered from/to a marine terminal. This could evenly distribute truck arrivals across the day and hence increase the efficiency of terminal operations and reduce truck waiting times at terminal gates. There is good evidence for the emergence and effectiveness of truck appointment systems to reduce congestion at seaport terminals (Giuliano and O'Brien, 2007). A recent comprehensive review of the research is given in (Abdelmagid et al., 2021). The earliest discussion on truck appointment systems emerged around early 1970 in Canada (Huynh, 2009) and soon after got

considerable attention from both practitioners and researchers. In this section, we summarize the advances on TAS research, discuss them from a methodological and a design perspective.

From a methodological perspective, researchers use mathematical programming, queueing theory, simulation, or a combination of them (Abdelmagid et al., 2021, Guan and Liu, 2009a). The mathematical programming method computes the best match between truck arrival patterns and service availability according to particular objective functions (e.g. waiting time costs, terminal handling, emissions) and constraints. Various types of mathematical programming such as binary (M. Abdelmagid et al., 2021) and mixed-integer programming (linear and nonlinear (Xu et al., 2021)) have been utilized to solve time slot scheduling problems in truck appointment systems. Queueing theory is very close to the physics of the system and can provide decision-makers a good approximation of truck queue length and truck waiting times even before the real-world application of TAS (Zhang et al., 2013). Finally, discrete event simulation mimics the real-world dynamics of terminal operations in response to the application of TAS.

Building on the advantages of these methods, a large and growing body of literature has used a combination of these methods to develop a TAS. For example, M. Abdelmagid et al. (2021) proposed a binary integer programming formulation to develop a TAS that maximizes the resource utilization of the terminal while minimizing the density of trucks inside the terminal. Their approach gives higher priority to the quayside operations as compared to the truck gate operations. Guan and Liu (2009b) proposed an optimization model that minimizes various gate system costs. The cost includes labour cost, terminal operation cost, and truck waiting cost. They used queuing modelling validated with field observations to approximate these costs. Li et al. (2020) proposed a bi-objective optimization formulation to balance the trade-off between trucks and terminal operation. In their formulation, they minimize the total amount of cranes assigned to the hinterland yard blocks and the waiting time at blocks for trucks. Similarly, Mar-Ortiz et al. (2020) designed a decision support system that balances capacity management (supply) and demand management (truck arrivals). Their research aims at determining the optimum quota considering crane productivity indicator and level of service for trucks. Besides operational costs, some authors have considered emission costs while optimizing slot assignments. Do et al. (2016) developed a simulation-based genetic algorithm approach aiming at minimizing total emissions produced from trucks and cranes at import yards. They utilized discrete event simulation to estimate total truck waiting times and the total moving distance of cranes, and then minimized the associated emissions from both truck and cranes. Fan et al. (2019) also considered reducing carbon emissions while determining the optimum number of truck arrivals in each appointment period. Zhao and Goodchild (2010) evaluated the impact of arrival information on terminal rehandling and later in 2013, they proposed a hybrid queue modelling and simulation technologies to improve terminal efficiency in terms of container retrieval operation, crane productivity, and truck turn-time (Zhao and Goodchild, 2013). One year later Zehendner and Feillet (2014) involved more stakeholders from the port side in the TAS modelling. Their model not only improves the quality of service for trucks at terminal gates but also considers the crane operation for trains and vessels in a multimodal terminal.

All the above studies, although different in terms of their methodology and problem formulation, are similar in terms of the scope of the TAS design, in that they all focus on the operations of stakeholders in the port area and ignore those in the hinterland. The TAS belongs to the terminals, and trucking companies only can reserve a time slot if the slot is not fully occupied. All these methods determine the maximum number of trucks that can reserve a time

slot or identify the optimum number of time slots during working hours. In recent years, a few but growing number of studies have started expanding the design of TAS towards a multi-stakeholder system. In such a design, TAS supports stakeholders to work together and optimize their objectives accordingly. (Phan and Kim, 2015) introduce a decentralized decision-making model to manage the negotiation between terminal operators and the dispatcher of trucking companies. They propose two sub-models. The first sub-model considers the scheduling of trucks based on the time window constraints posed by the terminal operator and the second sub-model approximates the waiting time based on the number of requests. This is an iterative process, which dynamically identifies the best truck schedules according to the current terminal operation. In this process, trucking companies apply for a time slot, the terminal operator approximates the waiting time for the requested time windows and sends this information to the carriers, carriers then reschedule their time slot based on the new constraint and propose a new time slot to the terminal operator. Yi et al. (2019) used a similar approach but added new constraints to the problem and proposed an efficient algorithm to solve this problem in a reasonable computation time. Torkjazi et al. (2018) expanded this collaborative method by considering tour scheduling of carriers accordingly. Here, the time slots imposed by the terminal operation influence the tour of the carrier as they are also constrained by other time windows imposed by customers. Their method helped to improve the truck tour scheduling by 11.5% compared to the model without rescheduling. The idea of a decentralized collaborative TAS has made this system more realistic and efficient. However, the drawback of is that it requires a robust communications channel and a relatively long planning horizon, as trucking companies have to dynamically change their plan. This long replanning time may keep the carrier on pause as the situations at terminals can rapidly change.

In summary, most existing studies assume that trucks can follow the optimum design of the appointment system at no cost. Such truck appointment systems usually enforces truckers to choose another time slot even though this shift in their arrival may have a domino effect on their operation schedules in the hinterland. In other words, these models consider the cost that truckers would have if they had to wait in a queue but not the cost that they would have if they had to shift their arrival time due to the lack of available spots in their preferred timeslot. Only a few recent works have considered this in designing a collaborative dynamic truck appointment system. However, this requires several backs-and-forth communication, as well as negotiation between carriers and terminal operator, which significantly increases planning time. Therefore, it is essential for terminal operators to consider truckers' scheduling preferences. This helps the collaborative TAS to minimize the communication need between stakeholders. In addition, previous research so far has neglected the role of road traffic in the design of the TAS. Traffic plays a key role in the scheduling of trucks and varies according to the variation of truck demands.

In this research, we introduce a new centralized truck appointment system, integrated with a traffic model. This central TAS communicates with all stakeholders i.e. terminal operators, traffic agencies and trucking companies, and collects required information. Then, it gives accurate advice to each stakeholder according to their conditions or preferences. The focus of this research is not only on providing a normative solution for companies, as is usually the case in the literature of TAS, but also to provide a quantitative description of the system for managers and policymakers, helping them to understand the port-hinterland ecosystem better. One can think of this method as a combined logistics and traffic system where traffic at port and hinterland can be understood together and controlled simultaneously. With our method, the benefits and cost of the system can be realized and distributed in both logistics and traffic

systems. This system can also deal with heterogeneity in trucking operations of multiple industries. Previous TAS studies have assumed a first-reserve-first-serve policy, and make no distinction between different preferences of shippers, e.g. between perishable agricultural merchandise with a morning delivery window and fashion products destined to warehouses in the afternoon. With the power of data-driven models deployed, our proposed method gives priority to some industries for specific time slots according to their inferred scheduling preferences.

To conclude the review, it appears that site managers at container ports require more disaggregate behavioural insights to have a better grip on demand. Data-driven methods allow exploring trucker behaviour at a disaggregated level but have, to our knowledge, never been applied to TAS. In this research, we aim to help fill this gap by introducing an optimal control policy derived from an analysis of a large historical sample of carriers and terminal gate operations. In addition, we use extensive traffic data to assure that control policies on logistics sites do not deteriorate traffic conditions on road networks. This also has never been considered in the literature.

6.3 Methodology for design of the truck appointment system

In this section, we propose a methodology for the design of a data-driven centralized time slot management system that takes multiple stakeholders into account. The building blocks of the proposed system are depicted in Figure 6-2.

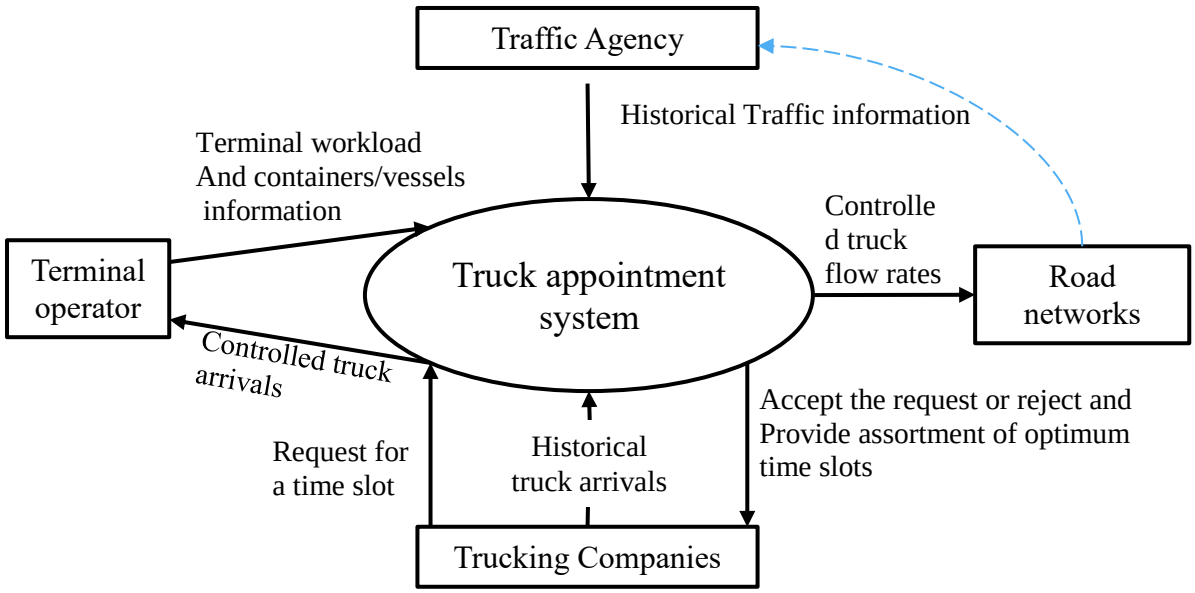


Figure 6-2: Centralized Multi-stakeholder truck appointment system

The process begins with carriers (trucking companies) who send a request for a time slot to the truck appointment system once they receive an order from shippers. This request may be part of a planned tour, which may include several other pickups and deliveries in the hinterland. For the TAS however, this request is the only observable part of the tour. The TAS should learn from this partially observed tour planning information to infer the scheduling of carriers for different times of the day. On the other hand, the system looks into the terminal activities. Terminals can provide more information about their activities through the port community

system. Based on this information, the TAS system can predict the workload of the terminal at the requested time slot, and predict the truck waiting time cost at the terminal gate. Simultaneously, the system receives information from traffic agencies about the historical traffic conditions on road networks. This can provide an accurate prediction of delays on the surrounding road networks for trucking companies at the requested time slot. Based on this information, the system can give suggestions to each actor to help them keep their operations efficient, while controlling the congestion at terminal gates. For example, it could advise the terminal operator to balance the workload of cranes between the yard side and the hinterland side based on the number of trucks approaching the gates (see Figure 6-3). On the other hand, if the terminal is overloaded with containers discharged from a big vessel, it may advise the terminal operator to assign more cranes to the yard operation and, at the same time, propose new time slots to carriers based on their scheduling preferences and road traffic conditions. From the traffic agency perspective, the system also can notify a traffic manager beforehand about a surge of trucks heading into the hinterland at a particular time of the day, so that they can apply appropriate traffic control measures. Our system of interest is illustrated in Figure 6-3.

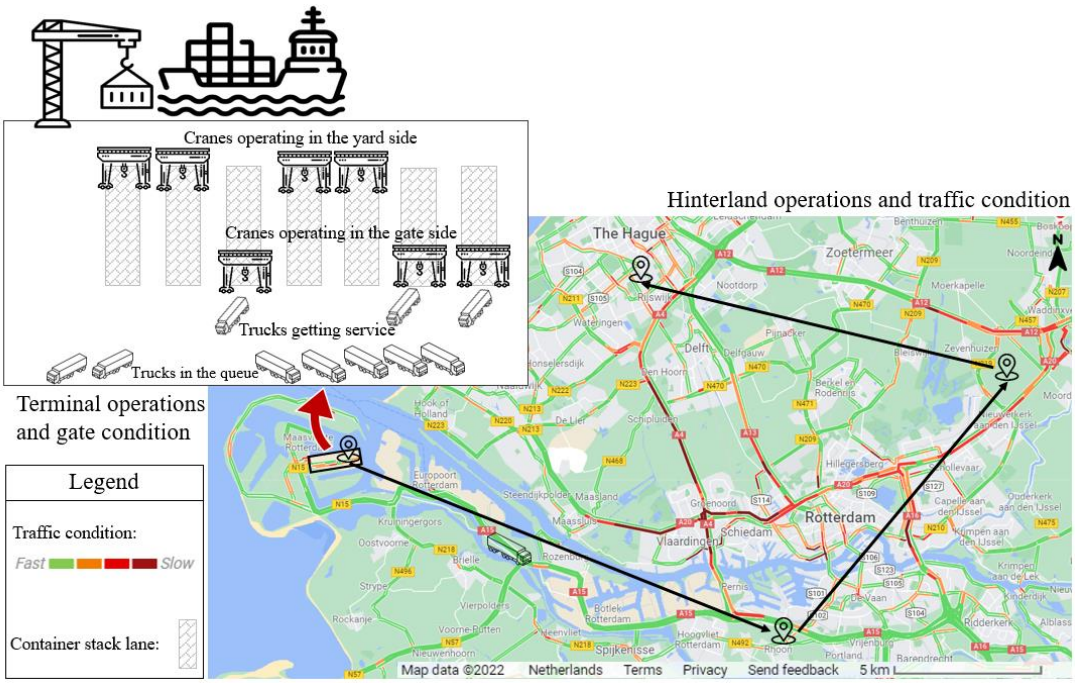


Figure 6-3: Logistics operations and traffic conditions in the study area

We assume that the decision time in this system begins at the end of a pre-defined planning time window (e.g. 1 hour) which can be 3 days to 3 hours before the actual operation time. The width of the planning time window is flexible and can be set by a system administrator. During this time, the system collects all the information from multiple stakeholders, and at the end of this time windows reveals the expected system states and associated prescriptions.

The objective is to minimize the total daily cost of the system which consists of three categories of costs:

1. Terminal gate operation cost C_t^g which is a function of C_t^S, the service cost of using a crane on the gate side during time slot *t*, and *S*, the number of cranes giving service to the trucks on the gate side.

$$C_t^g = C_t^S \times S_t \tag{6.1}$$

2. Carriers' costs (C_t^C) which includes two components:
    a. C_t^w: Total cost of waiting in a queue at the terminal gate during time slot t
    b. C_t^p: Planning or scheduling cost of visiting terminal during time slot t due to the hinterland operations
3. Societal costs of traffic systems C_t^{tr} due to the traffic intensity on the main roads from and to the port

In this section, we propose a port-hinterland modelling framework considering all these costs. This framework requires 5 modules. The first module is a demand model which predicts (generates) truck demands along the day. The second is the container terminal gate module which includes queueing models to represent terminals gate operations. This module uses truck inflows (generated by the first module) as an input and generates truck outflows, terminal operation costs C_t^g and waiting costs C_t^w. The third module is a data-driven traffic module that can estimate the societal cost of traffic system C_t^{tr} learning from day-to-day traffic patterns. This module uses the truck outflow generated by the second module to predict the monetary value of vehicle loss hours on the road network. The fourth module is the data-driven truck scheduling model which can predict the associated hinterland cost of trucks for picking up containers at the port in different time windows along the day i.e. C_t^p. The last module is an optimization model which minimizes the total system costs by assigning appropriate time slots to trucks. Figure 6-4 illustrates the complete modelling procedure and connection between these modules.

In the next subsections, we explain the method used to develop each component as well as the data used for a case study in the port of Rotterdam.

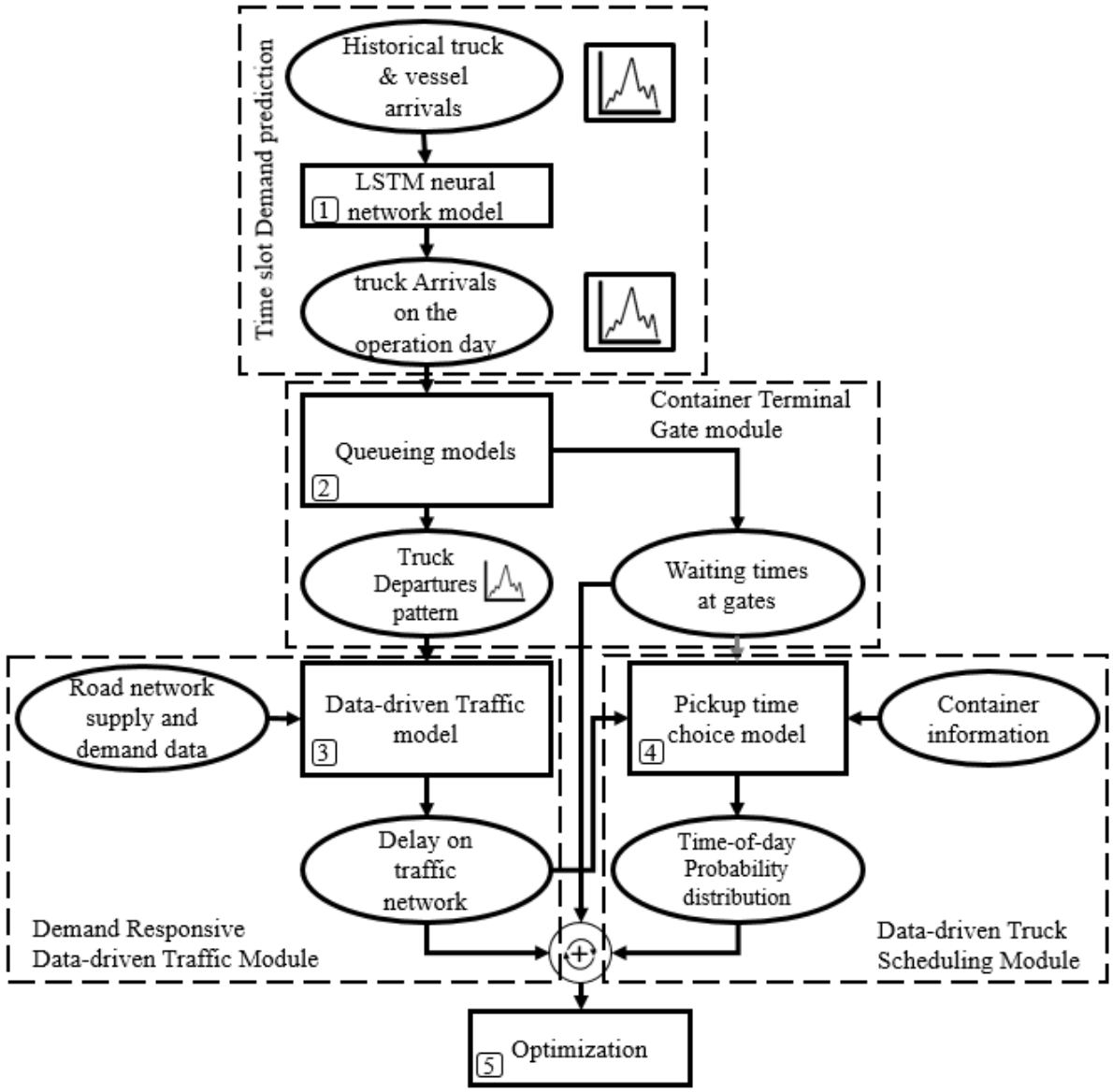


Figure 6-4: Modelling procedure for time-slot management system

6.3.1 Time slot demand prediction

We first divide the time horizon of the system into the planning and operation days. Assume that on the planning day at time windows t_p, the system receives requests for different time slots t_o in the operation day. The operation day could be the same as the planning day or one to two days after that. In either case, we assume that t_p is at least 3 hours before t_o. To predict the waiting time for each time slot on the operation day, we are first required to predict the number of trucks that arrive at each terminal. This prediction takes place on the planning day and the only information we can use for this prediction belongs to the day of planning and the days before that. To understand what information can be a good predictor of the truck demand for a time slot, we should look at the process at terminals. In a seaport terminal, once a vessel arrives, the cranes discharge containers. Then straddle carriers take the containers to the stackyard where the yard-hinterland cranes stack the container on the lanes. A container can stay in the terminal for weeks until a truck comes to the port and pick it up. Figure 6-5 shows an example of a typical distribution of pickup latencies, based on data from the Port of Rotterdam. The average pickup latency is approximately 4488 minutes (74.8 hours). That means it takes more

than 3 days on average for containers to be picked up. We can see that the distribution is skewed and has a wide right-side tail. This buffer time depends on the container discharge time, administration time, stacking time, and carriers planning latency which may differ from container to container and firm to firm. Since not all this information is available to the system at the time of prediction, we bypass the detailed process and predict truck demands for each time slot using a time series model.

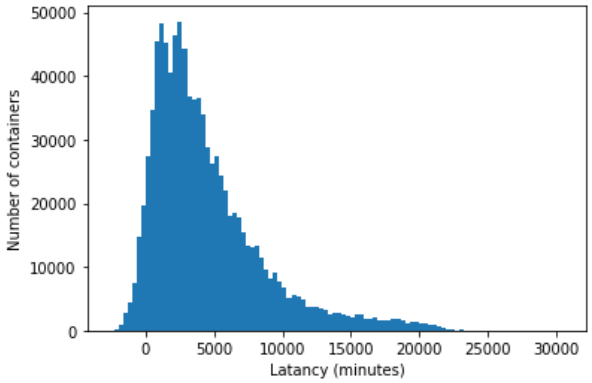


Figure 6-5: Distribution of pickup time planning latency

This requires a multivariate time series that can predict truck demand based on multiple input signals. The first input signal is the historical time series of the number of road freight containers arrived by vessels at the terminal and the second signal is the time series of trucks arrivals, which has autocorrelation and can deal with the recurrent daily pattern of truck arrivals.

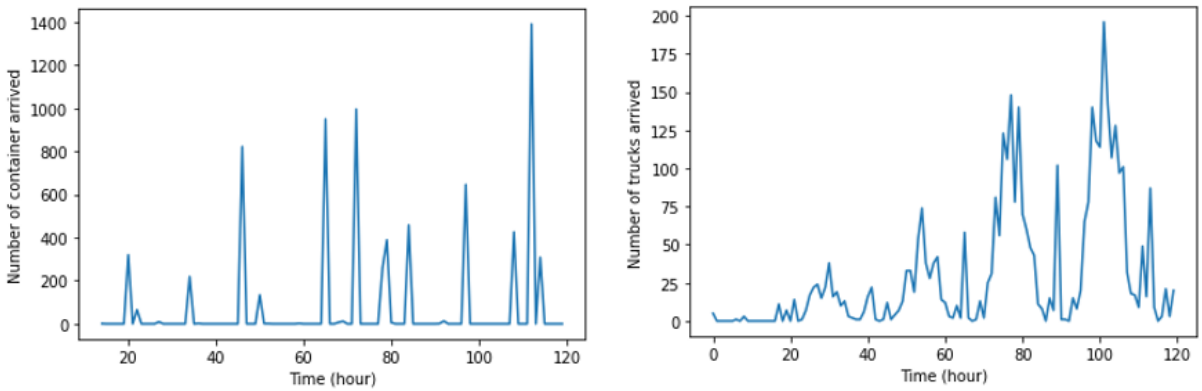


Figure 6-6: Comparison between deep-sea container arrival signal and truck arrival signal

Looking at Figure 6-6, we can see that there is a clear distinction between container and truck arrival signals. the distinction relates to the fact that the hundreds of containers arrive at the terminal all at once at a particular time of day during the call of a vessel. In other words, the deep-sea container arrival signal is zero unless a vessel arrives at the port and then the signal spikes. These containers are then distributed across different time slots in the following days as trucks arrive gradually over the day to pick them up. Since several points in the deep-sea container arrival signal contribute to multiple points in the future truck arrival signal, it is not straightforward to use this signal regularly to predict only one point in time. Instead, we

formulate this problem as a multivariate multi-step time series prediction model. Equation (6.2) shows a general representation of this problem.

$$\begin{bmatrix} y_{t_o} \\ \vdots \\ y_{t_o+k} \end{bmatrix} = f\left(\begin{pmatrix} x_{1,t_p-l} & \cdots & x_{1,t_p} \\ \vdots & \ddots & \vdots \\ x_{n,t_p-l} & \cdots & x_{n,t_p} \end{pmatrix}\right) \tag{6.2}$$

Our model aims at forecasting the next k values (e.g. k=24 hours) of truck arrivals time series in operation day t_o from two input time series (n=2) of deep-sea container arrivals and truck arrivals in the planning t_p with history lookup size of l days before the planning day. To deal with this problem we propose a sequence to sequence deep recurrent neural network (RNN) for multivariate multi-step time series prediction, where we can predict the truck arrivals at all time slots of the operation day based on the combined information sequence (i.e. both deep-sea arrivals and truck arrivals) of the past days. Sequence-to-sequence deep neural network architectures have seen successful applications in various multivariate time series prediction problems (Du et al., 2020, Du et al., 2018), but not yet in our context. Figure 6-7 depicts a graphical illustration of the proposed model.

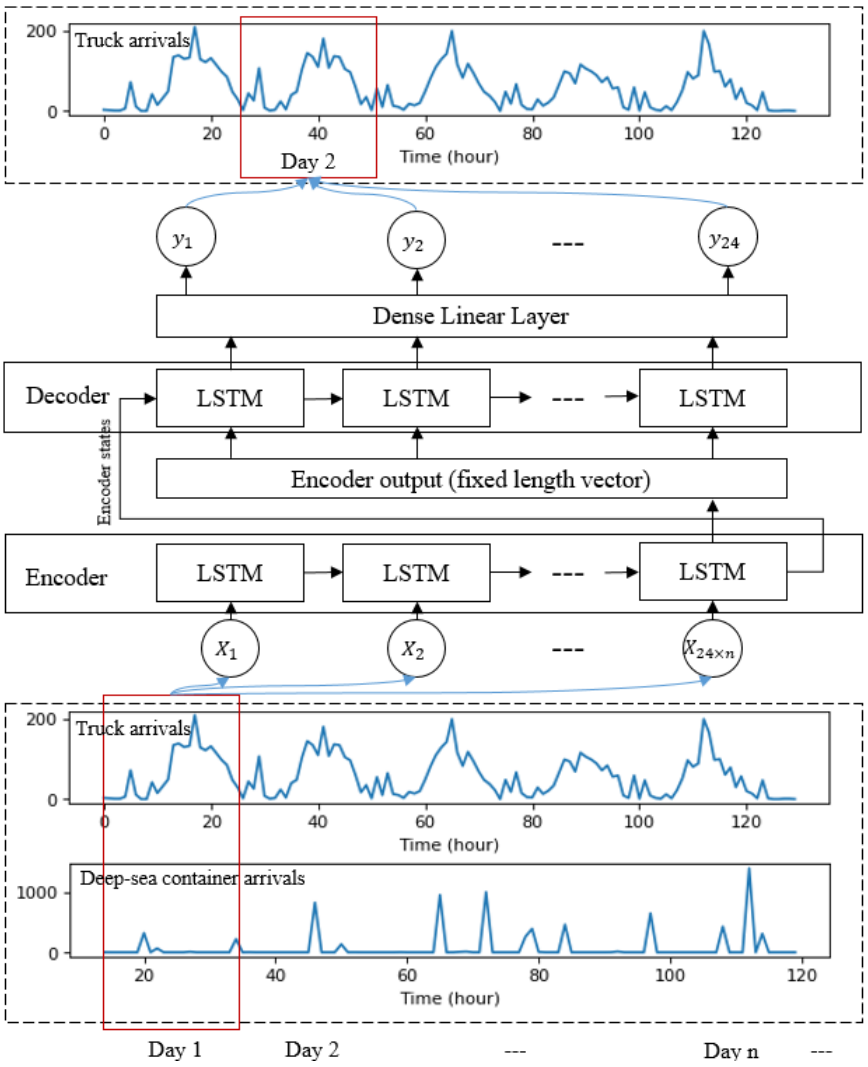


Figure 6-7: Architecture of the sequence to sequence deep RNN for truck arrival prediction

This model consists of two-three layers: The first layer is the encoder with several long-short term memory (LSTM) units which takes the historical multivariate time series samples as input to capture the temporal representation (fixed-length vector) of the past time series input. The second layer is an LSTM decoder, which generates the future time series as forecasting output. We used the fixed-length vector as the input and the encoder's final state as the initial state of the decoder layer. Finally, we used a linear dense (fully connected neural network) layer on top of the decoder layer to predict the target values for each time slot in the target sequence. LSTM is a type of recurrent neural network which was first introduced by Hochreiter and Schmidhuber (1997) and is known for its good performance in sequence to sequence prediction. LSTM was developed to deal with the vanishing gradient problem of traditional recurrent neural networks. A unit of LSTM includes 4 components, an internal memory cell c_t, an input gate i_t , a forget decision gate f_t, and the output decision gate o_t. Collaboration between these components enables the unit to learn and memorize long-term dependencies in a sequence. For example, in our case, the LSTM cell can learn when to consider or forget the impact of a spike in the deep-sea container arrival signal occurring in the past and, also, how big this impact would be. The compact representation of the equations for an LSTM with a forget gate is

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \tag{6.3}$$

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \tag{6.4}$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \tag{6.5}$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{6.6}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \tag{6.7}$$

$$h_t = o_t \circ \tanh(c_t) \tag{6.8}$$

Where x_t is the input vector to the LSTM unit, σ_g is the sigmoid activation function, W and U matrixes are the weights for input and recurrent links respectively. b vector is the bias parameters and $\tilde{c} \in (-1,1)$ is the cell input activation vector. As we can see in Equation (6.7), c_t is the summation of two terms. First is the element-wise multiplication of self-recurrent state c_{t-1} and forget gate f_t and the second term is element-wise multiplication of input gate i_t and input activation vector.

In sum, the proposed model for demand prediction of time slots in the operation day is based on an encoder-decoder end-to-end sequence deep learning architecture that can predict the most probable future time series. In the next section, we explain how the predicted demand can be used as inputs to estimate truck outflows, waiting times, and costs at terminal gates.

6.3.2 Container terminal gate module

There are two sides to a marine container terminal gate system. On one hand, we have the supply side which relates to the operations and gates and can be determined by the number of productive lanes and service rates (number of trucks being served per unit of time). On the

other hand, we have the demand side which is the process of truck arrivals. Given the physics of the system, we can formulate the terminal gate as a multi-server queueing system with a non-stationary average truck arrival rate (λ_t), the average gate service time of trucks (μ), number of active cranes at the hinterland side (S), average truck waiting time (T_t), and average number of trucks waiting (N_t) at time slot t. According to the characteristics of terminal gates, we assume that both the inter-arrival time and the service time are independent and identically distributed random variables with an exponential distribution and gates serve trucks with an integer number of cranes i.e. servers. Therefore, our method is the M/M/S queueing model. To calculate the waiting time at each time of the day we can use discrete event simulation software or the approximation method in Equation (6.9) developed by Cosmetatos (1976) and used in several similar studies (Guan and Liu, 2009b).

$$T_t = \{\frac{{a_t}^S}{\mu_t (S-1)!(S-a_t)^2}[\sum_{n=0}^{S-1}\frac{{a_t}^n}{n!} + \frac{{a_t}^S}{(S-1)!(S-a_t)}]^{-1} \tag{6.9}$$

where $a_t = \frac{\lambda_t}{\mu_t}$ denotes traffic intensity.

To set up the queue model of a terminal gate, the mean service time μ_t and the number of servers S_t have to be known. In our case, this information is missing. Therefore, we estimate these two parameters from field measurements by iterative examination of different settings for the queue simulation while minimizing the mean squared differences between the simulated and observed departure profiles. After calibration, we can use the model with estimated parameters to approximate the waiting time of trucks, the number of trucks waiting at the queue, and the departure profile of trucks heading towards the hinterland all based on the predicted truck demand explained in the previous section. Using Equation (6.9) and Little's law, we can calculate the number of trucks waiting in the queue ($N_t = \lambda_t T_t$). Given the average truck idling cost per hour C_I, Equation (6.10) calculates the total cost of waiting in a queue at the terminal gate during time slot t:

$$C_t^w = C_I \times N_t \tag{6.10}$$

As explained in section 6.3, another cost of truckers in this system is the cost of visiting the terminal at the timeslot t according to their tour planning in the hinterlands. In the next section, we explain how this cost can be inferred from observed data.

6.3.3 Data-driven truck scheduling module

For carriers, the cost of visiting container terminals is a part of the total cost of routing and scheduling tours along the day. Each tour may include several locations, one of which being the container terminal. The time at which a truck arrives at a container terminal depends on the sequence (in time and space) of the other pickup and drop-off locations in the hinterland. We could estimate the total hinterland operation cost of carriers for different terminal visit times if we had information about the daily activity of trucks in the hinterland. This information, however, is not observable to the system. The only observable characteristics are the containers and the arrival time of trucks to the terminal for each container. Carriers often solve a class of vehicle routing problems to plan their port and hinterland operation, minimizing their cost or maximizing their utility. Therefore, we can estimate the cost of visiting a time slot for each truck due to the relevant assumption that the utility of the truckers is maximized by their planned arrival time t.

$$\max \sum_{n=1}^{N} \sum_{t=1}^{T} U_{nt} T_{nt} \tag{6.11}$$

$$\sum_{t=1}^{T} T_{nt} = 1 \qquad \forall n \in N \tag{6.12}$$

$$T_{nt} = \{0,1\} \tag{6.13}$$

where U_{nt} is the utility of truck n if visiting terminal during time windows *t*, T_{nt} is a binary variable which is 1 if the truck *n* visit terminal at time *t* and 0 otherwise.

Equation (6.12) makes sure that each truck can only choose one arriving time window. We can assume that truckers use the above mathematical formulation to find the optimum visiting time T_{nt} that maximizes their utility. To our system, however, the inverse of this problem matters, where the optimum vising time T_{nt} is observable to the system and the system has to infer the utility of carriers U_{nt}. We can use utility maximization theory to calculate the probability of choosing a timeslot alternative over the others. The utility of one particular time window depends on some observed and unobserved factors that are related to the characteristics of tours and hinterland operation. We can formulate the utility function as follows:

$$U_t = V_t + \varepsilon_t \tag{6.14}$$

where V_t is the deterministic part of the utility, which includes the observable tour attributes χ_i that influence the probability of visiting terminals in a particular time window *t*.

$$V_t = \sum \beta_i \chi_i + ASC \tag{6.15}$$

where β_i are parameters of the model to be estimated from observations and capture the impact of each attribute on the utility of time windows and ASC is the alternative specific constant. The second part of the utility function contains an error term ε_t. This error term represents the unobserved factors that influence the utility of time windows alternatives.

We can estimate the parameters β_i using maximum log-likelihood estimation. In the maximum log-likelihood estimation, the model aims to estimate the parameters such that the model has the highest probability of fitting the observed data. Equation (6.11) can be transferred to Equation (6.16) which presents the maximum log-likelihood function.

$$\max L(\hat{\beta}_1, ..., \hat{\beta}_k) = \sum_{n=1}^{N} \sum_{t \in T_n} T_{tn} \ln P_n(t \in T_n \mid \chi), \tag{6.16}$$

where L indicates the log-likelihood and $P_n(t|\chi)$ represents the probability that a trucker *n* visits terminal during time windows t given attributes χ. Using the logit function to calculate this probability will lead us to equation (6.17).

$$P(t \in T_n \mid \chi) = \frac{e^{V_{tn}}}{\sum_{j \in T_n} e^{V_{jn}}}, \tag{6.17}$$

After estimating the parameters of this model, Equation (6.17) generates the probability distribution of time windows for an individual truck given its attributes. This probability can relate to the utility of this truck at different time windows. We can then translate this utility as the unitless planning cost of trucker C_t^p due to its hinterland operation (see section 6.3). In other words, if the probability of visiting terminal for a given attribute is high, it means that the cost of carriers for visiting terminal at that time windows, C_t^p, is low.

$$C_t^p = \eta(1 - P(t \in T_n \mid \chi)) \tag{6.18}$$

where η is a factor to scale this cost component compared to the others presented in section 6.3. In the next section, we explain the method to calculate the last cost component required for the design of the time slot management system.

6.3.4 Demand responsive data-driven traffic module

As explained in section 6.3, we also have to consider societal cost of traffic systems attributed to the dynamics of the port-hinterland activities. Here this cost relates to the delays per km that vehicles face under different conditions of the road network. Delays can be calculated as a deviation of travel times from free-flow speed conditions. To calculate the travel time of a particular path in a network, we require the average speed of vehicles along the path. The average speed of vehicles changes in time and space and can relate to the attributes of supply and demand on a road network. For our purposes, the model should be able to assess the impact of changes in truck inflow and outflow in the port area on the traffic system and estimate the traffic-related cost of these changes. Candidate approaches are (1) macroscopic traffic models (model-driven) that use traffic flow theory to describe the spatial and temporal dynamics of principal traffic flow variables and (2) data-driven methods that use historical data to predict traffic conditions in time and space. In general, each of these methods has advantages and disadvantages and cannot outperform the other under all conditions (Calvert et al., 2015). The main difference between them is that parameters of the data-driven approach typically have no physical interpretation, while the traffic flow approach has some parameters with physical interpretation. Additionally, data-driven models can follow the day-to-day variation of traffic states, whereas model-driven approaches typically deviate from the actual traffic states (Calvert et al., 2015). We refer to Vlahogianni et al. (2014) and Van Lint and Van Hinsbergen (2012) for a comprehensive overview of these methods. In this paper, we opt for data-driven approaches to reduce the calibration effort and to allow daily operations with short calculation times, as well as lower computational complexity. However, we integrate aspects of model-driven approaches from traffic flow theory to add to its transparency and interpretability.

A natural way of modelling traffic in a data-driven approach is through a non-linear mapping between demand and supply patterns. Demand concerns time-dependent inflows and outflows (summation of trucks and other vehicle classes) to the road network and supply concerns time-dependent road network characteristics like flow and speed. The correlation between the demand and supply dynamics happens in a very short-term time horizon (up to an hour ahead)

(Nadi et al., 2021). However, neither demand nor supply information of operation day is observable at the planning day (often one or two days before the operation day). In sections 6.3.1 and 6.3.2 we explained how we can predict the inflow and outflow of truck demand for the operation day given the available information in the planning day and history. To estimate the impact of this predicted demand on the traffic system and calculate the traffic-related costs associated with changes in the demand, we identify two required functionalities for a traffic model. First, it should be able to predict spatial-temporal traffic states for the operation day (one or two days ahead) based on available historical traffic data in the planning day. Second, it should be able to capture the evolution of truck demand and its impact on the road network in time and space. In this paper, we skipped the first component assuming that these long-term predictions are available for the system. We, therefore, use the observed spatial-temporal traffic states of the operation day. For the real application of the system, however, we recommend the use of advanced deep neural network methods proposed for spatial-temporal long-term traffic predictions which have shown high accuracy (Zang et al., 2018).

For the second component, we need to consider temporal and spatial interaction between demand nodes (trip generation zones) and various locations on road networks. The complexity of this task is mainly there because traffic by nature is spatially correlated and highly dependent on the topology of the road network. To cope with this complexity, we use a similar method as we proposed in chapter 5 (section 5.4.1). For the convenience of the readers, we briefly explain how this method works in this chapter. This method is a graph-based convolutional recurrent neural network. Graph convolution filters have recently gained special attention for short-term traffic prediction (Li et al., 2021) due to their performance in extracting spatial and temporal patterns. The choice of this neural network architecture is also based on macroscopic traffic theory where the traffic state dynamics of a link can be described in state-space by considering the upstream link, the link under consideration, and the downstream link. In this way, the evolution of demand dynamics on the consecutive links' properties can be modeled close to the physical properties and composition of a real traffic system. We represent the topology of the road network as a weighted directed graph *G(V, E, A)* where *V* is a set of nodes ($V \in \mathbb{N}$) representing loop detectors on motorways, *E* is a set of links representing road segments and $A \in \mathbb{R}^{N \times N}$ is a weighted adjacency matrix that represents node connectivity and proximity. For this study, we aim to predict the impact of truck demand generated from the port of Rotterdam on the surrounding 5 motorways up to the boundaries of our study area (see Figure 6-8). In this model, each node on the road network is a neuron or a hidden layer where the output of one layer is incorporated into the inputs of the next layer (node). Figure 6-8 shows how nodes in the graph are stacked together representing the topology of a road network.

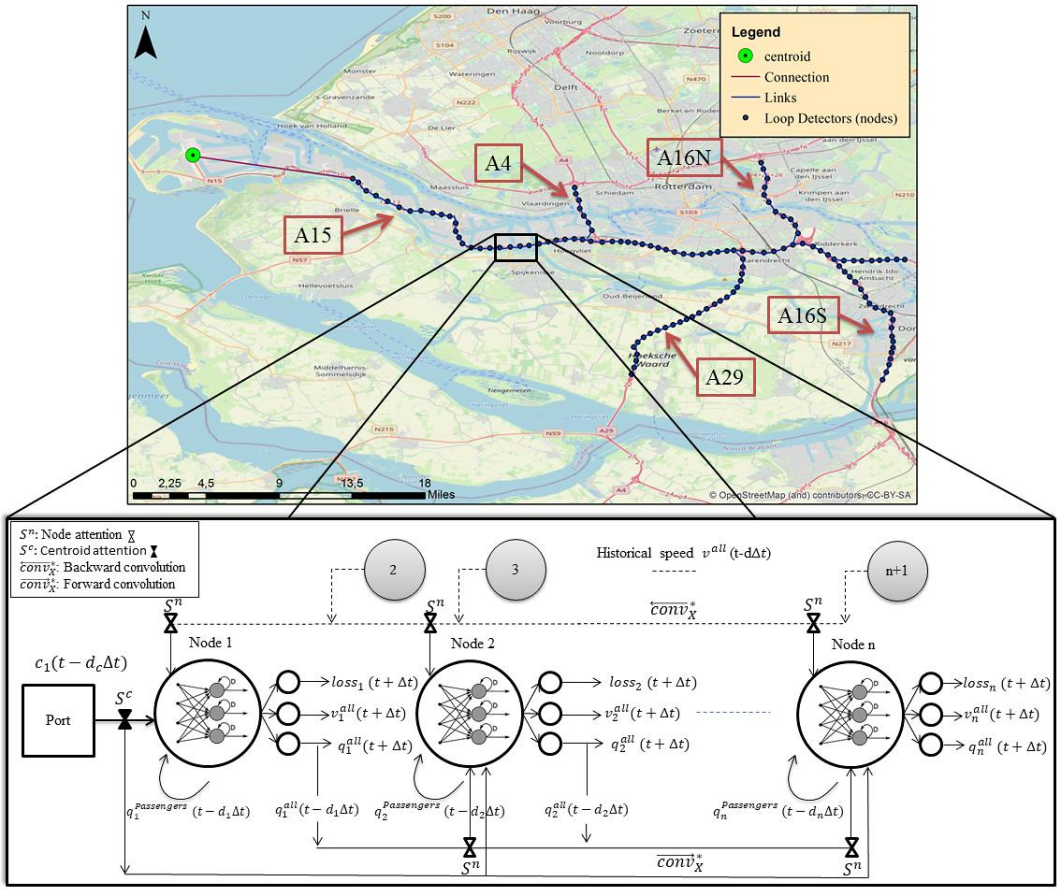


Figure 6-8: Graphical representation of the proposed graph-based convolutional neural network model

Let $X \in \mathbb{R}^{N\times M}$ be the signal input to each node of the graph G, where M is the number of features of each node (e.g. volume, speed) and N denotes the number of nodes. Also, let $C \in \mathbb{R}^n$ represent the signal input of the truck demand generation centroids added to the passenger cars signal collected from loop detectors in the vicinity of the truck generating zones and connected to $n \in N$ nodes in the graph G. Having X^t and C^t representing observed signals at time t, this model aims to learn a generalized non-linear approximation function $f(.)$ for each node in a given graph G that maps $d \in \mathbb{N}^{N\times(N+n)}$ historical input signals to the future signal $X^{t+\Delta t}$. It is important to mention that the off-ramps and on-ramps are well connected to the other nodes to ensure that the inflow and outflow to the road network match well. As Figure (6-8) shows, information from one node in a graph is passed to its neighbour nodes through the adjacency matrix. The adjacency matrix keeps the conservation law valid in the graph as the prediction of flow/speed from one node, is an input to its neighbours. To capture the spatial evolution of demand on the entire road network, we use the graph convolution (also known as message passing) technique. This elegant technique computes the spatial correlation of one link to its upstream or downstream links (Zhou et al., 2018). In this paper, we enhance the message passing rules proposed by Kipf and Welling (2016) by adding a new attention mechanism that produces a dynamic k-order weighted adjacency matrix for propagating each node's feature to its neighbours. This is done as follows:

$$conv_X^* = \bar{D}^{-0.5}(\Theta^n \otimes \tilde{A})\bar{D}^{-0.5}\bar{X}W_X \tag{6.19}$$

$$D_{ii} = \sum_j \tilde{A}_{ij} \tag{6.20}$$

$$\tilde{A} = A + I \tag{6.21}$$

$$\Theta_{ij}^{n} = \exp(-\frac{1}{2}(\frac{p_{ij}}{\sigma_i^n})^2) \tag{6.22}$$

The convolution operator $conv^*$ in Equation (6.19) calculates the aggregate (i.e. normalized weighted sum) of features for all nodes where D is the diagonal degree matrix, I is the identity matrix and $\tilde{A}$ is the adjacency matrix with added self-loops. The self-loop considers the features of the node itself along with its neighbors. W is a matrix of trainable weights and Θ^n is the attention mechanism that produces a dynamic k-order weighted adjacency matrix. Θ_{ij}^n are elements of the attention matrix Θ^n that can adaptively capture spatial correlations on the road network. With the attention mechanism introduced in equation (6.22), some nodes can benefit from the information coming from their second, third, or higher-order neighbors. Logically, these correlations fade out as nodes in the graph are getting away from each other. Therefore, we assume that these distance-related weights in the adjacency matrix are random variables with Gaussian distributions as presented in Equation (6.22), where p is the node proximity matrix. The parameter σ_i^n needs to be estimated for each node i; σ^n works as an adaptive spatial memory that helps the model remember the spatial dependency of various locations on the road network. In other words, the node i gets information from its close-by nodes when σ_i^n for the i^{th} node is small. This operator works in two directions in the graph. The forward pass is used to let the information of previous nodes (i.e. upstream) influence the prediction of the volume downstream and the backward pass sends the speed information of the next nodes (i.e. downstream) to the upstream for the speed prediction. This bi-directional aggregation helps the neural network to capture spillback phenomena, once we have congestion on a link (see Figure 6-8). Since this model is demand-responsive (meaning that the graph is connected to some demand nodes generating inflows and outflows to the graph) we introduce a similar special attention mechanism as in Equation (6.22) with spatial memory σ^c, to let the demand nodes be the k^{th} neighbor of all sensors. This mechanism is centroid attention Θ^c with which the neural network can adaptively learn correlations between demand and different nodes on the road network.

The neural network structure also includes a Softmax decision gate E_{mi} that can decide if adding inputs c_m from a demand node m improves predictions in a specific node i on the road network. This decision gate is similar to that of forgetting gates in typical LSTM networks.

$$E_{mi} = \frac{e^{c_m}}{\sum_{j=1}^{M} e^{c_j}} \qquad \forall i \in N \tag{6.23}$$

The resulted matrix $\bar{E}$ filled with elements E_{mi} is then applied to the inputs of the graph as a convolutional operator presented in Equation (6.24).

$$conv_C^* = (\bar{E}\Theta^c \otimes \bar{O})\bar{C}W_c \tag{6.24}$$

where $\bar{O}$ is a matrix of all ones, as all demand nodes are initially considered as the neighbour of all nodes in the network and W_c is a trainable matrix of parameters for convolution on centroids inputs. Since the decision gate matrix E automatically decides the normalized contribution rate of input signals to the predictions of traffic at a given node on the road network, it is like helping the network to simultaneously predict the destinations of the demands generated at each specific centroid. After the graph coevolution on the input signals, the adjusted inputs are transferred to g which is a linear activation function of a hidden layer as presented in Equation (6.25).

$$f(\bar{X}, \bar{C} \mid G) = g(conv_X^* \times \theta_1 + b_1, conv_C^* \times \theta_2 + b_2) \tag{6.25}$$

In sum, our proposed method lets the network pay relatively more attention to the most valuable information coming from different nodes and centroids and take care of the evolution of flows from demand nodes to the road network in time and space.

As far as monetary costs in the traffic system are concerned, we design this model to be able to predict the monetary value of vehicle loss hours regarding variations in speed and flow. In other words, inputs to the model are expected inflows and outflows, speeds, and demand signals at timestamp *t*, and outputs are speeds, flows, and monetary losses on the graph at time stamps t+Δt. Flows and speeds of vehicles are measured directly from the sensors for each link. Monetary losses are calculated based on vehicle loss hours which is the number of hours that all vehicles lose passing a link during time t as compared to if they would have passed the similar link under free-flow condition. Equations (6.26) to (6.28) show how this monetary loss for each node *i* at time *t* is calculated.

$$VLH_{i,t}^{Trucks} = \max\left(\frac{q_{i,t}^{Trucks} l}{v_{i,t}^{Trucks}} - \frac{q_{i,t}^{Trucks} l}{FFS^{Trucks}}, 0\right) \tag{6.26}$$

$$VLH_{i,t}^{Passengers} = \max\left(\frac{q_{i,t}^{Passengers} l}{v_{i,t}^{all}} - \frac{q_{i,t}^{Passengers} l}{FFS^{Passengers}}, 0\right) \tag{6.27}$$

$$Loss_{i,t} = VoT^{Passengers} \cdot VLH_{i,t}^{Passengers} + VoT^{Trucks} \cdot VLH_{i,t}^{Trucks} \tag{6.28}$$

where $q_{i,t}^{passenegers}$ and $q_{i,t}^{Trucks}$ are the passenger and truck flow passing section *i* at time *t* derived directly from field measurements. Class-specific vehicle-loss-hours is the deviation between flow-weighted travel time of each vehicle class in current and free flow conditions (see Equations (6.26) and (6.27)). Finally, the *Loss* denotes the space-time monetary loss matrix of all vehicles which its elements are the summation of vehicle loss hours for passengers ($VLH^{Passengers}$) multiplied by the value of time for passengers (10 euros per hour) and vehicle loss hours for trucks (VLH^{Truck}) multiplied by the value of time for trucks (45 euros per hour).

The third cost component for time slot management introduced in section 6.3 was C_t^{tr} which is the societal cost of the traffic system due to the changes in the arrival and departure of trucks

at terminals. Given the traffic mode introduced above, we can now compute this cost for each time slot as follows:

$$C_t^{tr} = \sum_{i=1}^{N} Loss_{i,t} \tag{6.29}$$

We use the mean square error (MSE), the most commonly used error function, and the Levenberg-Marquardt (LM) algorithm to estimate the model weights and parameters. For more detail about parameter estimation of this method, we refere readers to chapter 5, section 5.4.1.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{6.30}$$

Here N is the number of observations, y_i is the i^{th} observed value and $\hat{y}_i$ is its corresponding predicted value. We also use the root mean square error (RMSE) and mean absolute percentage error (MAPE), to evaluate the performance of the model.

6.3.5 Mathematical formulation and simulation-based optimization

Given all the models introduced in the above sections, we can compute all the costs in the system associated with changes in the demand and terminal service rates. In this section, we use simulation-based optimization to first simulate all these models and calculate costs and then formulate an optimization problem to manage time slots minimizing these costs. Figure (6-9) is a graphical representation of how this system is simulated and optimized.

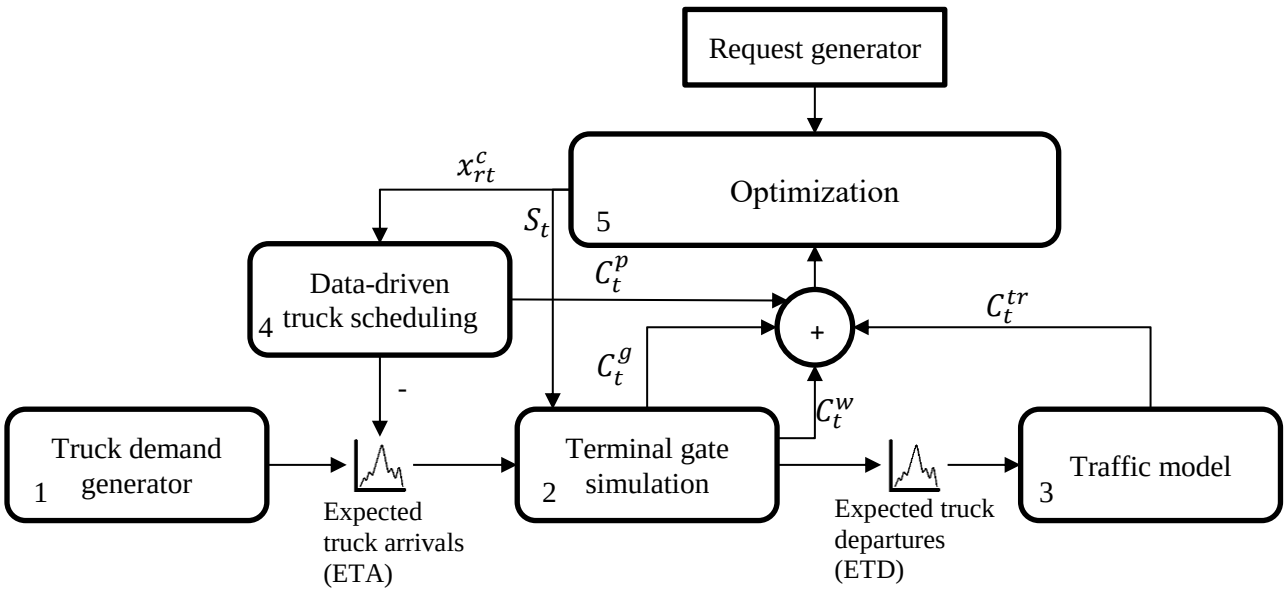


Figure 6-9: Simulation-based optimization framework for the time slot management system

We expand Equation 6.1 to formulate the optimization process in this simulation experiment. For readability and content, we present the notation of variables used in this methodology as follows:

x_{rt} A binary decision variable equal to 1 if request r is assigned to time window $t' \in T'$

S	An integer decision variable denoting the number of active cranes in the hinterland side
ETA_t	Expected number of truck arrivals at time slot t in the operation day
$ETA_t^{t_p}$	Adjusted expected number of truck arrivals at time slot *t* in the operation day at time t_p in the planning day
ETD_t	Expected number of truck departures at time t based on the output of the queueing models
$R_t^{t_p}$	A list of requests at t_p in the planning day for time slot *t* in the operation day
λ_t	Average truck arrival rate at time t in the operation day
μ_t	Average service time for each lane at terminal gates
C_t^p	Planning cost of visiting terminal during time slot t due to the hinterland operations
C_t^w	Truck waiting cost per hour at terminal gates
C_t^S	The service cost of using a crane at time slot *t*
C_t^{tr}	The cost traffic system due to the intensities on road networks at time slot *t*

This problem is a multi-objective optimization. For a given truck arrival rate (constant demand), an increasing number of active lanes (increasing supply) will increase the cost of terminal gate operation while decreasing the waiting time and vice versa. Therefore, there is a trade-off between demand and supply costs. A similar relations also hold for traffic systems where assigning trucks to times slots with less traffic at gates may coincide with peak hours on road networks.

$$\min\ [z_1, z_2, z_3, z_4] \tag{6.31}$$

$$z_1 = \sum_{t \in T'} \sum_{r \in R_t} C_t^p X_{rt} \tag{6.32}$$

$$z_2 = \sum_{t} C_t^w = \sum_{t \in T'} f(ETA_t^{t_o}, S_t, \lambda_t, \mu_t) \tag{6.33}$$

$$z_3 = \sum_{t \in T'} C_t^S \times S_t \tag{6.34}$$

$$z_4 = \sum_{t} C_t^{tr} \sum_{t \in T'} \sum_{n \in N} loss_{nt}(ETD_t) \tag{6.35}$$

S.t.

$$ETA_t^{t_o} = ETA_t^{t_o-1} - (\| R_t^{t_o-1} \| - \sum_{r \in R_t^{t_o-1}} X_{rt}) \tag{6.36}$$

$$\sum_{t\in T'} ETD_t = \sum_{t\in T'} ETA_t \tag{6.37}$$

$$\sum_{t_o\in T_o} R_t^{t_o} \le ETA_t \tag{6.38}$$

$$S_t < S_{\max} \;\; \& \;\; S_t > 0 \;\; \& \;\; S_t \in \mathbb{N} \tag{6.39}$$

$$X_{rt} \in \{0,1\}, e_{ij} \in \{0,1\} \tag{6.40}$$

In the above formulation, we have five objective functions. Equations (6.32) to (6.36) belong to the planning, waiting, service, and road network traffic costs respectively. Equation (6.37), updates the expected number of truck arrivals recursively based on the assigned trucks in the previous planning time window. This implies that the optimization problem should be solved for each planning time window t_o. Equation (6.38) concerns the conservation of flow. Equation (6.39) guarantees that the cumulative number of requests for each time slot does not exceed the expected number of truck arrivals. Finally, Equations (6.40) and (6.41) show the search space and boundaries of the decision variables.
The problem presented above is a multi-objective non-linear mixed-integer optimization problem, in which the simulation optimization procedure presented in Figure 6-9 is used to obtain an optimal assignment of requests for each time slot, as well as an optimal value for the *S* (number servers) to minimize the total daily cost of the port and hinterland system. This procedure is as follows:

1- First of all, a comparison between the container data and traffic data is used to convert the number of containers to the number of trucks for each time slot t in the planning day. For each planning, time window t_p, a Monte Carlo simulation method is used to draw and generate requests for each time slot in the operation day. These requests are proportionally selected from the observed distribution of truck pick-up time reported in the PCS data. These requests have all the information about the requested time slot, container and commodity type, and the deep-sea arrival time of the container to the terminal.
2- We use the requested time slots as the initial solution in the first iteration of the optimization process.
3- We calibrate and validate the queueing models for each terminal based on field measurements.
4- Next is estimating data-driven truck scheduling model and simulating this model to calculate the cost (using equations (6.18) and (6.33)) of each time slot for each request generated in step 1 (attributes of requests are inputs to the truck scheduling model).
5- Build and simulate the truck generator module to generate the expected number of truck arrivals for the operation day and adjust it based on the assignment solutions (using equation (6.37)).
6- Use the adjusted truck arrivals and number of active container lanes as an input to the terminal models to calculate the waiting costs and expected number of truck departures.
7- Train the traffic model as explained in section 6.3.4 and use the sum of time series of the expected number of truck departures of all terminals with time series of passenger cars. Use the result as an input to the traffic model to predict the loss hours of all vehicles on the network.

8- Finally, use a multi-objective optimization algorithm to find the Pareto frontier of solutions making a trade-off between all these costs.

We use the optimization toolbox of MATLAB R2019b as a solver to solve this problem. This toolbox uses a controlled, elitist genetic algorithm, a variant of NSGA-II (Deb et al., 2002), to find the Pareto front of solutions.

6.4 Experimental setup

As explained in the previous section and illustrated in Figure 6-9, our method consists of five main modules. The first module is to predict time series of truck demand for each terminal; the second is the marine terminal gate model to use these predictions to estimate waiting times and trucks' departure time series; the third the data-driven traffic model to predict traffic states (flow, speed, and vehicle loss hours) across the road network. We use the port of Rotterdam as our case study. The fourth is the data-driven truck scheduling module which estimates the cost of scheduling a truck to pick up a container at a particular time window within a day; and finally, In this section, we demonstrate and validate these components before being used in the simulation-based optimization procedure. We begin by introducing the available datasets and then explain and validate the parameters of all models.

6.4.1 Logistics and traffic data

Two main categories of data i.e. logistics and traffic data are used in this research for estimation of models and their validations.

- The logistics data is twofold. One is the PCS container data (provided by the Portbase) for five terminals operating in the Port of Rotterdam. These data include the deep-sea arrival time of vessels, container discharge time, container size/type, commodity type, and the time of loading containers on trucks. This data is used to estimate the data-driven truck scheduling model as well as to generate time slot requests for the simulation. The truck demand generator also utilizes the historical deep-sea arrival data stored in this data set in combination with historical truck inflows to predict truck demand. We also use data from two trucking companies reporting on their idling time at terminal gates. This data is used to validate the marine terminal gate queuing model.

- Traffic data are collected from loop detectors located on the road network (provided by the national data warehouse NDW) and also in the vicinity of terminals gates (provided by the Port of Rotterdam). Road network data includes a stream of flow and speed data in 1-minute resolution and is used to estimate the traffic model and the gate data includes inflow and out flow to the terminal per hour which is used to calibrate marine terminals' gate models.

6.4.2 Demand model estimation

As explained in section 6.3.1, we use a sequence to sequence deep LSTM network with encoder-decoder architecture for multi-step forward prediction of truck demand trough out the operation day. We used the data collected for the whole year of 2017 from PCS for this model to predict hourly demand for the operation day. We used 10 months (January to October) of data for training and the months November and December for testing and validating the model respectively. We used the Adam algorithm, 40 units for each hidden layer, the learning rate of 0.005, the dropout rate is 0.2, the validation frequency is 20, and the gradient threshold is set

to 1 to prevent gradient explosion. The mean square error (MSE) is used as the loss function. Figure 10 shows the prediction results and performance of the model on the tests set for all four terminals. Terminals are anonymously labeled A to D due to the privacy regulations of the data provider.

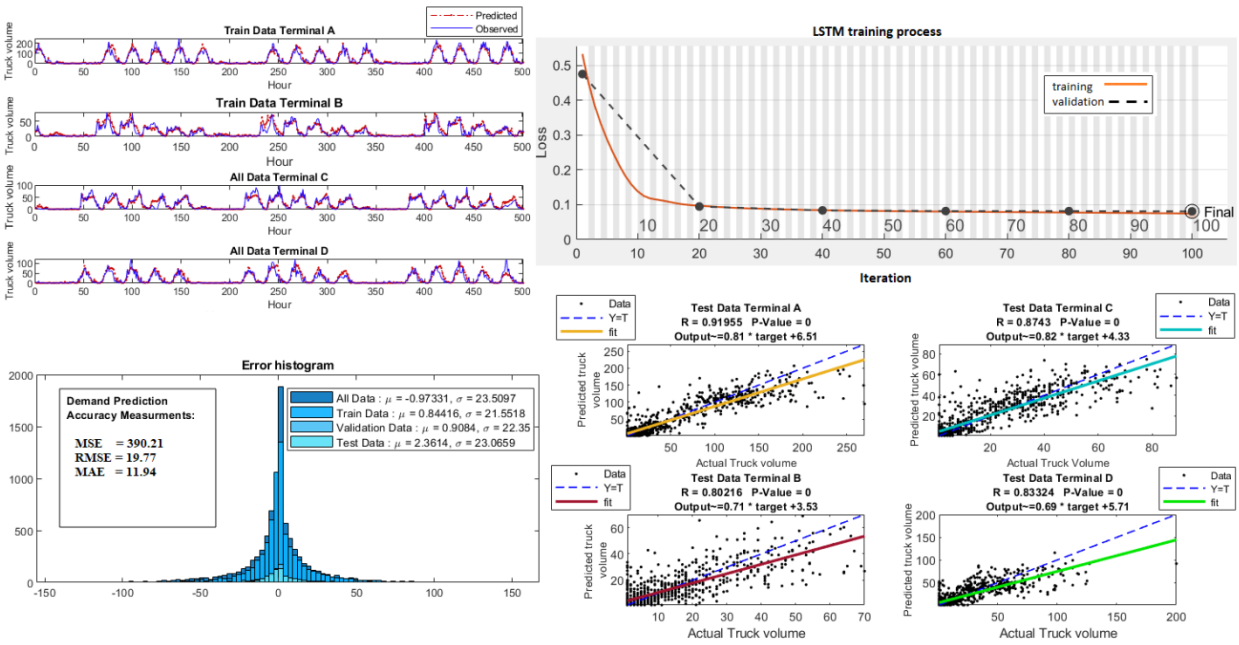


Figure 6-10: Performance of the demand model for 1 day ahead hourly demand prediction for each terminal

Several common accuracy measurements are used to show the performance of the model. The route means square error (RMSE) of this model is 19.77. Mean Absolute Error (MAE) is the indication of how big are the model errors on average. Figure 10 shows that the model can predict the truck arrivals with on average 11.94 error. The correlation coefficient with predictions and observed values are in all models close to 1 (between R=0.8 for terminal B and R=91 for terminal A) on test data. This confirms the high performance and generalization of the model to predict unseen data. The model is tested on different lookup windows as it should be able to predict the demand depending on the day of planning. In scenario 1, the request comes on a similar day as the operation day. In this case, the lookup window will be one and two days before the operation day and a similar weekday in the last three weeks. In scenario 2, requests arrive one day before the operation day. In this scenario, the lookup time is two and three days before the operation day and similar weekdays in the last three weeks. Finally, if time slot reservation takes place two days in advance, then the step of forwarding prediction is 2 days and the lookup window will be three and four days before the operation day and the similar weekdays in the last two weeks. Table 6-1 compares the average performance of this model for all terminals under various scenarios.

Table 6-1: accuracy of the model for different prediction horizons on test data

Scenario	Lookup window	Forward steps	MSE	RMSE	MAE
1	$t_o-1, t_o-2, t_o-7, t_o-14, t_o-21$	1 day ahead	390.21	19.77	11.94
2	$t_o-2, t_o-3, t_o-7, t_o-14, t_o-21$	2 days ahead	481.7	21.94	13.03
3	$t_o-3, t_o-4, t_o-7, t_o-14, t_o-21$	3 days ahead	483.25	21.98	13.38

A comparison between the performance of this model (for 1 day ahead) with historical average and regular LSTM time series shows that long-term demand predictions can be improved by sequence to sequence deep neural network structure.

Table 6-2: comparison of the model performance with other time series models

model	MSE	RMSE	MAE
Seq2Seq LSTM	390.21	19.77	11.94
Regular LSTM	450.42	21.22	12.8
HA	604.39	24.58	14.96

6.4.3 Marine Terminal gate model calibration and validation

As explained in section 6.3.2, the terminals' gate models are parametric. These parameters have to be determined with field measurements to ensure that the models can simulate the gate operations close to an average day in the real world. These parameters are λ_t , μ_t, and S which are truck arrival rate, gate average service rate, and the number of servers (lanes). Please note that the service time of the servers is generated using a constrained exponential distribution with service rate μ_t which will not generate service times less than 20 minutes. We calibrate four terminal models based on an average daily profile for weekdays since the statistical tests, e.g. ANOVA and two-sided t-test, showed that there are no significant monthly or daily trends in arrival or departure profiles. We have used the method of least squares to estimate these parameters, based on observed average truck arrivals and departures. For terminal five, we did not have truck counts observations at the vicinity of its gates. Therefore, we cannot calibrate a queueing model for this terminal. We, therefore, scaled the aggregated result of the four calibrated terminal models using loop detector data at the beginning of the road network before using it as an input in the traffic model. Figure 6-11 shows that the simulated truck departure volumes fit the observed volumes with R-squares greater than 0.9 for all terminals.

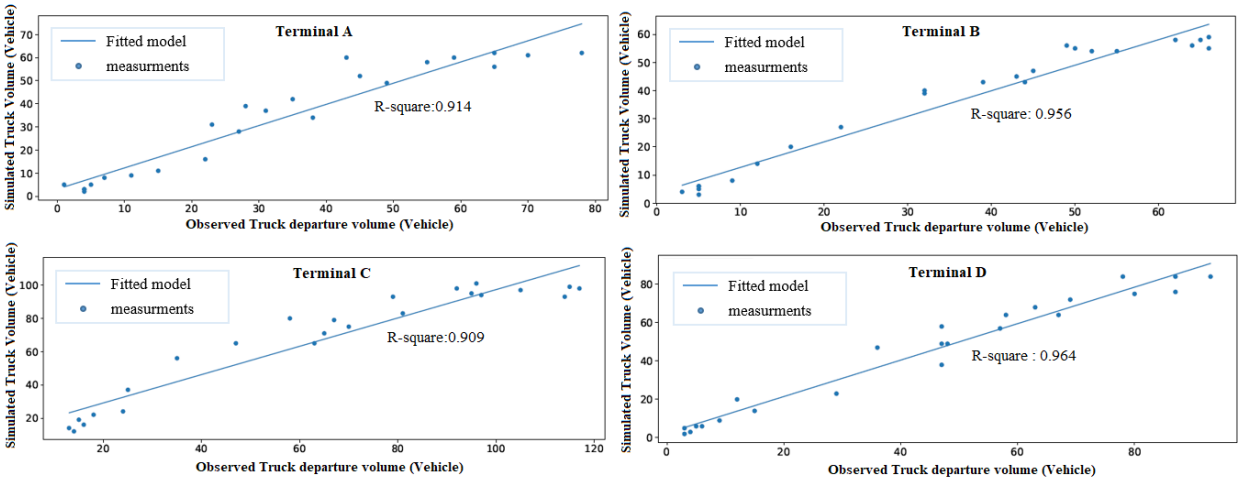


Figure 6-11: Terminals gate model fit

After parameter estimation, we used a two-sided t-test to statistically test the similarity between simulated and observed truck arrival signals as well as truck departure signals. The test results

show that the null hypothesis – The observed and simulated departure profiles are similar- is accepted with t-values 0.01, 0.014, 0.025, 0.029 for terminals A ,B, C, and D respectively.

6.4.4 Data-driven Truck scheduling model estimation

In this section, we explain the estimation process of the parameters of equation (6.15) and the results of this model. Assumptions about the distribution error term result in different model specifications. The simplest way of parameter estimation for such a model is restricting all convinces to be zero meaning that all the variables should be independent and identically distributed. This results in a multinomial logit model.

Table 6-3 shows the overview of attributes that we consider in this study to predict the utility of pick-up time windows. Since this model is one of the first in its kind, we selected these attributes based on expert knowledge at the port of Rotterdam and reviewing the most relevant studies (Irannezhad et al., 2019, Kourounioti and Polydoropoulou, 2018, Kourounioti et al., 2021). We believe the characteristics of containers like their type, size, and weight may logically influence the utility as some of these containers, e.g. refrigerated containers, are specialized for specific use at a particular time of day. Commodity type also represents the industry that uses these commodities in the hinterland which may have specific regulations in the time of day. Traffic conditions on surrounding road networks are another attribute that most logically influences the utility of time slots for carriers. Carriers often have some expectations about the traffic conditions at different times of the day. We, therefore, incorporate these expectations in our model using average inbound and outbound delays per km per time of day that truck drivers may experience approaching the port or heading towards the hinterland. Finally, we noticed that a noticeable number of carriers come to pick up the containers on the same day as the vessel arrives. Due to the discharge process time and container administration time, there is always some latency between the arrival of the vessels and the arrival of the trucks on the same day; therefore this could influence the utility of carriers for a particular time of the day.

The alternatives in our model are times of the day which are aggregated into four-time windows. The periods are formulated as morning (from 5:00 until 10:00), midday (from 10:00 until 15:00), and afternoon (from 15:00 until 19:00), night (from 21:00 until 5:00). These periods are based on observed arrival patterns and time slot categories used in practice at the terminals.

Table 6-3: Overview of attributes

Attributes	Description
Container type	GP: General purpose containers
	RE: refrigerated containers
	CC: chemical contained containers
	TC: Tank containers
Container Length	40ft: containers with 40 feet length
	20ft: containers with 20 feet length
Container weight	Categorized into:
	Heavy: 15000 to 35000 kg
	Light: 2000 to 15000 kg
	Empty: < 2000 kg
Commodity type	AGR: Agricultural
	Chem: Chemical products
	Food: Food products
	Fert: Fertilizers
	Pet: Petroleum
	RawMin : Raw minerals
	SolMin: Solid mineral fuels
	Ores
	Miss: miscellaneous
Vessels size	Large: if call size > 1250
	Small: if call size < 1250
Traffic conditions	Delay_Port: Average delay per km toward Port
	Delay_Hint: Average delay per km toward Hinterland
Planning latency	The differences between arrival of the vessels and arrival of trucks if that happens to be within similar days.

We present the result of this model in Table 6-4 including the model fit, estimates of the coefficients, and the level of significance (t-value). In this table, ASC represents the alternative specific coefficient which captures the mean on observed preferences for a specific alternative and β indicates the estimated coefficient for each attribute. The label of alternative-specific attributes is followed by the suffix, Mor, Mid, Aft, or Night. We have run several experiments with different specifications to get the best model fit with significant improvements in the likelihood ratio.

Table 6-4: Results of the truck scheduling model for all terminals (the base alternative is Night)

	Morning		Midday		Afternoon	
Coefficient	Estimate	t	Estimate	t	Estimate	t
ASC_{Night}	-0.251	-12.2	-0.251	-12.2	-	-
β_{Agr}	0.321	10.2	-0.128	-5.6	-0.147	-8.57
β_{Chem}	-0.101	-4.89	0.111	7.84	-	-
β_{Fert}	-	-	0.295	15.3	0.232	7.68
β_{Food}	-	-	0.411	14.3	0.275	7.11
β_{Iron}	-	-	0.278	5.88	0.356	6.36
β_{Miss}	-	-	0.2	11.8	0.161	5.85
β_{Ores}	0.144	5.04	0.263	11.8	-0.166	-5.31
β_{Petro}	-	-	0.197	10.1	0.089	3.01
β_{RawMin}	0.0891	3.26	0.0935	4.87	-	-
$\beta_{SolMinFu}$	-0.0838	-3.15	0.2	10.6	-	-
β_{GP}	-0.292	-14.2	0.312	18.1	-0.0808	-3.63
β_{RE}	-0.211	-7.11	0.194	8.26	-0.104	-3.42
β_{CC}	-0.352	-15.6	0.425	22.7	-	-
β_{TC}	-0.336	-11.8	0.277	12.7	-	-
β_{Vessel_Mor}	-0.233	-5.26	0.385	10.3	0.53	13.7
β_{Vessel_Mid}	-	-	0.425	13.5	0.771	24.2
β_{Vessel_Aft}	-	-	-	-	0.619	17.7
β_{Empty}	0.156	12.5	0.138	12.1	0.0979	8.14
$\beta_{HeavyWeight}$	0.0638	4.58	0.359	28.4	0.0826	6.1
$\beta_{LightWeight}$	-0.109	-8.1	0.151	12.2	-0.0374	-2.88
β_{Lenght_20ft}	-0.0563	-5.63	0.425	50	0.0758	7.63
β_{Length_40ft}	-0.204	-19.9	0.205	25.1	-0.132	-12.9
β_{Delay_Port}	-0.483	-13.9	-0.483	-13.9	-0.483	-13.9
β_{Delay_Hint}	-1.14	-29.7	-1.14	-29.7	-1.14	-29.7
Sample size: 303930						
Final log-likelihood: 397581.9						
Rho-squared with respect to constants: 0.19						
Number of parameters: 61						

Comparing the estimated coefficients of different commodity types, we can conclude that agricultural goods have a higher utility for morning time windows. This means that scheduling time slots for carriers with agricultural products will have a high cost for them in midday and afternoon. Ores and Raw mineral products also have a positive impact on the morning utility, However, their utility is slightly higher for midday. Iron is the only commodity type with the highest utility in the afternoon. The model does not suggest any significant preferences toward morning arrivals for fertilizers, foods, Iron, Miscellaneous, and petroleum which means that their preferences are more significant in the midday or the afternoon. All container types and lengths have a larger positive impact on utility for the midday. Please note the differences between the magnitude of the parameters indicated for the time slot management system to prioritize carriers based on their utilities. As opposed to the lightweight containers which have a negative impact on the utility of the morning time slots, the sign of impact for heavy containers is positive for all alternatives. However, the impact of afternoon time windows on the cost of carriers is lower than morning. Regarding planning latency, truckers that are willing to pick up a container within the same day of its arrival at the terminal have higher preferences for the next time windows. For example, if the vessel will arrive in the morning, there is strong disutility for the morning time slot for trucks - obviously because containers have to be first discharged from the vessel and only then become ready for pickup. Having this in the model

helps the time slot management system give priority to these truckers to reserve a slot in the afternoon.

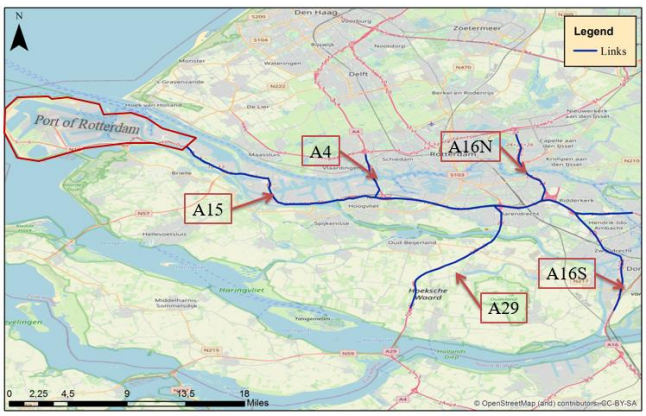
The sign of delay on surrounding road networks is plausible (negative). This means that the higher the delay per km, the lower the utility of the truckers. Comparing the magnitude of parameters for delays shows that the utility of carriers is more sensitive to the Hinterland direction as compared to the Port direction. It is important to mention that traffic conditions change rapidly on the road network. Therefore, we used the actual delay for the hour of the day that is associated with the pickup time of each container. In other words, delays are not aggregated for the range of alternatives. Please also note that delay is used as a generic parameter for all alternatives.

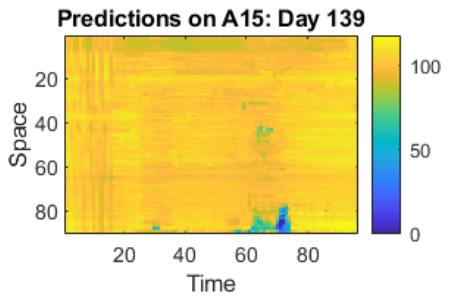
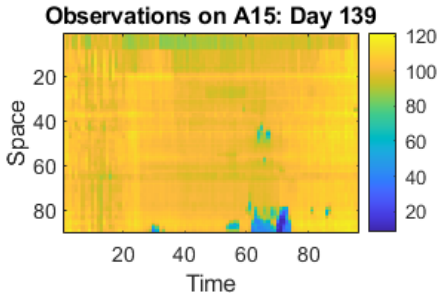
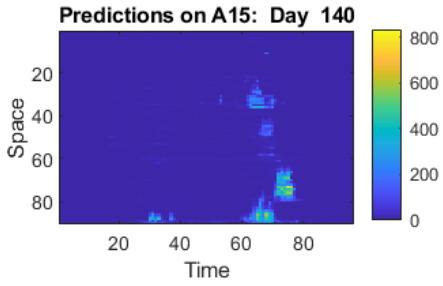
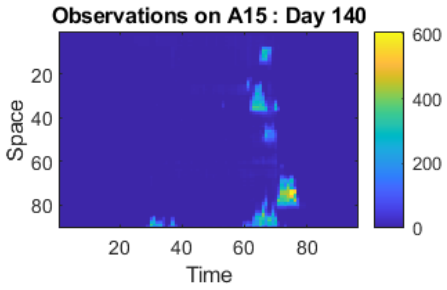
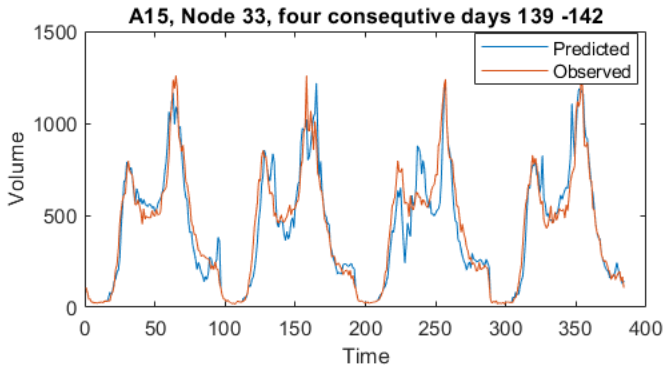
In sum, this model provides plausible estimates and will support the time slot management system to consider the utility or cost of each time window for carriers.

6.4.5 Traffic model estimation

The parameter of the proposed traffic model is estimated for the 181 days of data collected for the given road network. In **Table 6-5** we summarize these network specifications and prediction accuracy for speed, flow and vehicle loss hours, which are then converted to the monetary loss providing traffic cost for the time slot management system. As one can see, the model can predict time-space traffic conditions accurately.

Table 6-5: Data and Network description

Network description	
Aggregation in time	15 min
Aggregation in space	600 m
Number of nodes	166

The maximum length of the path	A15 : 55 km ; A4: 4,8 km; A29: 11,4 km; A16 North: 7.1 km; A16 North: 20.1 km	
	Predictions	Observations
Speed		
Loss hours		
Flow		

We used the Mean absolute percent error (MAPE) to evaluate the accuracy of this model.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{t_i - y_i}{t_i}\right| \times 100\% \tag{6.41}$$

where t_i is ground truth and y_i is the predictions. We calculated MAPE under various equally distributed error bins. This will result in a more realistic percentage of errors. the overall MAPE for speed prediction in the network is 11%, for loss hours 14%, and for flow 9%.

6.4.6 Time slot management simulation and optimization

The simulation experiment follows the steps explained in section 6.3.5. We selected a day with congestion between 15:00 and 18:00 as the simulation day. We use the container pick-up time reported in the PCS data for the selected day to generate requests. For this, we randomly but proportionally draw from the PCS data. We considered 1 hour for the time to respond on the planning day (t_o). The request generator generates requests for each hour between 7:00 and 17:00 which is aligned with the working hours of the trucking companies. The sum of all requests for each hour should be equal to the total demand of that hour. For each planning window, we use a MATLAB solver to re-schedule requests solving the optimization problem presented in section 6.3.5. In the year 2017, the waiting cost for trucks in container transport was estimated at around 38 euros per hour. The hourly cost of a crane is twofold. One relates to the labour cost which is equal to 2.5 people per gate lane (Guan and Liu, 2009b) and the other is the cost associated with the yard operation if the crane will be active in the hinterland side. Given 50 Euros for labour costs, we can calculate the cost of using a gate as follows:

$$C^S = 2.5\times 50 + \alpha[C^W_{peak}(s) - C^W_{peak}(S+1)] \tag{6.42}$$

where $C^W_{peak}(s)$ is waiting time cost during peak hours. The idea is that the value of a crane for yard operation is at least equal to the benefit of the waiting-time reduction in the hinterland adding one more crane to serve outbound trucks. Since for the calibrated terminal models we have the estimated S for the peak hour (see section 6.3.2), we can now estimate this cost for the constant service and arrival rates running the terminal model with S+1. In equation (6.42), $\alpha > 0$ is a factor that can be tuned by terminal operators depending on the priority they may want to give to the yard operation, for instance, if large vessels are at berth. Based on our models and with $\alpha = 1$ we came up with approximately 80 Euros for the disutility of the terminal operator to add an extra crane. Therefore, we consider $C^S = 205$ Euro per hour per crane. We refer readers to equations (6.18) and (6.29) for the calculation of the re-scheduling cost for carriers changing their schedules C^P_t and the cost associated with the delays on road networks C^{tr}_t respectively.

Given these costs, the optimizer gives a set of solutions that are equally efficient and which together represent the Pareto frontier of search space. This Pareto frontier is difficult to visualize as it has four dimensions. The selection of one solution among all optimal solutions is straightforward, however, it depends on the policy of the decision-maker. These solutions range between an extreme focus on supply (terminal costs) and an extreme focus on demand (carriers). If the decision-maker focuses only on terminal costs (supply) and therefore selects solutions with only the lowest number of cranes as compared to the base case, this increases the waiting time for a constant arrival rate. Therefore, the optimizer reduces the arrival rate by assigning more trucks to off-peak hours which results in reduced waiting times. This, however, increases the disutility of carriers since they have to bear more re-organization costs. The cost of the traffic system changes as well due to the changes in the arrival and departure of trucks. The magnitude of the traffic cost, however, depends on the traffic conditions. In this scenario, although the solution is optimal, all the pressure will be on the carriers. In another scenario, the focus may be on the demand side (waiting time). In this approach, the system increases the number of cranes, and the waiting time drops accordingly. There are no changes needed in the arrival profile of trucks and therefore the re-organization and traffic costs will be zero in this scenario (highest carriers satisfaction). In this scenario, also the solution is optimal, but all the pressure will be on terminal operators.

In both above scenarios, only one stakeholder should take the action for the benefit of the other stakeholders in the system. Such solutions are less probable to succeed and more difficult to implement in practice since the costs of one stakeholder will be relatively high and it requires extra efforts and governance to collect a part of the gain from other stakeholders to compensate and incentivize the actor. Due to these difficulties, the most logical way is to select solutions that involve all stakeholders in the operation and balance out their costs and benefits. Among the solutions that satisfy this requirement, we select the one with the maximum monetary gain by comparing the waiting times before and after applying time slot management. The simulation result of TSMS with respect to the reduction in waiting time is illustrated in Figure 6-12. We can see that the waiting times are spread across other times of the day resulting in extra waiting times during the morning. The overall state of the system, however, obtains significant improvements.

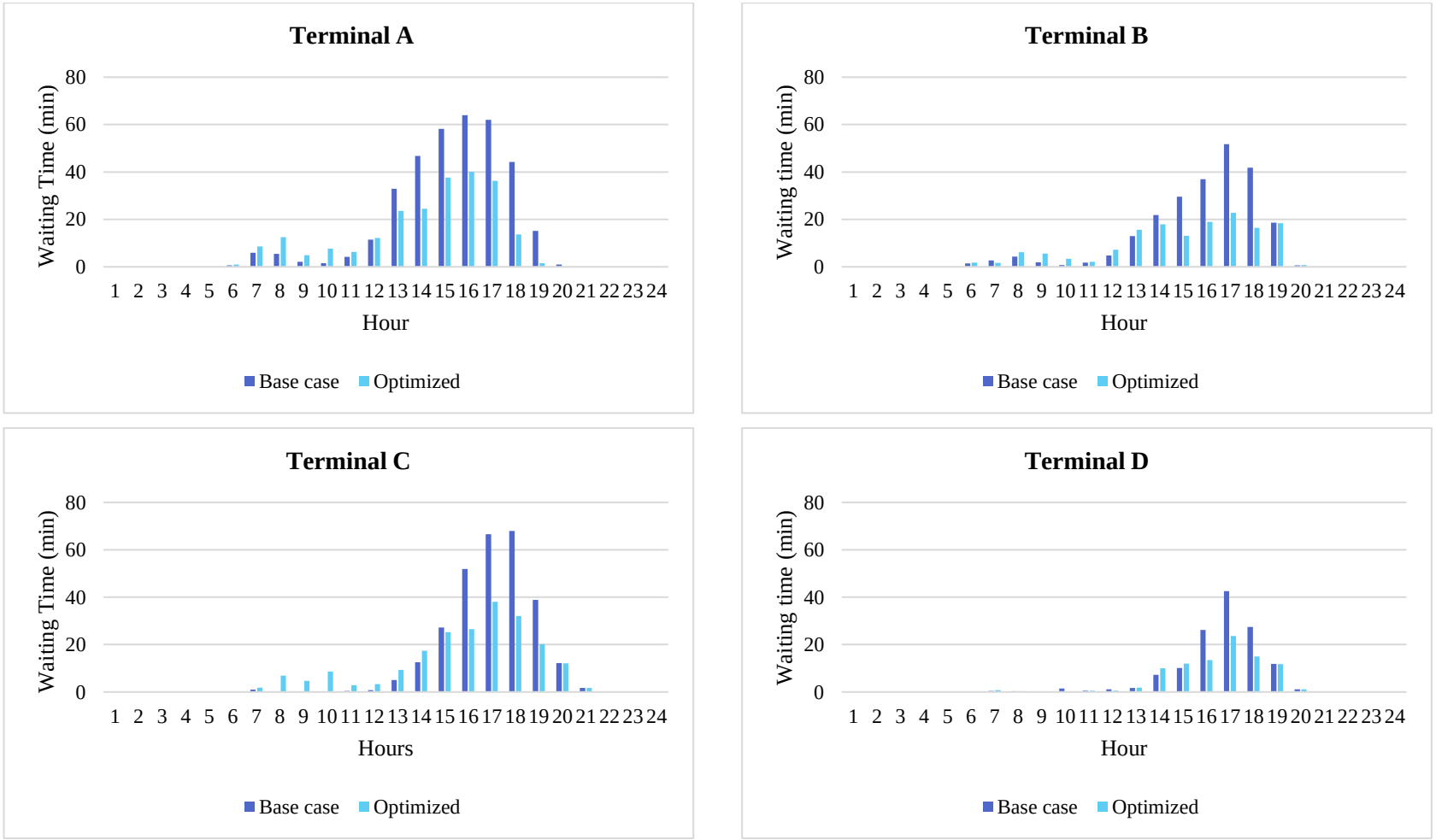


Figure 6-12: waiting time comparison between the base case and optimized day

In Table 6-6, we show the cost and benefits from the perspective of all stakeholders in the system.

Table 6-6: Cost and benefit of stakeholders with TSMS (simulation of one working day)

Stakeholder	Total gain [€]	Road Network [€]					Terminals [€]			
		A15	A4	A16N	A16S	A29	A	B	C	D
Terminal operator	1025						615	205	205	0
Trucking companies	15011						4890	2905	4922	2294
Trucking companies	-221.02*									
Traffic system	1310	1904.3	250.16	170.27	-1235.8	178.5				
# of trucks rescheduled	231 (~10%)									
Computation time	00:8:23									
Waiting cost Optimality	0.03									

*Disutility is unitless and should not be considered as monetary values

As we can see from Figure 6-6, Trucking companies can gain 15011 euros on the day of the experiment experiencing less waiting time with only 10% of requests being shifted to off-peak period. Besides this gain, truckers can be more productive during the day if they don't have to wait at terminal gates. If we use 62 euros per hour as the cost of transporting a container (van der Meulen et al., 2020), dividing the total gain from less waiting time by 62 gives us the number of hours that truckers can be more productive along the day. In our case, the productivity gain of this solution is 241 hours.

The shifts in time slot requests, however, result in a total of 221.02 disutility for 231 trucks that have to reschedule their hinterland operations. Please note that this number is unitless and its magnitude cannot be compared with other costs in the system. This indicator shows how dissatisfied these carriers are from the perspective of their hinterland operation. Lower waiting time at gates resulted in a lower number of cranes being used by terminal operation in the hinterland side which leads to 1025 Euros per day (this gain will be higher for $\alpha > 1$). To investigate the relation between the gain of terminal operators and the disutility of carriers, we visualize the Pareto frontier of optimal solutions. Figure 6-13 shows that if we reduce the number of active cranes (reducing the cost of terminal operations) the TSMS system has to redistribute more requests to control truck arrival rates and consequently reduce waiting time at gates. This increases the disutility of carriers exponentially as they have to reschedule their hinterland operations. This graph can help decision-makers to tradeoff between the satisfaction of these two actors.

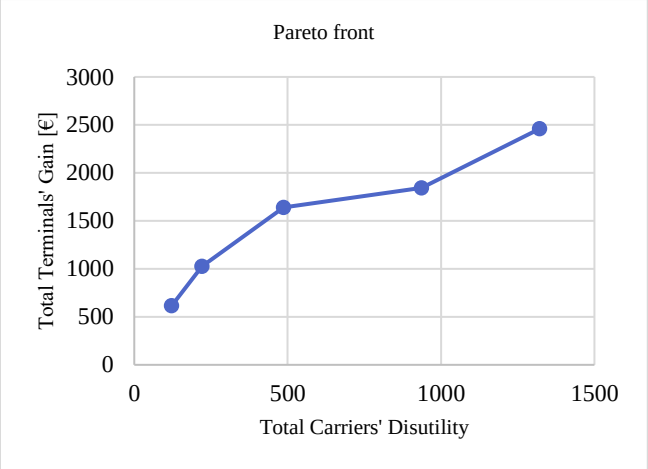


Figure 6-13: Terminal Gain Versus Carriers disutility in Pareto frontier of optimum solutions (only solutions with positive terminal gains are illustrated)

From the traffic perspective, we have small disturbances on the road networks between 15:00 and 18:00 on the day of the experiment and a large truck arrival rate for the same period. This is why TSMS has resulted in positive gains on 4 out of 5 motorways in the road networks by controlling truck arrival rates. The only motorway that has a negative gain (increased travel time along the day due to this operation) is A16-South which happened to have no congestion in the base case. The total gain of 1310 euros is relatively small and implies that TSMS not only prevents truck re-scheduling to deteriorate traffic conditions on road networks but also has a small impact on the improvement of traffic on the selected network. This impact however could be higher for larger networks with more severe congestion.

In Figure 6-14 we show the distribution of re-scheduled containers over commodity and container types. This shows which transport markets are more influenced by this system. As we can see that 34%, 27%, and 20% of all request that has been shifted belongs to Solid fuels, Chemical products, and Agricultural products respectively. General-purpose container, reefer, and chemical content containers have 74%, 15%, and 11% of the share of shifted requests. The reason is that these markets have lower disutility (see Table 6-4) for other times slots as compared to other markets which makes it possible for TSMS to shift their requests. These statistics only belong to the simulated day and more insight can be provided by TSMS if the system works over a longer period. Since this decision support system only rejects requests that have the minimum shift costs, such information or statistics can help decision-makers to identify the transport markets that have the potential to re-organize their hinterland activities and make a larger contribution to more efficient gate operation.

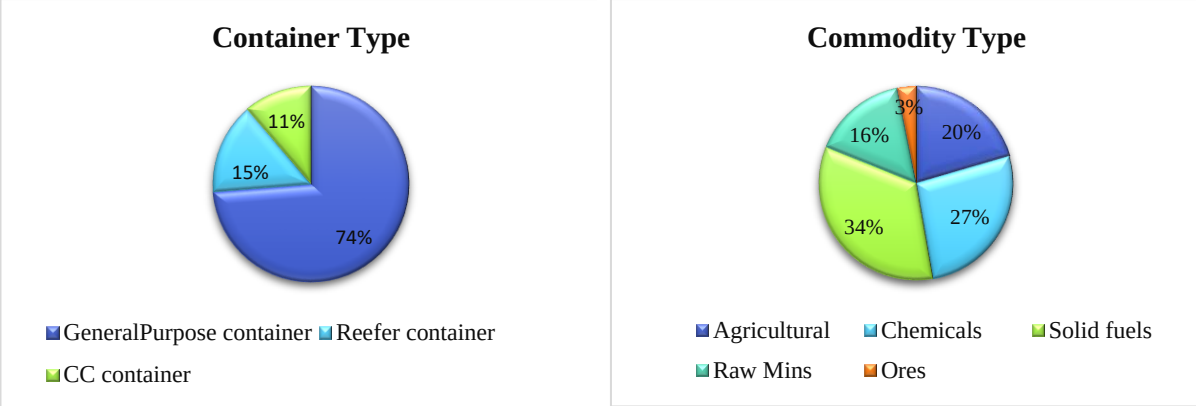


Figure 6-14: Share of shifted requests based on commodity and container types

6.5 Discussion

In this section, we summarize our main findings and discuss the potential of the proposed decision support system for time slot management in marine container terminals regarding its application for practice. We also discuss the limitations of the model and provide a set of recommendations for future research.

The results demonstrate the significance of integrative traffic and logistics modelling for use in terminal planning and decision support systems. This integration takes place considering the collective perspective of multiple actors which sheds light on the costs and benefits of the system. The data-driven truck scheduling model in the proposed TSMS provides insights into all these behaviours and preferences of truckers and hence can help decision-makers to interpret the result of this decision support system. For example, a decision-maker, by looking at the time slot preferences of all transport markets, can identify those in which their request is more frequently shifted to other time slots.

The costs and benefits of the stockholders involved in this system are not balanced. The gain for all stakeholders can be obtained at the cost of the disutility of a portion of carriers that are re-scheduled by the system. Our analysis shows that there is a benefit in the system for everyone but only truckers have to take action and bear the cost of re-organizing their schedules. From our simulation experiment based on the real case study, it is evident that the disutility of carriers increases if they have to re-schedule their tours to different pick-up times at terminal gates. Although taking into account the market preferences of carriers in TSMS minimizes the disutility of carriers, it still may put the success of the system at risk. Decision-makers may use the insight from TSMS to redistribute the collective benefit of the system to compensate the cost of carriers for equity. These incentives however require a gain-sharing mechanism to keep carriers motivated. The design of such a mechanism is beyond the scope of our study. However, we can picture it as tax reduction policies, giving priority to shifted trucks of specific markets that follow the TSMS recommendations, or defining incentive packages.

The first step towards designing such a system is to understand the monetary value of the disutilities. This requires a more in-depth analysis of hinterland pre-trip tour planning decisions like routing and scheduling of carriers. In this research, we only addressed hinterland scheduling decisions due to the available data. For calculating the monetary value of disutility

of carries, however, a combined analysis of routing and scheduling decisions on a set of observed tours is required.

Finally, the impact of TSMS on road network traffic is highly dependent on the conditions of the road networks under study. The TSMS is not aimed at traffic management to help improve traffic conditions but can help to reduce the negative impact of the changes in port activities on traffic conditions. In addition, TSMS can provide predictions for traffic agencies. These can support the application of traffic management measures to reduce the possible negative impact of truck flows heading towards the hinterland or port. This highlights the importance of incorporating traffic modelling in the design of TSMS.

6.6 Conclusions

To deal with large waiting times, this chapter presents a novel, centralized decision support system for time slot management at the marine container terminal. We explain the required steps toward the implementation of this system and evaluation of its performance under a simulation experiment, supported by real-world data. Our contribution is twofold.

Firstly, our method integrates logistics, freight and traffic modelling. It combines (1) advanced deep neural network structure for accurate truck demand predictions at terminal gates, (2) queuing modelling to simulate terminal gate operations, (3) utility maximization modelling to consider the disutility of truckers when assigning time slots, (4) a data-driven traffic model which can learn from day-to-day traffic to link logistics activities at the port to the dynamics of the traffic system and, finally (5) a nonlinear mixed-integer formulation to minimize multiple costs in the system. We conclude that the model can perform well in predicting truck demands (up to 91% accuracy) at terminals gates, and predict traffic states on road networks with a 14% average error while linking these predictions to truck demands. The performance of the time slot management system indicates a significant reduction in waiting times at terminals' gates giving accurate suggestions to all the stakeholders. Secondly, the model provides new insights into the trade-off between supply (terminal operations) and demand (trucking operations) considering their impact on traffic at the port (terminal gates) and on-road networks in the hinterland. We infer the relation between the disutility of carriers and their tour planning decisions. By rescheduling a small portion of requests of trucks for time slots, the system obtains remarkable waiting time and productivity gains. We also find, however, that these gains come at the cost of carriers' processes. The system gives priority to tucks based on their preferences hence minimizing their costs by suggesting appropriate timeslot to drivers. We, therefore, conclude that the system coordinates well between multiple stakeholders in the port-hinterland ecosystem and provides an inclusive assessment of the costs\benefits of the system.

With respect to future research, we recommend including internal operations in the terminal model including the vessel discharge process. This could help the system to develop a broader view of the system's shadow costs. We also believe that insights into the monetary value of disutility of carriers, which can translate as truck scheduling cost, require more in-depth analysis of tours and trip chains in the hinterland. In addition to the stakeholders considered in this study, some of the shippers' and freight forwarders' decisions may have a direct or indirect impact on this system and therefore, requires further investigations. The application of such a decision support system can also be tested in digital twins of the port-hinterland ecosystem where policymakers can monitor day-to-day interaction between port and hinterland stakeholders.

Chapter 7

Conclusion

This chapter discusses the salient outcomes, highlights the significance of the study, and discusses opportunities to extend the research in new directions. We first begin with a summary of our key findings. Next, we explain the contribution of this study to science and practice. Finally, we provide our recommendations for future research in the field as well as several important implications for future applications of the models in policy and management.

In this research, we identify current challenges of understanding the complex interaction between freight transport and traffic on road networks, in particular considering the spatial and temporal dynamics of both systems. These challenges arise, on one hand, from the limitations of theory-driven modelling of tour planning decisions and, on the other hand, a lack of empirical evidence on the impact of traffic on tour planning decisions. The methods presented in this thesis improve freight modelling in several ways and further the understanding of these mutual impacts:

- by discovering knowledge from mining large freight trip databases identifying generalized freight tour characteristics with respect to the time of day scheduling and type of tour routing decisions. [chapter2] ;
- by introducing a data-driven vehicle routing and scheduling model that can reproduce pre-trip tour planning decisions [chapter 3];
- by investigating the delayed correlation between the schedules of freight transport activities in logistic hubs (like the port of Rotterdam) and the variation of truck volumes on a motorway.[chapter 4];
- by exploring the evolution of the impact of freight trip scheduling spatially and temporally on a given road network and exploring the application for evaluation of truck departure time shift policy [chapter 5]; and

- by integrating logistics, freight transport, and traffic models in a unified framework and evaluating its use in decision support systems that help manage and provide insights into the port-hinterland ecosystem. [chapter 6]

Together these contributions lead to substantial insights into the freight transport system and its interaction with traffic phenomena.

7.1 Main findings

In this section, we discuss a summary of each chapter's key findings relating them to the research questions proposed in Chapter 1.

Research question 1: *How can freight trip data sources be enriched by combining them with logistics and traffic data?* [chapter 2]

In **chapter 2**, We developed a systematic and standard data preprocessing activity pipeline to enrich massive freight trip data with logistics and traffic data sources. This pipeline includes data transformation, data segmentation and data integration. In data transformation, we collected massive speed data from loop detectors on road networks and transformed them into aggregate contextual information about congestion levels at the course of geographical zones. This contextual information is added to the pickup and drop-off locations reported in the freight trip database. This provides us information on the variation of freight activity patterns based on the perception of the congestion. Freight trip data collected from the daily movement of trucks in between pickup and delivery locations in the Netherlands include detailed information about the characteristics of trips and commodities but do not contain information about the logistic characteristics of the senders or receivers of the goods. In our data preprocessing pipeline, we propose a systematic multi-scale data integration approach to enrich the freight trip diary with firm registry data. Using this approach, we label the origins and destinations of shipments with the most probable logistic activity that matches the attributes of the inflow and outflow trips. These labels include producers, wholesale warehouses, distribution centers, and consumers of goods. The enriched database, thus, provides a rich basis for comprehensive studies into logistic behavior around zones with particular logistics activities.

Research question 2: *How can patterns in departure time (scheduling) and space (structure of tours) be understood from mining freight trip data sources?* [chapter 2]

After data integration for enriching freight trip data, we developed a tour-based segmentation method in **chapter 2** to cluster data based on discovering frequent patterns in pickups and deliveries *between* industrial sectors. This helped us to identify 9 distinct transport markets with homogeneous transport activities. This provides improved possibilities to model the two key structural characteristics of tours in freight transport: time of day scheduling and tour patterns.

To distill meaningful and acceptable knowledge about scheduling and routing of these transport markets, we propose a new enhanced decision tree algorithm (Multi-Task Decision Tree) that predicts multiple discrete and continuous response variables. This algorithm can accurately predict the type of tour (direct, collection, and distribution) and link it to the number of stops with a high goodness-of-fit. Similarly, we used Decision trees to model the time of day characteristics of tours. This model can extract knowledge from the existing patterns in the departure time of tours with a slightly lower but still reasonable goodness-of-fit.

The results from this study have produced new knowledge that contributes to our understanding of the activities of freight transport markets. For example, our method brings into the discussion the differences and similarities of different transport markets in applying a particular type of tour when facing congested zones. It also provides deep insights into the within-peak or off-peak scheduling of tours. The study has led to many more detailed findings, as mentioned in the discussion section of **chapter 2.** The most salient one is that some transport markets have strict rules regarding the time of day and type of tour when visiting logistic hubs. For example, in general, visits to distribution centres occur during morning and afternoon peaks even if the zone of the distribution centre is congested during the visit. Surprisingly, this observation is not limited to the expected cases of relatively high-value goods like miscellaneous products, parcels, or flowers. These findings show the surprising behaviour of a group of carriers in some transport markets or around some logistics zones. This finding suggests that these differences between goods should be taken into account in freight transport models. Regarding the types of tour patterns, the results of this study indicate that tour planners tend to plan direct tours if they have to visit non-congested zones and rather plan distribution or collection tours otherwise. As expected, these rules differ from one market to another. It is, therefore, very difficult - if not impossible - for theory-driven methods and normative models to incorporate all this information. This may be a reason why the results of theory-driven models are often less accurate than expected. In conclusion, the results highlight the importance of developing new data-driven methods that can learn preferences of tour planners for different transport markets.

Research question 3: *How can we develop a generalized data-driven routing and scheduling model for the tour-based representation of time-dependent freight transport activities, based on partially observable tour data?* [chapter 3]

In **chapter 3**, we developed a data-driven time-dependent vehicle routing problem with trainable parameters to model collective tour planning behaviour of planners learning from a set of planned and executed freight tours. We formulated tours as a vehicle routing problem with a weighted sum of multiple objectives. The aim is to estimate these weights in such a way that the output of the model fits the observed tours' characteristics. Since parameter estimation of this model requires the extensive computation of the tours iteratively, we utilized Bayesian optimization as a fast and reliable global optimization approach to deal with this problem. The estimated parameter of this model is completely interpretable with respect to the distance-related costs and travel times. Also, it provides insights into the characteristics of tours starting during different times of the day. The performance of the model regarding estimated and observed tours show that the proposed model can reproduce tours with low error, with the highest errors belonging to tours that start during the afternoon. We tested the ability of the model to reproduce aggregate observed freight transport features like the distribution of the number of stops and total kilometres travelled in tours. The model reproduces these indicators with high accuracy. With these methods we successfully investigated how congestion and variations in travel times can influence tour planning decisions.

Research question 4: *How can we model the relation between freight trip scheduling and the volume of freight vehicles on a motorway?* [chapter 4]

In **chapter 4**, we analyze the short-term relation between schedules of trips resulting from activities in a logistics hub and dynamics of truck volumes on a motorway. We used Port of Rotterdam as the case for the logistic hub and developed a multi-layer feedforward neural network with automatic feature extraction (MLP-AFE) to probe into the delayed correlations

between trucks scheduled to pick up containers at terminals with the volume of trucks on the nearby motorway. The model predicts one hour ahead of the truck volumes using information up to 9 time-steps back in historical data of trip schedules. We tested the accuracy of the model for various prediction resolutions from 5 minutes to 1 hour and the results indicate high accuracy of the model for all resolutions. The feature importance analysis shows that 1 time step back in the history of trip schedules has the highest contribution to the prediction of truck volumes on the motorway. Our analysis also confirms the functionality of the model to be responsive to disruptions in port activities. In sum, the findings from this study provide evidence for the way freight trip schedules influence a nearby motorway and support the exploration of this impact on a road network. This, however, requires taking into account the evolution of truck volumes on road networks in time and space.

Research question 5: *How can we incorporate spatial and temporal dependencies in predicting traffic dynamics from freight trip schedules?* [chapter 5]

We used a graph representation of the road network to consider its topology as well as its temporal characteristics. In **chapter 5**, we developed a graph convolutional neural network to predict traffic states (flows, speeds, and delays) mapped on the graph representation of the road network. These predictions are based on the variations in freight trip schedules coming from the port of Rotterdam. To incorporate traffic flow theory in our data-driven traffic model, we proposed a special type of attention mechanism that can help the model learn and capture spillback phenomena once we have congestion on a link. All in all, this method lets us build a demand-responsive data-driven traffic model that can link space-time variations in the demand side of the system (freight and passenger activities) to space-time dynamics on the supply side (traffic on road networks). The results indicate that the prediction accuracy of the model is very high at link level, for speed, vehicles volumes and vehicle loss hours. Our sensitivity analysis also shows that this traffic model can accurately show traffic response to the changes in freight trip schedules. This brings the opportunity to assess the impact of shifts in the truck departure times on the traffic system.

Research question 6: *What are the impacts of truck departure times shift policy on traffic conditions on road networks?* [chapter 5]

In **chapter 5**, we designed an optimized truck departure time shift policy for the case of the port of Rotterdam to remove a portion of trucks from road networks during afternoon peak hours. The societal travel time gained from this policy is significant if the selected road network is congested during the study period. The results indicate a range of 4 to 10% travel time saving on the road network. The monetary value of this gain can be collected from other road users and returned to the participating carriers. The findings show that if only 10% of trucks change their schedules, they can receive around 15 Euros bonus for each container. This however requires a decent gain sharing mechanism to involve government organizations in this intervention.

Research question 7: *How can we design a time slot management system for seaport terminals to mitigate congestion at terminals' gates considering road networks conditions, scheduling of trucking companies, and terminal operations?* [chapter 6]

Changes in departure times of trucks to mitigate congestion require making alignments between several actors in the port and hinterland ecosystem. In **chapter 6**, we developed a centralized decision support system for a time slot management system to facilitate shifting trucks from

peak hours to off-peak periods. This system links port operations to the hinterland operations considering the costs and benefits of terminal operators, truckers, and the traffic on road networks. The system includes several modules working together to facilitate time slot management to work efficiently. To begin with, we developed a sequence-to-sequence deep learning structure to predict truck demands at the port after which a queuing model is used to simulate marine terminal gate operation generating terminal crane utilization costs, the costs from queue length, and truck departure volumes. After that, the generated truck volumes are used as the input for the traffic model developed in **chapter 5** to calculate the cost or benefit of the traffic system from rescheduling truck requests. Finally, we developed a data-driven truck scheduling model based on the theory of utility maximization. The later model helps calculate the disutility of carriers for a time slot with respect to their hinterland activities and preferences. We used this disutility as a proxy for the cost of carriers to re-schedule their hinterland tour activities due to the shifts imposed by the time slot management system. We utilized all these costs by means of simulation experiments and optimize the system using multi-objective nonlinear mixed-integer programming. The most prominent finding from this study is that our decision support system can significantly reduce the waiting time cost of all carriers (at the scale of million Euros per year) at terminal gates by only re-scheduling around 10% of the time slots requests. These shifts also help container terminals manage the use of cranes between the hinterland and yard sides which adds to the benefits of the terminal operations. In addition, carriers can gain 241 hours of productivity as a result of reduced waiting times at terminals' gates. This means that, with the fixed fleet size, carriers can transport around 10% more containers in the hinterland. All these gains come at the cost of the disutility of those carriers who are obliged to reschedule their tour planning in the hinterland due to shifts applied to their container pickup time slots. The proposed system quantifies this disutility and provides insights into the tradeoff between demand (shift trucks' requests) and supply (increasing number of cranes in the hinterland side) obtaining a maximum waiting-time gain at terminal gates while keeping the disutility of the carrier at a minimum. This system also ensures that shifts in truck schedules at terminals do not deteriorate traffic conditions on the road network, keeping the collective vehicle loss hours on the road network minimum. The system presented some improvements, although marginal, in the traffic system.

7.2 Main conclusions

In general, we conclude that integrating freight transport and traffic system expands our understanding of the system and can be of assistance to policy implications and decision support systems. In this section, we draw several conclusions from the main findings of this study.

- Insights into the integration of freight, logistics, and traffic data

We conclude that integrating logistics, freight transport, and traffic data in a systematic way provides new insights into freight transport activities and is necessary for more comprehensive freight transport modelling. In **chapter 2**, we show how this integration shed new light on the structure of tours. In **chapters 4,5** and **6** we also linked data from logistics activities in maritime container terminals to the traffic data which provides insights into the port-hinterland ecosystem.

- Insights into transport markets

From the transport market segmentation using the pickup and delivery activity of carriers, we conclude that this new way of classification provides deeper insight into the interrelation between different industries. This new method could help to improve freight transport modelling due to clustering transport markets into homogeneous transport markets. In

chapter 2 we showed the importance of this segmentation in analyzing activities of freight transport.

- Insights into the spatial dimension of tours

The spatial dimension of tours relates to the routing decisions of carriers, In **chapter 2**, we predict the type of tour and its associated number of stops and link these predictions to the level of congestion at the pickup and delivery locations at the time of visit. From this analysis, We conclude that the type of tours and the number of stops are sensitive to the level of congestion. Therefore, this is important to consider modelling type of tour decisions in freight transport taking into account congestions at logistic locations.

- Insights into the temporal dimension of tours

The time dimension of tours relates to the scheduling decision of tour planners. In **chapter 2**, we extract meaningful and strong rules that can explain departure time scheduling decisions of tour planners under various externalities, including congestion. From this analysis, we conclude that there is a diversity of behaviours around different logistics zones and different good industries regarding the scheduling of tours in congested areas. This highlights the importance of considering these differences in freight transport modelling.

- Insights into tour planning

The discovered knowledge from analyzing the time and space dimensions of tours emphasizes the need for developing a data-driven method for modelling tours. The new method proposed in **chapter 3** allows us to jointly capture routing and scheduling dimensions of tours from data. This makes it possible to predict time-dependent movements of freight vehicles in between geographical zones in a more realistic way that has agreements with observed tour activities. From this experiment, we conclude that although planners are sensitive to the travel time variations and try to avoid routes with high travel times, they value travel costs more in routing and scheduling. They also differentiate between the different time of day intervals. We also conclude that using a data-driven vehicle routing and scheduling model can realistically predict and represent freight tour activities in time and space.

- Insights on methods and parameter estimation

From the findings of **chapter 3**, we conclude that even with some aggregate tour information, we can formulate tour planning problems in such a way that the parameters of the model can be estimated efficiently and with reasonable confidence.

- Insights into the demand-responsive short term traffic state predictions

In **chapter 4**, we proposed a shallow supervised neural network that expressed the strong correlation between traffic supply and demand considering logistics-related flows. In **chapter 5**, we showed that the spatial and temporal traffic dynamics are predictable from the spatial and temporal dynamics of freight transport demand. From this experiment, we learned that using neural networks to predict short-term traffic can become demand-responsive and explainable if we use our knowledge of the traffic flow to design the structure of the model. We, therefore, conclude that policymakers and traffic agencies can rely on these predictions to make short-term decisions and apply appropriate interventions to deal with the unreliability of travel times on a given road network.

- Insights into time-shifted freight transport

In **Chapter 5**, we explore the possibility of a time-shifted freight transport policy at the Port of Rotterdam. From the finding of this research, we conclude that by shifting the departure time of only 10% of trucks, there are measurable monetary gains that can be obtained from the traffic system. Collecting this monetary gains from the traffic system, however, requires the intervention of multiple actors, especially governments. Governments can collect this money by applying taxes, for instance, and give the money to the shifted carriers as incentives.

- Decision support systems for port-hinterland alignments

In **Chapter 6**, We extensively look into the application of our methods in a decision support system for the case of the port of Rotterdam to make alignments between terminal operators, carriers, and traffic on road networks. We conclude that using an integrated logistics, freight transport, and traffic model explains jointly the costs and benefits of different actors in the system while controlling the truck inflow demands at terminal gates. This method, since having support from real-world data, can assist multiple stakeholders in their tactical and operational decisions.

7.3 Recommendations

We believe that the contributions of this study to the field open up several important opportunities for future research and practice.

7.3.1 Recommendations for research

This thesis devotes the research to improving freight transport modelling taking the mutual impact of routing and scheduling of carriers on the traffic system. For this aim, we developed several methodologies and discussed the key findings. In this subsection, we discuss the potential research path that can benefit from the outcome of this study.

- Comparison of transport markets with standard industrial classification
  In chapter 2, we proposed a new classification technique to identify transport markets from their tour activities in between different industries. A comparison study will cast new lights on freight transport modelling showing if and how these transport markets are different or similar to that of standard industrial classifications like ISIC and NACE which are mainly only based on the economic activity of industries.

- Validation of discovered rules regarding structure of tours
  In chapter 2, we discovered a set of significant rules that can explain the characteristics of types and departure times of tours. Although these rules have strong support from data and are statistically significant, We are required to validate them with views from practitioners to better understand the reason behind the found patterns. In chapter two we justified some of these rules with previous studies and our expectations. However, continued efforts are needed to validate and reason these findings with expert knowledge from practice.

- Estimation of Data-driven routing and scheduling with fully observed tour data
  The data-driven model we developed in chapter 3 for routing and scheduling decisions of carriers (chapter 3) is specialized for the cases that the tour data is partially observable. In the cases where tours are fully observable, the method has to be adapted. This requires defining a new similarity measurement that can calculate the differences

between the sequence of trips generated by the model in a tour and observed trip sequences.

- Integration of the methods in simulation platforms
  We recommend the integration of our data-driven routing and scheduling model (proposed in chapter 3) with agent-based traffic or freight simulation platforms like MATSIM and MASS-GT respectively. These two platforms have their tour planning module which either is normative (like in MATSIM) or with a simplified tour formation module that cannot completely capture spatial-temporal aspects of tours (like in MASS-GT). Integrating these systems together with our tour model can sum up the advantage of agent-based freight and traffic simulation with more realistic freight operations.

- Incorporating artificial intelligence in learning and generating freight tour patterns
  One may also explore the possibilities for more advanced methods like AI to learn more about the sequence of trips in a tour. In recent years, learning in the context of combinatorial problems, like tour planning with AI, has gained some attention. It can be an extension of our efforts, if large data sets of planned tours will become available, with full information about the sequence of trips.

- OD estimation and model validation
  The tour model developed in this thesis can be used in disaggregate freight transport simulation experiments to help policymakers get a more realistic grip on spatial and temporal characteristics of freight tours. More specifically, this model can be used for dynamic truck OD estimation with a tangible connection to freight activities. The output of the data-driven routing and scheduling model proposed in this study is a set of tours with all the pick-ups and deliveries. These tours can be transferred into time-dependent OD matrixes which can be linked to a traffic simulation. This integration allows predicting truck counts on a specific location on road networks. comparing the predicted and observed truck counts can help to validate the tour model.

- Investigating tour chaining and transport channels
  In this thesis, we have explored several dimensions of freight transport and traffic systems with respect to the pre-trip decisions at the tactical level. A further study could investigate more strategic decisions like shipper-consumer interactions for the long-term impact of logistics and freight transport on the traffic system. Decisions like tour chaining and transportation channels could shed more light on the flow of commercial vehicles which can also be linked to the production plan and inventory of products.

- Interpretability of demand-responsive data-driven traffic model
  The data-driven traffic model developed in chapters 4 and 5 includes considerations from theory-driven traffic flows models. This helped us to develop a model with a level of interpretability particularly in learning fundamental diagrams. Since our traffic model is demand-responsive, it is also important to ensure the expandability of the model with respect to the demand. In chapter 4, we have taken very initial steps toward this issue using the perturbation feature importance approach. However, more advanced techniques like shapely values and visualization of decision gates in the developed graph neural networks can increase the transparency of the model.

7.3.2 Recommendations for practice

We have explored two applications of our method in chapters 5 and 6, showing its ability to be used in departure time shift policy assessment and decision support systems. However, there are three promising directions for the practical applications of the methods developed in this thesis.

- One major example is a contribution to the digital twins of the port-hinterland ecosystem or city logistics. Such systems help decision-makers to monitor the day-to-day dynamics of freight operations and traffic states. Integrating methods developed in this dissertation into the digital twins can give realistic predictions for time to action for policies and management.

- Although our methods can calculate and trade-off the costs and benefits of multiple stakeholders. Our practical experiments show that efforts, costs and benefits are not balanced among stakeholders for the defined policies. We recommend for practice diving into the cost and benefit analysis to design a comprehensive gain-sharing mechanism between stakeholders that guarantee the successful application of time-shifted freight transport policy.

- Finally, We recommend to practitioners investigate how departure time shift policy and decision support system for time slot management can be implemented in practice. This requires efforts to qualitatively search for key actors and their role in the system and encourage them to see the benefit of the time-shifted transport operation and its significance.

In sum, with recent advances in data collection and the availability of disaggregated data in freight transport and logistics systems, both society and academia have realized the need for data-driven approaches that can convert data into knowledge. This thesis has taken important initial steps and paved the path for further research and developments in this field.

Appendix A

Accuracy and goodness-of-fit

In this appendix, we calculate the goodness of fit of the decision tree models developed in chapter 2. Assume that ρ is the probability of predicting the response variable for a given observation *x*.

$$\rho = \sum_{t} Pr(x \to l) \sum_{i} \Pr(i \mid l) \Pr'(i \mid l) \tag{A.1}$$

In this formulation, *Pr(x→l)* is the probability that the model use leaf node *l* to classify observation *x*, *Pr(i|l)* is the probability of the observed class *i*, and *Pr'(i|l)* is the probability of the predicted class *i* in leaf the node *l*. It is also possible to calculate the relative performance of the models by comparing the goodness-of-fit of a model with that of its root model. The root model for a decision tree is the tree with just its root node. This measurement is comparable to the maximum likelihood ratio and indicates the model performance improvement compared to the root model.

$$\rho_0 = \sum_{i} \Pr(i) \Pr'(i) \tag{A.2}$$

$$\rho_{incr} = \frac{\rho - \rho_0}{1 - \rho_0} \tag{A.3}$$

The most common metrics to evaluate the predictive performance of a classification tree are metrics that use precision and recall driven from the confusion matrix.

$$Accuracy(Acc) = \frac{TP + TN}{TP + TN + FP + FN} \tag{A.4}$$

Where TP is the number of true positives, TN is the number of true negatives, FP is the number of false-positive, and FN is the number of false negatives. However, this metric is not suitable if the class distribution is imbalanced. Therefore we also report on the one-vs-all balanced accuracy and the F1-score which measure the performance of multi-class classifiers with an imbalanced class distribution. We also present Cohen's kappa metric which is useful to assess if the performance of the model is better than the prediction by chance.

$$Sensitivity = \frac{TP}{TP + FN} \tag{A.5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{A.6}$$

$$Precision = \frac{TP}{TP + FP} \tag{A.7}$$

$$Balanaced - Accuracy(BAcc) = \frac{1}{2}(Sensitivity + Specificity) \tag{A.8}$$

$$F1 - score = \frac{2 \times Sensitivity \times Precision}{Sensitivity + Precision} \tag{A.9}$$

As for the regression capability of the model, we report on the conventional r^2.

Appendix B

Decision Tree quantifications

The time of day and the type of tour decision tree models are quantified using Equations (2.22), (2.23), and (2.24). We present the results of these quantifications for each model in Tables B-1 and B-2.

Table B-1: quantitative effect of covariates for time-of-day model

Market	Variables	MI	MI1	MI2	MI3	MI4	DI1	DI2	DI3	DI4
Agricultural	FNS_cat	0.254	0.242	0.312	0.333	0.018	-1	-1	-1	0.11
	LaterDestCong	0.173	0.011	0.156	0.166	0.328	-1	-1	1	1
	FirstDestCong	0.169	0.188	0.156	0.166	0.182	-1	1	1	1
	Vehtype	0.157	0.217	0.156	0.166	0.098	1	1	1	1
	DCVisited	0.103	0.176	0.048	0.062	0.215	1	1	1	1
	Day	0.073	0.132	0.059	0.051	0.095	0.01	0.09	0.02	-0.02
	TripDist_cat	0.071	0.034	0.115	0.055	0.065	-0.12	-0.15	0.34	0.19
Fresh Fruits and Vegetable	TTVisited	0.179	0.247	0.116	0.048	0.221	-1	-1	-1	-1
	FNS_cat	0.174	0.223	0.246	0.290	0.077	-1	-1	-1	-1
	FirstDestCong	0.140	0.085	0.123	0.065	0.192	-1	1	1	1
	Vehtype	0.139	0.246	0.123	0.065	0.119	1	1	1	1
	DCVisited	0.117	0.039	0.123	0.025	0.167	1	1	-1	1
	LaterDestCong	0.099	0.005	0.123	0.088	0.126	-1	-1	1	1
	TripDist_cat	0.097	0.056	0.101	0.392	0.048	0.18	-0.12	0.30	0.06
	Day	0.055	0.099	0.044	0.027	0.050	-0.02	0.11	0.10	-0.03
Flowers	TripDist_cat	0.285	0.252	0.173	0.267	0.406	0	-0.10	1	-0.08
	FNS_cat	0.260	0.042	0.275	0.573	0.217	-1	-1	-1	-1
	Vehtype	0.185	0.257	0.137	0.022	0.268	1	1	1	1
	LaterDestCong	0.167	0.326	0.315	0.010	0.033	1	-1	1	-1
	Day	0.102	0.122	0.101	0.130	0.076	-0.16	-0.01	0.09	0.05
Food industries	FNS_cat	0.242	0.144	0.276	0.423	0.168	1	-1	-1	-1
	TTVisited	0.216	0.265	0.138	0.046	0.294	-1	-1	-1	-1
	Vehtype	0.200	0.078	0.138	0.202	0.220	1	-1	1	1
	FirstDestCong	0.153	0.288	0.115	0.066	0.178	-1	1	1	1
	Day	0.074	0.107	0.072	0.068	0.072	0.05	0.67	-0.01	0.00
	tripDist_Cat	0.066	0.053	0.122	0.147	0.026	0.17	-0.73	-0.42	0.38
	DCVisited	0.050	0.066	0.138	0.047	0.042	1	-1	1	1
Agricultural-food	LaterDestCong	0.187	0.004	0.335	0.149	0.224	-1	1	-1	1
	TTVisited	0.185	0.216	0.062	0.105	0.243	-1	-1	-1	-1
	FNS_cat	0.158	0.070	0.200	0.386	0.012	-1	-1	-1	1
	Vehtype	0.121	0.181	0.100	0.166	0.086	1	-1	1	1
	DCVisited	0.096	0.048	0.002	0.020	0.159	1	-1	1	1
	FirstDestCong	0.093	0.082	0.100	0.051	0.121	-1	-1	1	1
	tripDist_Cat	0.080	0.289	0.101	0.067	0.064	-0.40	-0.10	0.26	0.35
	Day	0.080	0.111	0.100	0.055	0.092	0.28	-0.22	-0.10	0.04
Chemical and petroleum industries	FNS_cat	0.200	0.150	0.328	0.191	0.128	1	-1	0	-0.38
	TTVisited	0.199	0.175	0.159	0.034	0.285	-1	-1	1	-1
	DCVisited	0.129	0.222	0.154	0.075	0.033	1	1	1	1
	Empty_container	0.126	0.113	0.162	0.137	0.102	-1	-1	-1	-1
	Vehtype	0.105	0.025	0.010	0.210	0.244	1	1	1	1
	Day	0.078	0.105	0.044	0.104	0.081	0.07	-0.02	0.17	0.03
	tripDist_Cat	0.066	0.047	0.079	0.175	0.054	0.59	-0.24	0.32	-0.06
	LaterDestCong	0.051	0.091	0.059	0.009	0.016	1	-1	1	1
	FirstDestCong	0.046	0.073	0.005	0.065	0.057	1	1	1	1
Construction materials	FNS_cat	0.289	0.370	0.263	0.261	0.271	0.95	-1	0	-1
	tripDist_Cat	0.145	0.206	0.080	0.325	0.160	-0.55	-0.14	0.73	-0.50
	DCVisited	0.174	0.157	0.219	0.086	0.145	1	1	1	1
	TTVisited	0.278	0.210	0.219	0.194	0.390	-1	-1	1	-1
	FirstDestCong	0.114	0.057	0.219	0.135	0.033	-1	1	1	-1
Parcels & packages	DCVisited	0.030	0.110	0.048	0.024	0.000	1	1	1	1
	Vehtype	0.162	0.055	0.233	0.205	0.032	1	1	1	1
	Day	0.254	0.516	0.228	0.137	0.478	0	0.03	-0.07	0.09
	tripDist_Cat	0.123	0.078	0.130	0.142	0.084	-1	-0.50	0.48	-0.48
	LaterDestCong	0.244	0.055	0.254	0.248	0.301	-1	-1	1	1
	FirstDestCong	0.187	0.188	0.106	0.244	0.104	1	1	1	1
Miscellaneous goods	LaterDestCong	0.177	0.017	0.055	0.207	0.349	-1	-1	1	-1
	Vehtype	0.159	0.039	0.166	0.236	0.111	1	1	1	1
	tripDist_Cat	0.156	0.226	0.240	0.150	0.045	-0.75	-0.09	0.45	-0.34
	DCVisited	0.147	0.289	0.055	0.031	0.311	-1	1	1	1
	FNS_cat	0.134	0.242	0.272	0.044	0.110	-0.94	-0.92	0	-1
	Day	0.075	0.101	0.040	0.100	0.030	0.04	0.09	-0.059	0.026
	FirstDestCong	0.153	0.086	0.172	0.232	0.044	1	1	1	1

Table B-2: quantitative effect of covariates for type-for-tour model

Market	variables	MI	MI1	MI2	MI3	DI1	DI2	DI3
Agricultural	FNS_cat	0.254	0.356	0.208	0.251	-1	-1	0.13
	LaterDestCong	0.195	0.241	0.191	0.170	1	-1	1
	FirstDestCong	0.048	0.084	0.036	0.039	1	1	1
	WF_Cat	0.126	0.023	0.189	0.105	-0.60	-0.44	-0.75
	TTVisited	0.169	0.161	0.185	0.151	-1	-1	-1
	Vehicle Type	0.134	0.087	0.138	0.161	1	1	1
	TripDist_cat	0.074	0.048	0.052	0.123	-0.25	-0.28	0.14
Vegetable and Fresh Fruits	FNS_cat	0.294	0.380	0.228	0.385	-1	-1	0
	FirstDestCong	0.057	0.090	0.048	0.050	1	1	1
	Vehtype	0.143	0.129	0.153	0.130	1	1	1
	WF_Cat	0.145	0.072	0.197	0.077	-0.81	-0.48	-0.46
	LaterDestCong	0.190	0.190	0.189	0.192	1	-1	1
	TTVisited	0.171	0.140	0.184	0.167	-1	-1	-1
Flowers	FirstDestCong	0.082	0.160	0.081	0.061	1	1	1
	FNS_cat	0.471	0.529	0.423	0.565	-1	-1	0
	LaterDestCong	0.347	0.265	0.385	0.283	1	-1	1
	TripDist	0.100	0.046	0.111	0.090	-0.20	-0.52	0.19
Food industries	FNS_cat	0.266	0.341	0.190	0.323	-1	-1	1
	DCVisited	0.044	0.041	0.012	0.098	1	1	1
	LaterDestCong	0.187	0.171	0.211	0.162	1	-1	1
	FirstDestCong	0.050	0.039	0.060	0.045	1	1	1
	Day	0.055	0.081	0.041	0.058	-0.16	0.07	0.03
	tripDist_Cat	0.082	0.120	0.074	0.065	-0.09	-0.64	0.20
	WF_Cat	0.155	0.124	0.208	0.097	-1.00	-0.52	-0.51
	TTVisited	0.160	0.083	0.205	0.152	1	-1	-1
Agricultural-food	LaterDestCong	0.168	0.154	0.170	0.177	1	-1	1
	WF_Cat	0.115	0.180	0.133	0.073	-0.94	0.33	-0.12
	FNS_cat	0.354	0.381	0.339	0.338	-1	0	0.058
	Vehtype	0.090	0.045	0.029	0.121	1	1	1
	DCVisited	0.086	0.027	0.001	0.126	1	-1	1
	tripDist_Cat	0.069	0.055	0.197	0.072	0.52	0.22	-0.71
	Empty_Pallets	0.118	0.157	0.132	0.093	-1	-1	-1
Chemical and petroleum industries	FNS_cat	0.187	0.345	0.091	0.243	-1	-0.89	1.00
	WF_Cat	0.143	0.067	0.242	0.062	-0.31	-0.71	-0.25
	DCVisited	0.085	0.136	0.017	0.141	1	1	1
	Vehtype	0.052	0.032	0.063	0.045	1	1	1
	Empty_container	0.104	0.105	0.159	0.047	-1	-1	-1
	Day	0.074	0.073	0.043	0.106	0.06	-0.05	0.00
	tripDist_Cat	0.039	0.019	0.040	0.043	0.16	0.14	0.06
	TTVisited	0.135	0.148	0.158	0.109	-1	-1	-1
	LaterDestCong	0.153	0.054	0.172	0.159	1	-1	1
	FirstDestCong	0.029	0.021	0.015	0.045	1	1	1
Construction materials	FNS_cat	0.207	0.365	0.133	0.420	-1.00	-0.82	1
	tripDist_Cat	0.124	0.074	0.155	0.029	0.06	-0.68	0.83
	DCVisited	0.159	0.051	0.170	0.170	1	1	1
	Vehicle	0.230	0.182	0.251	0.170	1	1	1
	LaterDestCong	0.252	0.182	0.289	0.142	1	-1	1
	FirstDestCong	0.027	0.145	0.001	0.068	1	-1	1
Parcels & packages	Vehicle	0.107	0.082	0.153	0.047	1	1	1
	Empty_Pallets	0.107	0.082	0.153	0.047	1	1	1
	Day	0.048	0.047	0.035	0.134	0.04	0.04	-0.18
	tripDist_Cat	0.086	0.090	0.084	0.068	0.09	-0.58	-0.13
	LaterDestCong	0.141	0.166	0.109	0.134	1	-1	1
	FNS_cat	0.268	0.344	0.164	0.238	-1	-1	0.03
	WF_cat	0.142	0.074	0.220	0.242	0.01	-0.72	-1
	FirstDestCong	0.100	0.115	0.082	0.089	1	1	1
Miscellaneous goods	LaterDestCong	0.188	0.166	0.204	0.204	1	-1	1
	tripDist_Cat	0.087	0.090	0.071	0.201	0.04	-0.82	-0.10
	FNS_cat	0.258	0.331	0.212	0.163	-1	-1	0.44
	Empty_Pallets	0.188	0.166	0.204	0.204	-1	-1	0.03
	WF_cat	0.174	0.086	0.252	0.119	-0.12	-0.79	-0.63
	FirstDestCong	0.104	0.161	0.057	0.110	1	1	1

Bibliography

ABDELMAGID, A. M., GHEITH, M. S. & ELTAWIL, A. B. 2021. A comprehensive review of the truck appointment scheduling models and directions for future research. *Transport Reviews***,** 1-25.

AGRAWAL, R. & SRIKANT, R. Fast algorithms for mining association rules. Proc. 20th int. conf. very large databases, VLDB, 1994. 487-499.

AL-DEEK, H. M. 2002. Use of vessel freight data to forecast heavy truck movements at seaports. *Transportation research record,* 1804**,** 217-224.

AL-DEEK, H. M., JOHNSON, G., MOHAMED, A. & EL-MAGHRABY, A. 2000. Truck trip generation models for seaports with container and trailer operation. *Transportation Research Record,* 1719**,** 1-9.

ALHO, A. R., SAKAI, T., CHUA, M. H., JEONG, K., JING, P. & BEN-AKIVA, M. 2019. Exploring Algorithms for Revealing Freight Vehicle Tours, Tour-Types, and Tour-Chain-Types from GPS Vehicle Traces and Stop Activity Data. *Journal of Big Data Analytics in Transportation,* 1**,** 175-190.

ARENTZE, T. & TIMMERMANS, H. 2003. Measuring Impacts of Condition Variables in Rule-Based Models of Space-Time Choice Behavior: Method and Empirical Illustration. *Geographical Analysis,* 35**,** 24-45.

ARIAN, A., ERMAGUN, A., ZHU, X. & CHIU, Y.-C. 2018. An Empirical Investigation of the Reward Incentive and Trip Purposes on Departure Time Behavior Change. *Advances in Transport Policy and Planning.* Elsevier.

BANISTER, D. & STEAD, D. 2004. Impact of information and communications technology on transport. *Transport Reviews,* 24**,** 611-632.

BEN-AKIVA, M. E., TOLEDO, T., SANTOS, J., COX, N., ZHAO, F., LEE, Y. J. & MARZANO, V. 2016. Freight data collection using GPS and web-based surveys: Insights from US truck drivers' survey and perspectives for urban freight. *Case studies on transport policy,* 4**,** 38-44.

BEN-ELIA, E. & ETTEMA, D. 2011. Changing commuters' behavior using rewards: A study of rush-hour avoidance. *Transportation research part F: traffic psychology and behaviour,* 14**,** 354-368.

BLIEK, L., VERSTRAETE, H. R., VERHAEGEN, M. & WAHLS, S. 2016. Online optimization with costly and noisy measurements using random Fourier expansions. *IEEE transactions on neural networks and learning systems,* 29**,** 167-182.

BLIEMER, M. C. & VAN AMELSFORT, D. H. 2010. Rewarding instead of charging road users: a model case study investigating effects on traffic conditions. *European Transport\Trasporti Europei***,** 23-40.

BORCHANI, H., VARANDO, G., BIELZA, C. & LARRAÑAGA, P. 2015. A survey on multi-output regression. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* 5**,** 216-233.

BREIMAN, L., FRIEDMAN, J., STONE, C. J. & OLSHEN, R. A. 1984. *Classification and regression trees,* CRC press.

BURNS, L. D., HALL, R. W., BLUMENFELD, D. E. & DAGANZO, C. F. 1985. Distribution strategies that minimize transportation and inventory costs. *Operations research,* 33**,** 469-490.

CALVERT, S. C., SNELDER, M., BAKRI, T., HEIJLIGERS, B. & KNOOP, V. L. 2015. Real-time travel time prediction framework for departure time and route advice. *Transportation Research Record,* 2490**,** 56-64.

CASTELLUCCI, P. B., DARVISH, M. & COELHO, L. C. 2021. *A Benders Decomposition Algorithm for the Time-dependent Vehicle Routing Problem,* Bureau de Montreal, Université de Montreal.

CHAWLA, N. V., BOWYER, K. W., HALL, L. O. & KEGELMEYER, W. P. 2002. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research,* 16**,** 321-357.

CHEN, G. & YANG, Z. 2010. Optimizing time windows for managing export container arrivals at Chinese container terminals. *Maritime Economics & Logistics,* 12**,** 111-126.

CHEN, X., ZHOU, X. & LIST, G. F. 2011. Using time-varying tolls to optimize truck arrivals at ports. *Transportation Research Part E: Logistics and Transportation Review,* 47**,** 965-982.

CHRISTIDIS, P. & RIVAS, J. N. I. 2012. Measuring road congestion. *Institute for Prospective and Technological Studies, Joint Research Centre, Brussels*.

COSMETATOS, G. P. 1976. Some approximate equilibrium results for the multi-server queue (M/G/r). *Journal of the Operational Research Society,* 27**,** 615-620.

CUI, Z., KE, R., PU, Z., MA, X. & WANG, Y. 2020a. Learning traffic as a graph: A gated graph wavelet recurrent neural network for network-scale traffic prediction. *Transportation Research Part C: Emerging Technologies,* 115**,** 102620.

CUI, Z., KE, R., PU, Z. & WANG, Y. 2020b. Stacked Bidirectional and Unidirectional LSTM Recurrent Neural Network for Forecasting Network-wide Traffic State with Missing Values. *Transportation Research Part C: Emerging Technologies,* 118.

DE BOER, C., SNELDER, M., VAN NES, R. & VAN AREM, B. 2017. The impact of route guidance, departure time advice and alternative routes on door-to-door travel time reliability: Two data-driven assessment methods. *Journal of Intelligent Transportation Systems,* 21**,** 465-477.

DE BOK, M. & TAVASSZY, L. 2018. An empirical agent-based simulation system for urban goods transport (MASS-GT). *Procedia computer science,* 130**,** 126-133.

DE BOK, M., TAVASSZY, L. & THOEN, S. 2020. Application of an empirical multi-agent model for urban goods transport to analyze impacts of zero emission zones in The Netherlands. *Transport Policy*.

DE JONG, G., KOUWENHOVEN, M., BATES, J., KOSTER, P., VERHOEF, E., TAVASSZY, L. & WARFFEMIUS, P. 2014. New SP-values of time and reliability for freight transport in the Netherlands. *Transportation Research Part E: Logistics and Transportation Review,* 64**,** 71-87.

DE JONG, G., KOUWENHOVEN, M., RUIJS, K., VAN HOUWE, P. & BORREMANS, D. 2016. A time-period choice model for road freight transport in Flanders based on stated preference data. *Transportation Research Part E: Logistics and Transportation Review,* 86**,** 20-31.

DEB, K., AGRAWAL, S., PRATAP, A. & MEYARIVAN, T. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. International conference on parallel problem solving from nature, 2000. Springer, 849-858.

DEB, K., PRATAP, A., AGARWAL, S. & MEYARIVAN, T. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation,* 6**,** 182-197.

DHINGRA, S., MUJUMDAR, P. & GAJJAR, R. H. 1993. Application of time series techniques for forecasting truck traffic attracted by the Bombay metropolitan region. *Journal of advanced transportation,* 27**,** 227-249.

DI FEBBRARO, A., SACCO, N. & SAEEDNIA, M. 2016. An agent-based framework for cooperative planning of intermodal freight transport chains. *Transportation Research Part C: Emerging Technologies,* 64**,** 72-85.

DO, L. N., VU, H. L., VO, B. Q., LIU, Z. & PHUNG, D. 2019. An effective spatial-temporal attention based neural network for traffic flow prediction. *Transportation Research Part C: Emerging Technologies,* 108**,** 12-28.

DO, N. A. D., NIELSEN, I. E., CHEN, G. & NIELSEN, P. 2016. A simulation-based genetic algorithm approach for reducing emissions from import container pick-up operation at container terminal. *Annals of Operations Research,* 242**,** 285-301.

DONNELLY 2009. A hybrid microsimulation model of urban freight transport demand. *PhD dissertation,* University of Melbourn.

DU, S., LI, T. & HORNG, S.-J. Time series forecasting using sequence-to-sequence deep learning framework. 2018 9th international symposium on parallel architectures, algorithms and programming (PAAP), 2018. IEEE, 171-176.

DU, S., LI, T., YANG, Y. & HORNG, S.-J. 2020. Multivariate time series forecasting via attention-based encoder–decoder framework. *Neurocomputing,* 388**,** 269-279.

ELIASSON, J. 2008. Lessons from the Stockholm congestion charging trial. *Transport Policy,* 15**,** 395-404.

ETTEMA, D., KNOCKAERT, J. & VERHOEF, E. 2010. Using incentives as traffic management tool: empirical results of the" peak avoidance" experiment. *Transportation Letters,* 2**,** 39-51.

FAN, H., REN, X., GUO, Z. & LI, Y. 2019. Truck scheduling problem considering carbon emissions under truck appointment system. *Sustainability,* 11**,** 6256.

FIGLIOZZI, M. A. 2010. The impacts of congestion on commercial vehicle tour characteristics and costs. *Transportation research part E: logistics and transportation review,* 46**,** 496-506.

FU, R., ZHANG, Z. & LI, L. Using LSTM and GRU neural network methods for traffic flow prediction. 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), 2016. IEEE, 324-328.

FURTADO, M. G. S., MUNARI, P. & MORABITO, R. 2017. Pickup and delivery problem with time windows: a new compact two-index formulation. *Operations Research Letters,* 45**,** 334-341.

GARRIDO, R. A. & MAHMASSANI, H. S. 2000. Forecasting freight transportation demand with the space–time multinomial probit model. *Transportation Research Part B: Methodological,* 34**,** 403-418.

GIANNOPOULOS, G. A. 2004. The application of information and communication technologies in transport. *European journal of operational research,* 152**,** 302-320.

GIULIANO, G. & O'BRIEN, T. 2007. Reducing port-related truck emissions: The terminal gate appointment system at the Ports of Los Angeles and Long Beach. *Transportation Research Part D: Transport and Environment,* 12**,** 460-473.

GONZALEZ-CALDERON, C. A. & HOLGUÍN-VERAS, J. 2017. Entropy-based freight tour synthesis and the role of traffic count sampling. *Transportation Research Part E: Logistics and Transportation Review*.

GONZALEZ-CALDERON, C. A. & HOLGUÍN-VERAS, J. 2019. Entropy-based freight tour synthesis and the role of traffic count sampling. *Transportation Research Part E: Logistics and Transportation Review,* 121**,** 63-83.

GUAN, C. & LIU, R. R. 2009a. Container terminal gate appointment system optimization. *Maritime Economics & Logistics,* 11**,** 378-398.

GUAN, C. Q. & LIU, R. 2009b. Modeling gate congestion of marine container terminals, truck waiting cost, and optimization. *Transportation Research Record,* 2100**,** 58-67.

HAGAN, M. T. & MENHAJ, M. B. 1994. Training feedforward networks with the Marquardt algorithm. *IEEE transactions on Neural Networks,* 5**,** 989-993.

HAHSLER, M., BUCHTA, C., GRUEN, B., HORNIK, K., JOHNSON, I., BORGELT, C. & HAHSLER, M. M. 2021. Package 'arules'.

HAHSLER, M. & CHELLUBOINA, S. 2011. Visualizing association rules: Introduction to the R-extension package arulesViz. *R project module,* 223-238.

HASSOUN, M. H. 1995. *Fundamentals of artificial neural networks,* MIT press.

HEILIG, L. & VOß, S. 2017. Information systems in seaports: a categorization and overview. *Information Technology and Management,* 18**,** 179-201.

HEINITZ, F. M. & LIEDTKE, G. T. 2010. Principles of Constraint–Consistent Activity-Based Transport Modeling. *Computer-Aided Civil and Infrastructure Engineering,* 25**,** 101-115.

HOCHREITER, S. & SCHMIDHUBER, J. 1997. Long short-term memory. *Neural computation,* 9**,** 1735-1780.

HOFSTEDE, G. J., JONKER, C. M., VERWAART, T. & YORKE-SMITH, N. 2019. The lemon car game across cultures: Evidence of relational rationality. *Group Decision and Negotiation,* 28**,** 849-877.

HOLGUÍN-VERAS, J. 2008. Necessary conditions for off-hour deliveries and the effectiveness of urban freight road pricing and alternative financial policies in competitive markets. *Transportation Research Part A: Policy and Practice,* 42**,** 392-413.

HOLGUÍN-VERAS, J. & AROS-VERA, F. 2015. Self-supported freight demand management: pricing and incentives. *euro Journal on transportation and Logistics,* 4**,** 237-260.

HOLGUÍN-VERAS, J., JALLER, M., SÁNCHEZ-DÍAZ, I., CAMPBELL, S. & LAWSON, C. T. 2014. Freight generation and freight trip generation models. *Modelling freight transport.* Elsevier.

HOLGUIN-VERAS, J. & PATIL, G. R. 2007. Integrated origin–destination synthesis model for freight with commodity-based and empty trip models. *Transportation Research Record,* 2008**,** 60-66.

HOLGUÍN-VERAS, J. & PATIL, G. R. 2005. Observed trip chain behavior of commercial vehicles. *Transportation Research Record,* 1906**,** 74-80.

HOLGUÍN-VERAS, J. & PATIL, G. R. 2008. A multicommodity integrated freight origin–destination synthesis model. *Networks and Spatial Economics,* 8**,** 309-326.

HOLGUÍN-VERAS, J., SILAS, M., POLIMENI, J. & CRUZ, B. 2008. An investigation on the effectiveness of joint receiver–carrier policies to increase truck traffic in the off-peak hours. *Networks and Spatial Economics,* 8**,** 327-354.

HOLGUÍN-VERAS, J., WANG, Q., XU, N., OZBAY, K., CETIN, M. & POLIMENI, J. 2006. The impacts of time of day pricing on the behavior of freight carriers in a congested urban area: Implications to road pricing. *Transportation Research Part A: Policy and Practice,* 40**,** 744-766.

HONG, W.-C. 2011. Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm. *Neurocomputing,* 74**,** 2096-2107.

HORNIK, K., GRÜN, B. & HAHSLER, M. 2005. arules-A computational environment for mining association rules and frequent item sets. *Journal of statistical software,* 14**,** 1-25.

HUANG, B., BUCKLEY, B. & KECHADI, T.-M. 2010. Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications. *Expert Systems with Applications,* 37**,** 3638-3646.

HUNT, J. D. & STEFAN, K. 2007. Tour-based microsimulation of urban commercial movements. *Transportation Research Part B: Methodological,* 41**,** 981-1013.

HURVICH, C. M. & TSAI, C. L. 1993. A corrected Akaike information criterion for vector autoregressive model selection. *Journal of time series analysis,* 14**,** 271-279.

HUTTER, F., HOOS, H. H. & LEYTON-BROWN, K. Sequential model-based optimization for general algorithm configuration. International conference on learning and intelligent optimization, 2011. Springer, 507-523.

HUYNH, N. 2009. Reducing truck turn times at marine terminals with appointment scheduling. *Transportation research record,* 2100**,** 47-57.

HUYNH, N., SMITH, D. & HARDER, F. 2016. Truck appointment systems: where we are and where to go from here. *Transportation Research Record,* 2548**,** 1-9.

IM, H., YU, J. & LEE, C. 2021. Truck appointment system for cooperation between the transport companies and the terminal operator at container terminals. *Applied Sciences,* 11**,** 168.

IRANNEZHAD, E., PRATO, C. & HICKMAN, M. 2019. A joint hybrid model of the choices of container terminals and of dwell time. *Transportation Research Part E: Logistics and Transportation Review,* 121**,** 119-133.

KHAN, M. & MACHEMEHL, R. 2017a. Analyzing tour chaining patterns of urban commercial vehicles. *Transportation research part A: policy and practice,* 102**,** 84-97.

KHAN, M. & MACHEMEHL, R. 2017b. Commercial vehicles time of day choice behavior in urban areas. *Transportation Research Part A: Policy and Practice,* 102**,** 68-83.

KIM, I. Y. & DE WECK, O. 2006. Adaptive weighted sum method for multiobjective optimization: a new method for Pareto front generation. *Structural and multidisciplinary optimization,* 31**,** 105-116.

KIM, S., RASOULI, S., TIMMERMANS, H. & YANG, D. 2018. Estimating panel effects in probabilistic representations of dynamic decision trees using bayesian generalized linear mixture models. *Transportation Research Part B: Methodological,* 111**,** 168-184.

KIPF, T. N. & WELLING, M. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

KIRSCHNER, J., MUTNY, M., HILLER, N., ISCHEBECK, R. & KRAUSE, A. Adaptive and safe Bayesian optimization in high dimensions via one-dimensional subspaces. International Conference on Machine Learning, 2019. PMLR, 3429-3438.

KLEFF, A., BRÄUER, C., SCHULZ, F., BUCHHOLD, V., BAUM, M. & WAGNER, D. Time-Dependent Route Planning for Truck Drivers. International Conference on Computational Logistics, 2017. Springer, 110-126.

KLODZINSKI, J. & AL-DEEK, H. M. 2004. Methodology for modeling a road network with high truck volumes generated by vessel freight activity from an intermodal facility. *Transportation research record,* 1873**,** 35-44.

KNOCKAERT, J., TSENG, Y.-Y., VERHOEF, E. T. & ROUWENDAL, J. 2012. The Spitsmijden experiment: A reward to battle congestion. *Transport Policy,* 24**,** 260-272.

KOUROUNIOTI, I. & POLYDOROPOULOU, A. 2018. Application of aggregate container terminal data for the development of time-of-day models predicting truck arrivals. *European Journal of Transport and Infrastructure Research,* 18.

KOUROUNIOTI, I., POLYDOROPOULOU, A. & KANAAN, A. 2021. Predicting container terminal daily workload: a Middle East port case study. *Freight Transport Modeling in Emerging Countries.* Elsevier.

KUMAR, S. V. 2017. Traffic flow prediction using Kalman filtering technique. *Procedia Engineering,* 187**,** 582-587.

KUMAR, S. V. & VANAJAKSHI, L. 2015. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review,* 7**,** 21.

LANA, I., DEL SER, J., VELEZ, M. & VLAHOGIANNI, E. I. 2018. Road traffic forecasting: recent advances and new challenges. *IEEE Intelligent Transportation Systems Magazine,* 10**,** 93-109.

LANGE, A.-K., KREUZ, F., LANGKAU, S., JAHN, C. & CLAUSEN, U. Defining the quota of truck appointment systems. Hamburg International Conference of Logistics (HICL) 2020, 2020. epubli, 211-246.

LANGEVIN, A. & RIOPEL, D. 2005. *Logistics systems: design and optimization*, Springer Science & Business Media.

LI, B., TAN, K. W. & TRAN, K. T. Traffic simulation model for port planning and congestion prevention. Proceedings of the 2016 Winter Simulation Conference, 2016. IEEE Press, 2382-2393.

LI, G., KNOOP, V. L. & VAN LINT, H. 2021. Multistep traffic forecasting by dynamic graph convolution: Interpretations of real-time spatial correlations. *Transportation Research Part C: Emerging Technologies,* 128**,** 103185.

LI, N., CHEN, G., NG, M., TALLEY, W. K. & JIN, Z. 2020. Optimized appointment scheduling for export container deliveries at marine terminals. *Maritime Policy & Management,* 47**,** 456-478.

LIEDTKE, G. 2009. Principles of micro-behavior commodity transport modeling. *Transportation Research Part E: Logistics and Transportation Review,* 45**,** 795-809.

LIEDTKE, G. & SCHEPPERLE, H. 2004. Segmentation of the transportation market with regard to activity-based freight transport modelling. *International Journal of Logistics Research and Applications,* 7**,** 199-218.

LOH, W. Y. 2014. Fifty years of classification and regression trees. *International Statistical Review,* 82**,** 329-348.

LV, Y., DUAN, Y., KANG, W., LI, Z. & WANG, F.-Y. 2014. Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems,* 16**,** 865-873.

M. ABDELMAGID, A., S. GHEITH, M. & B. ELTAWIL, A. A Binary Integer Programming Formulation and Solution for Truck Appointment Scheduling and Reducing Truck Turnaround Time in Container Terminals. 2021 The 8th International Conference on Industrial Engineering and Applications (Europe), 2021. 126-131.

MA, Y., VAN ZUYLEN, H. J., CHEN, Y. & VAN DALEN, J. Modeling and analyzing departure time slots allocation to optimize dynamic network capacity—the case of A15-motorway to rotterdam port. 5th Advanced Forum on Transportation of China (AFTC 2009), 2009. IET, 214-222.

MAHMASSANI, H. S. & JAYAKRISHNAN, R. 1991. System performance and user response under real-time information in a congested traffic corridor. *Transportation Research Part A: General,* 25**,** 293-307.

MAR-ORTIZ, J., CASTILLO-GARCÍA, N. & GRACIA, M. D. 2020. A decision support system for a capacity management problem at a container terminal. *International Journal of Production Economics,* 222**,** 107502.

MOMMENS, K., LEBEAU, P., VERLINDE, S., VAN LIER, T. & MACHARIS, C. 2018. Evaluating the impact of off-hour deliveries: An application of the TRansport Agent-BAsed model. *Transportation Research Part D: Transport and Environment,* 62**,** 102-111.

MONIRUZZAMAN, M., MAOH, H. & ANDERSON, W. 2016. Short-term prediction of border crossing time and traffic volume for commercial trucks: A case study for the Ambassador Bridge. *Transportation Research Part C: Emerging Technologies,* 63**,** 182-194.

NADI, A., SHARMA, S., SNELDER, M., BAKRI, T., VAN LINT, H. & TAVASSZY, L. 2021. Short-term prediction of outbound truck traffic from the exchange of information in logistics hubs: A case study for the port of Rotterdam. *Transportation Research Part C: Emerging Technologies,* 127**,** 103111.

NADI NAJAFABADI, A., SHARMA, S., SNELDER, M., TAVASSZY, L. & VAN LINT, J. Analyzing the role of seaport operations in generating inbound/outbound truck traffic demand and its implications on traffic system. TRAIL PhD Congress 2019, 2019.

NICHOLSON, H. & SWANN, C. 1974. The prediction of traffic flow volumes based on spectral analysis. *Transportation Research,* 8**,** 533-538.

NOTTEBOOM, T. 2009. The relationship between seaports and the intermodal hinterland in light of global supply chains: European challenges.

NUZZOLO, A. & COMI, A. 2014. Urban freight demand forecasting: a mixed quantity/delivery/vehicle-based model. *Transportation Research Part E: Logistics and Transportation Review,* 65**,** 84-98.

PHAN, M.-H. & KIM, K. H. 2015. Negotiating truck arrival times among trucking companies and a container terminal. *Transportation Research Part E: Logistics and Transportation Review,* 75**,** 132-144.

POLSON, N. G. & SOKOLOV, V. O. 2017. Deep learning for short-term traffic flow prediction. *Transportation Research Part C: Emerging Technologies,* 79**,** 1-17.

POONIA, P., JAIN, V. & KUMAR, A. 2018. Short Term Traffic Flow Prediction Methodologies: A Review. *Mody University International Journal of Computing and Engineering Research,* 2**,** 37-39.

POPE, J. A., RAKES, T. R., REES, L. P. & CROUCH, I. W. 1995. A network simulation of high-congestion road-traffic flows in cities with marine container terminals. *Journal of the Operational Research Society,* 46**,** 1090-1101.

PORT-OF-ROTTERDAM 2019. Continuing container growth pushes throughput at port of Rotterdam to a new high.

QUINLAN, J. R. 2014. *C4. 5: programs for machine learning*, Elsevier.

RAEDER, T. & CHAWLA, N. V. 2011. Market basket analysis with networks. *Social network analysis and mining,* 1**,** 97-113.

RAJABZADEH, Y., REZAIE, A. H. & AMINDAVAR, H. 2017. Short-term traffic flow prediction using time-varying Vasicek model. *Transportation Research Part C: Emerging Technologies,* 74**,** 168-181.

RUAN, M., LIN, J. J. & KAWAMURA, K. 2012. Modeling urban commercial vehicle daily tour chaining. *Transportation Research Part E: Logistics and Transportation Review,* 48**,** 1169-1184.

SAKAI, T., ALHO, A. R., BHAVATHRATHAN, B., DALLA CHIARA, G., GOPALAKRISHNAN, R., JING, P., HYODO, T., CHEAH, L. & BEN-AKIVA, M. 2020. SimMobility Freight: An agent-based urban freight simulator for evaluating

logistics solutions. *Transportation Research Part E: Logistics and Transportation Review,* 141**,** 102017.

SÁNCHEZ-DÍAZ, I., GEORÉN, P. & BROLINSON, M. 2017. Shifting urban freight deliveries to the off-peak hours: a review of theory and practice. *Transport reviews,* 37**,** 521-543.

SÁNCHEZ-DÍAZ, I., HOLGUÍN-VERAS, J. & BAN, X. J. 2015. A time-dependent freight tour synthesis model. *Transportation Research Part B: Methodological,* 78**,** 144-168.

SARVAREDDY, P., AL-DEEK, H., KLODZINSKI, J. & ANAGNOSTOPOULOS, G. 2005. Evaluation of two modeling methods for generating heavy-truck trips at an intermodal facility by using vessel freight data. *Transportation research record,* 1906**,** 113-120.

SCHREITER, T., VAN LINT, H., TREIBER, M. & HOOGENDOORN, S. Two fast implementations of the adaptive smoothing method used in highway traffic state estimation. 13th International IEEE Conference on Intelligent Transportation Systems, 2010. IEEE, 1202-1208.

SCHRÖDER, S. & LIEDTKE, G. T. 2017. Towards an integrated multi-agent urban transport model of passenger and freight. *Research in Transportation Economics,* 64**,** 3-12.

SHARIF, O., HUYNH, N. & VIDAL, J. M. 2011. Application of El Farol model for managing marine terminal gate congestion. *Research in Transportation Economics,* 32**,** 81-89.

SIRIPIROTE, T., SUMALEE, A. & HO, H. 2020. Statistical estimation of freight activity analytics from Global Positioning System data of trucks. *Transportation Research Part E: Logistics and Transportation Review,* 140**,** 101986.

TAVASSZY, L. & DE JONG, G. 2013. *Modelling freight transport*, Elsevier.

TAVASSZY, L. & REIS, V. 2021. Appraisal of freight projects and policies. *Advances in Transport Policy and Planning.* Elsevier BV.

TAVASSZY, L. A. 2008. Freight modelling: an overview of international experiences.

TAVASSZY, L. A., SMEENK, B. & RUIJGROK, C. J. 1998. A DSS for modelling logistic chains in freight transport policy analysis. *International Transactions in Operational Research,* 5**,** 447-459.

THOEN, S., TAVASSZY, L., DE BOK, M., CORREIA, G. & VAN DUIN, R. 2020. Descriptive modeling of freight tour formation: A shipment-based approach. *Transportation Research Part E: Logistics and Transportation Review,* 140**,** 101989.

THORHAUGE, M., CHERCHI, E. & RICH, J. 2016a. How flexible is flexible? Accounting for the effect of rescheduling possibilities in choice of departure time for work trips. *Transportation Research Part A: Policy and Practice,* 86**,** 177-193.

THORHAUGE, M., HAUSTEIN, S. & CHERCHI, E. 2016b. Accounting for the Theory of Planned Behaviour in departure time choice. *Transportation Research Part F: Traffic Psychology and Behaviour,* 38**,** 94-105.

TIAN, Y. & PAN, L. Predicting short-term traffic flow by long short-term memory recurrent neural network. 2015 IEEE international conference on smart city/SocialCom/SustainCom (SmartCity), 2015. IEEE, 153-158.

TLN. 2019. *Economische Wegwijzer 2019 nader verklaard* [Online]. Available: https://www.transport-online.nl/site/verkeers-nieuws/uploads/attachment-1100-Economische_Wegwijzer_Nader_Verklaard.pdf [Accessed].

TLN & EVOFENEDEX. 2020. *Blijf ondanks coronacrisis investeren in infrastructuur In: Economische Wegwijzer* [Online]. Available: https://www.tln.nl/app/uploads/2020/11/TLN_EcoWegwijzer_2020_A4_2P_DEF_RGB_HR.pdf [Accessed 20-11-2021].

TORKJAZI, M., HUYNH, N. & SHIRI, S. 2018. Truck appointment systems considering impact to drayage truck tours. *Transportation Research Part E: Logistics and Transportation Review,* 116**,** 208-228.

TREIBER, M. & HELBING, D. 2003. An adaptive smoothing method for traffic state identification from incomplete information. *Interface and Transport Dynamics.* Springer.

UKKUSURI, S. V., OZBAY, K., YUSHIMITO, W. F., IYER, S., MORGUL, E. F. & HOLGUÍN-VERAS, J. 2016. Assessing the impact of urban off-hour delivery program using city scale simulation models. *EURO Journal on Transportation and Logistics,* 5**,** 205-230.

VAN ASPEREN, E., BORGMAN, B. & DEKKER, R. 2013. Evaluating impact of truck announcements on container stacking efficiency. *Flexible Services and Manufacturing Journal,* 25**,** 543-556.

VAN DER MEULEN, S., GRIJSPAARDT, T., MARS, W., VAN DER GEEST, W., ROEST-CROLLIUS, A. & KIE, J. 2020. *Cost Figures for Freight Transport – final report* [Online]. Available: https://www.kimnet.nl/publicaties/formulieren/2020/05/26/cost-figures-for-freight-transport [Accessed].

VAN HINSBERGEN, C., VAN LINT, J. & SANDERS, F. Short term traffic prediction models. PROCEEDINGS OF THE 14TH WORLD CONGRESS ON INTELLIGENT TRANSPORT SYSTEMS (ITS), HELD BEIJING, OCTOBER 2007, 2007.

VAN LINT, J. 2010. Empirical evaluation of new robust travel time estimation algorithms. *Transportation Research Record,* 2160**,** 50-59.

VAN LINT, J. & VAN HINSBERGEN, C. 2012. Short-term traffic and travel time prediction models. *Artificial Intelligence Applications to Critical Transportation Issues,* 22**,** 22-41.

VEGELIEN, A. G. & DUGUNDJI, E. R. A Revealed Preference Time of Day Model for Departure Time of Delivery Trucks in the Netherlands. 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018. IEEE, 1770-1774.

VLAHOGIANNI, E. I., KARLAFTIS, M. G. & GOLIAS, J. C. 2007. Spatio-temporal short-term urban traffic volume forecasting using genetically optimized modular networks. *Computer-Aided Civil and Infrastructure Engineering,* 22**,** 317-325.

VLAHOGIANNI, E. I., KARLAFTIS, M. G. & GOLIAS, J. C. 2014. Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies,* 43**,** 3-19.

WAN, Y., ZHANG, A. & LI, K. X. 2018. Port competition with accessibility and congestion: a theoretical framework and literature review on empirical studies. *Maritime Policy & Management,* 45**,** 239-259.

WATLING, D., CONNORS, R. & CHEN, H. Sensitivity analysis of optimal routes, departure times and speeds for fuel-efficient truck journeys. 2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2019. IEEE, 1-7.

WIBOWO, B. & FRANSOO, J. 2021. Joint-optimization of a truck appointment system to alleviate queuing problems in chemical plants. *International Journal of Production Research,* 59**,** 3935-3950.

WILLIAMS, B. M. & HOEL, L. A. 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. *Journal of transportation engineering,* 129**,** 664-672.

WISETJINDAWAT, W., YAMAMOTO, K. & MARCHAL, F. 2012. A commodity distribution model for a multi-agent freight system. *Procedia-Social and Behavioral Sciences,* 39**,** 534-542.

WU, Y., TAN, H., QIN, L., RAN, B. & JIANG, Z. 2018. A hybrid deep learning based traffic flow prediction method and its understanding. *Transportation Research Part C: Emerging Technologies,* 90**,** 166-180.

XIE, Y. & HUYNH, N. Kernel-Based Machine Learning Methods for Modeling Daily Truck Volume at Seaport Terminals. 51st Annual Transportation Research Forum, Arlington, Virginia, March 11-13, 2010, 2010. Transportation Research Forum.

XU, B., LIU, X., YANG, Y., LI, J. & POSTOLACHE, O. 2021. Optimization for a multi-constraint truck appointment system considering morning and evening peak congestion. *Sustainability,* 13**,** 1181.

XU, D., WEI, C., PENG, P., XUAN, Q. & GUO, H. 2020. GE-GAN: A novel deep learning framework for road traffic state estimation. *Transportation Research Part C: Emerging Technologies,* 117**,** 102635.

YI, S., SCHOLZ-REITER, B., KIM, T. & KIM, K. H. 2019. Scheduling appointments for container truck arrivals considering their effects on congestion. *Flexible Services and Manufacturing Journal,* 31**,** 730-762.

YOSHII, T., AJISAWA, S. & KUWAHARA, M. Impacts on traffic congestion by switching routes and shifting departure time of trips. 5th World Congress on Intelligent Transport Systems Proceedings, 1998.

YOU, S. I., CHOW, J. Y. & RITCHIE, S. G. 2016. Inverse vehicle routing for activity-based urban freight forecast modeling and city logistics. *Transportmetrica A: Transport Science,* 12**,** 650-673.

ZANG, D., LING, J., WEI, Z., TANG, K. & CHENG, J. 2018. Long-term traffic speed prediction based on multiscale spatio-temporal feature learning network. *IEEE Transactions on Intelligent Transportation Systems,* 20**,** 3700-3709.

ZEHENDNER, E. & FEILLET, D. 2014. Benefits of a truck appointment system on the service quality of inland transport modes at a multimodal container terminal. *European Journal of Operational Research,* 235**,** 461-469.

ZHANG, M. & TAVASSZY, L. A. 2012. *Intermodal terminal network of The Netherlands* [Online]. Available: https://www.researchgate.net/publication/345898422_Intermodal_terminal_network_of_The_Netherlands [Accessed].

ZHANG, X., ZENG, Q. & CHEN, W. 2013. Optimization model for truck appointment in container terminals. *Procedia-Social and Behavioral Sciences,* 96**,** 1938-1947.

ZHANG, Y. & YE, Z. 2008. Short-term traffic flow forecasting using fuzzy logic system methods. *Journal of Intelligent Transportation Systems,* 12**,** 102-112.

ZHAO, W. & GOODCHILD, A. V. 2010. The impact of truck arrival information on container terminal rehandling. *Transportation Research Part E: Logistics and Transportation Review,* 46**,** 327-343.

ZHAO, W. & GOODCHILD, A. V. 2013. Using the truck appointment system to improve yard efficiency in container terminals. *Maritime Economics & Logistics,* 15**,** 101-119.

ZHENG, H., LIN, F., FENG, X. & CHEN, Y. 2020. A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*.

ZHOU, J., CUI, G., ZHANG, Z., YANG, C., LIU, Z., WANG, L., LI, C. & SUN, M. 2018. Graph neural networks: A review of methods and applications. *arXiv preprint arXiv:1812.08434*.

ZHOU, W., CHEN, Q. & LIN, J. J. 2014. *Empirical study of urban commercial vehicle tour patterns in Texas* [Online]. [Accessed].

ZOU, M., LI, M., LIN, X., XIONG, C., MAO, C., WAN, C., ZHANG, K. & YU, J. 2016. An agent-based choice model for travel mode and departure time and its case study in Beijing. *Transportation Research Part C: Emerging Technologies,* 64**,** 133-147.

Summary

This thesis takes initial steps towards introducing a data-driven integrated logistics and traffic modelling framework. The main objective is to unravel the complex interaction between freight transport and traffic systems and to incorporate this knowledge into measures for improving the performance of traffic and logistics operations. Using large databases of observed truck trips and empirical research, a data-driven modelling pipeline is developed and applied, leading to new knowledge about the organization of road transport, in time and space. The modelling pipeline includes all required steps, from data pre-processing and data fusion to pattern recognition in the routing and scheduling of freight movements. We describe the main results below.

First, we use the power of machine learning techniques to identify clusters of transport markets linking them to sectors. Using this new method, we identify 9 transport markets with homogeneous pickup and delivery patterns. We also visualize the interrelation between different sectors using the tour patterns of freight transport. This sheds new light on the relationship between the inputs and outputs of different industries. The new segmentation of the transport market also improves the accuracy of freight transport demand models.

Next, a new decision tree algorithm is proposed to characterize the structure of tours from time of the day and type of tour perspectives. This method investigates the spatial-temporal correlation between the structure of tours with zonal congestion levels. The result of this study shows that the type of tour and time of day decisions in tour planning has a strong correlation with the congestion level of the visiting locations in tours. It also indicates that trip patterns in different transportation markets depend on the type of logistics activities. This underlines the importance of including the type of logistic activities in freight transport models.

The insight gained into the characteristics of planned trips makes it interesting to develop a model that can systematically incorporate planners' preferences into assigning routes and planning times of freight transport. However, these preferences cannot be observed and must be derived from known trips. The proposed model is a descriptive and parametric routing and scheduling algorithm which can be used to learn trip patterns and scheduling preferences. These parameters link a set of zonal and route features to the total cost of routing and scheduling and thus incorporate the perceptions of planners. To estimate the learning parameters of this model, an efficient and fast algorithm based on Bayesian optimization is used. We used a set of disaggregate tour data with total tour distance, duration, departure and end time. The results show that carriers value distance-related travel costs more than travel time and are generally more sensitive to travel time and costs during the afternoon peak period. This method can then generate synthetic tours for different times of the day that are close to reality. It is therefore a

strong new tool for modelling and forecasting freight transport and thus for evaluating policies related to freight traffic.

The research then focuses on predicting the impact of trips on traffic dynamics. This impact was investigated using a novel multi-layer neural network with automatic feature extraction. With this method, short-term patterns were found in the dynamics of truck flows, in relation to known truck departure times. The method has been extended to a network-wide traffic prediction model that can predict short-term traffic states (flow, speed and delays) on a given road network, also based on given truck schedules. We applied insights from traffic flow theory and transportation modelling to design a deep learning network with a specific topology that can capture meaningful spatial and temporal patterns on a congested road network, and link these patterns to the truck schedules. The results are promising, showing that the model can predict traffic states with high accuracy and that the predictions respond logically to truck demand variation.

The proposed data-driven traffic model has been applied in a predictive control mechanism to assess the impact of optimally shifting truck departure times out of peak hours on the traffic system. The result of this experiment shows that we can obtain measurable gains from the traffic system by shifting the departure time of only 10% of trucks from a logistic hub. Despite this interesting result, the applicability in practice is still limited. Achieving such a change requires active involvement and cooperation of different actors. Road users benefit directly from this policy, while carriers have to bear the reorganization costs to change their business model. A public-private gain-sharing model is needed to make compensation possible. This could be developed under the direction of a neutral third party.

Finally, the model is further extended to port and hinterland operations. Optimization of departure times is performed for a system-optimal state. The descriptive model of the combined logistics and traffic system is thereby coupled to an optimization model, which incorporates preferences and business models of various actors, including carriers, terminals, hinterland hubs, and traffic agencies. The system includes an LSTM deep neural network for arrival time forecasting, data-driven models for traffic and truck scheduling, and a descriptive queuing model for terminals. This integrated model of logistics, freight, and traffic consistently shows the costs and benefits of joint optimization of different actors. The result is a decision support system that achieves an optimal effect with realistic, minimal shifts in carriers' schedules. For the deep-sea terminal, new scenarios emerge with significantly lower waiting times and new optimization possibilities for the deployment of transshipment equipment.

In summary, the new modelling framework leads to new optimization opportunities for the port community and road users in the hinterland, addressing both the interests and preferences of the actors. The developed techniques show that data-driven approaches are feasible and pave the way for large-scale applications in a new generation of freight transport models.

Samenvatting

Deze dissertatie zet eerste stappen naar de introductie van een data-gedreven, geïntegreerd logistiek en verkeersmodelleringsraamwerk. Het doel hiervan is om de complexe interactie tussen vrachtvervoer en verkeerssystemen te ontrafelen en deze kennis te integreren in maatregelen voor het verbeteren van de prestaties van het verkeer en van logistieke operaties. Gebruikmakend van grote databases met waargenomen vrachtwagenritten en met empirisch onderzoek is een data-gedreven modelleerpijplijn ontwikkeld, die leidt tot nieuwe kennis over de organisatie van het wegvervoer, in tijd en ruimte. Het modelleerraamwerk omvat alle vereiste stappen, van data-voorbewerking en datafusie tot patroonherkenning in de *routing and scheduling* van vrachtbewegingen.

Allereerst hebben wij de kracht van *machine learning* technieken gebruikt om clusters van transportmarkten vast te stellen en deze aan sectoren te koppelen. Met deze nieuwe methode identificeren wij 9 transportmarkten met homogene ophaal- en afleverpatronen. Wij hebben ook het verband tussen de verschillende sectoren gevisualiseerd aan de hand van de rondritpatronen van het vrachtvervoer. Dit werpt nieuw licht op de relatie tussen de inputs en outputs van verschillende industrieën. De nieuwe segmentatie van de vervoersmarkt verbetert ook de nauwkeurigheid van de modellering van de vervoersvraag.

Vervolgens is een nieuw beslisboomalgoritme ontwikkeld om de structuur van ritten te karakteriseren vanuit het oogpunt van tijdstip en vorm van ritten, en is de ruimtelijk-temporele correlatie onderzocht tussen de structuur van ritten en zonale congestieniveaus. Het resultaat van deze studie laat zien dat keuzes omtrent het rittype en –timing een sterke correlatie hebben met het congestieniveau van de bezoekende locaties. Ook blijkt dat ritpatronen in de verschillende vervoersmarkten afhankelijk zijn van het type logistieke activiteit. Hiermee is het belang van het meenemen van het type logistieke activiteiten voor het modelleren van het goederenvervoer aangetoond.

Het opgedane inzicht in de kenmerken van geplande ritten maakt het interessant om een model te ontwikkelen dat de voorkeuren van planners systematisch kan integreren in route- en tijdplanning. Deze voorkeuren kunnen echter niet worden waargenomen en moeten afgeleid worden van bekende ritten. Het voorgestelde model is een beschrijvend en parametrisch routerings- en planningsalgoritme, waarmee ritpatronen en planningsvoorkeuren geleerd kunnen worden. De parameters koppelen een set van zonale en routekenmerken aan de totale kosten van routing en scheduling en bevatten daarmee de perceptie van de planners. Om de leerparameters van dit model te schatten, wordt een efficiënt en snel algoritme gebruikt op basis van Bayesiaanse optimalisatie. We gebruikten een set van gedesaggregeerde tour data met totale ritafstand, -duur, vertrek en eindtijd. De resultaten laten zien dat vervoerders afstand-

gerelateerde reiskosten hoger waarderen dan reistijd, en in het algemeen gevoeliger zijn voor reistijd en kosten tijdens de middagspitsperiode. Deze methode kan vervolgens synthetische ritten genereren, voor verschillende tijdstippen op de dag, die dicht bij de realiteit liggen. Het is daarom een sterk nieuw instrument voor het modelleren en prognotiseren van vrachtvervoer en daarmee voor het evalueren van beleid gerelateerd aan vrachtverkeer.

Het onderzoek heeft zich vervolgens gericht op het voorspellen van de impact van ritten op de dynamiek van het verkeer. Deze impact is onderzocht met een nieuw meerlagig neuraal netwerk met automatische kenmerkextractie. Op deze wijze zijn korte-termijn patronen gevonden in de dynamiek van vrachtwagenstromen, mede in relatie tot de bekende vertrektijden van vrachtwagens. De methode is uitgebreid tot een netwerkbreed verkeersvoorspellingsmodel dat verkeerssituaties op korte termijn kan voorspellen (doorstroming, snelheid en vertragingen) op een gegeven wegennet, eveneens op basis van gegeven vrachtwagenschema's. We hebben inzichten uit verkeersstroomtheorie en transportmodellering toegepast om een *deep learning* netwerk met een specifieke topologie te ontwerpen dat zinvolle ruimtelijke en temporele patronen kan vastleggen op een overbelast wegennetwerk, en deze patronen kan koppelen aan de vrachtwagenschema's. De resultaten zijn veelbelovend en tonen aan dat het model verkeerstoestanden met hoge nauwkeurigheid kan voorspellen en dat de voorspellingen logisch reageren op de variatie in de vraag naar vrachtwagens.

Het voorgestelde data-driven verkeersmodel is toegepast in een voorspellend controle-mechanisme om de impact van een optimale verschuiving van de vertrektijden van vrachtwagens uit de spitsuren op het verkeerssysteem te beoordelen. Het resultaat van dit experiment toont aan dat we meetbare winst uit het verkeerssysteem kunnen halen door het verschuiven van de vertrektijd van slechts 10% van de vrachtwagens van een logistiek knooppunt. Ondanks dit interessante resultaat is de toepasbaarheid in de praktijk vooralsnog beperkt. Om een dergelijke verandering te realiseren, is actieve betrokkenheid en samenwerking nodig van verschillende actoren. Weggebruikers hebben direct baat bij dit beleid, terwijl vervoerders de reorganisatiekosten moeten dragen om hun bedrijfsmodel aan te passen. Een publiek-privaat winstdelingsmodel is nodig om compensatie mogelijk te maken. Deze zou onder regie van de overheid ontwikkeld kunnen worden.

Tenslotte is het model verder uitgebreid naar hubs in het achterland. Hierbij wordt een optimalisatie van vertrektijden uitgevoerd voor een systeem-optimale toestand. Het beschrijvende model van het gecombineerde logistieke- en verkeerssysteem is daarbij gekoppeld aan een optimalisatiemodel, waarin voorkeuren en business modellen van diverse actoren zijn meegenomen, inclusief vervoerders, terminals, achterlandhubs en verkeersdeelnemers. Het systeem omvat een LSTM deep neuraal netwerk voor prognose van aankomsttijden, data-gedreven modellen voor verkeer en truck scheduling en een beschrijvend wachtrijmodel voor terminals. Deze integrale aanpak van logistiek, goederenvervoer en verkeer laat op consistente wijze de kosten en baten zien van gezamenlijke optimalisatie van verschillende actoren. Het resultaat is een beslissingsondersteunend systeem waarmee met realistische, minimale verschuivingen van de planning van vervoerders een optimaal effect wordt geboekt. Voor de deepsea terminal ontstaan nieuwe scenario’s met significant lagere wachttijden en nieuwe optimalisatiemogelijkheden voor de inzet van overslagmaterieel.

Samengevat leidt het nieuwe modelleringsraamwerk tot nieuwe optimalisatiemogelijkheden voor de havencommunity en de weggebruikers in het achterland, waarbij zowel aan de belangen als de voorkeuren van de actoren recht wordt gedaan. De ontwikkelde technieken laten zien dat

data-gedreven aanpakken haalbaar zijn en effenen het pad voor grootschalige toepassingen in een nieuwe generatie goederenvervoermodellen.

About the Author

Ali Nadi was born in Isfahan, Iran. He finished high school with a special focus on mathematics and physics. In 2007, he left his hometown to start his bachelor in Civil engineering at Arak University. In 2012, he went to Tehran to study Transportation Engineering at one of the best universities in Iran, K.N. Toosi University of Technology. His research area was on managing logistics in natural disasters to improve emergency responses. He was awarded his MSc degree with the top grade in 2014.

After his graduation, Ali worked in a consultancy and was involved in several research projects in collaboration with Esfahan University and regional Municipalities. After 4 years of working between practice and science he decided to explore the scientific career path further.

In 2018, Ali commenced his PhD education at Delft University of Technology (TU Delft), the Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, in The Netherlands. He started his PhD journey by joining the Freight and Logistics Lab and the Data Analytics and Traffic Simulation Lab (DiTTLab) to work on the “ToGRIP-Grip on Freight Trips" project. His research was funded by Netherlands Organization for Scientific Research (NWO) and supported by TKI Dinalog, Commit2data, Port of Rotterdam, SmartPort, Portbase, TLN, Deltalinqs, Rijkswaterstaat, and TNO.

In his PhD research, Ali worked on developing methods for integrative modelling of Logistics and traffic systems using advanced data-driven and optimization techniques. He applied his knowledge in mathematical modelling and data analytics to integrate machine learning and artificial intelligence with operational research and theory-driven methods to model and control demand and supply interactions in both freight transport and traffic systems.

List of Journal Articles

1. Nadi, A., Sharma, S., Snelder, M., Bakri, T., van Lint, H., & Tavasszy, L. (2021). Short-term prediction of outbound truck traffic from the exchange of information in logistics hubs: A case study for the port of Rotterdam. *Transportation Research Part C: Emerging Technologies*, *127*, 103111. DOI: https://doi.org/10.1016/j.trc.2021.103111

2. Nadi, A., Sharma, S., van Lint, J. W. C., Tavasszy, L., & Snelder, M. (2022). A data-driven traffic modeling for analyzing the impacts of a freight departure time shift policy. Transportation Research Part A: Policy and Practice, 161, 130-150. DOI: https://doi.org/10.1016/j.tra.2022.05.008

3. Nadi, A., Snelder, M., van Lint, J.W.C., & Tavasszy, L. (2022). Spatial and temporal characteristics of freight tours: a data-driven exploratory analysis – submitted to a journal.

4. Nadi, A., Nugteren, A., Snelder, M., van Lint, J.W.C., Rezaei, J. (2022). Truck Arrival Shift: An *Advisory*-Based Time Slot Management System to Mitigate Waiting Time at Container Terminal gates. *Transportation Research Record, 1-14*. DOI: https://doi.org/10.1177/03611981221090940

5. Nadi, A., Snelder, M., van Lint, J.W.C., Tavasszy, L. (2022). A Data-driven and multi-actor decision support system for time slot management at container terminals: A case study for the Port of Rotterdam – submitted to a journal.

6. Nadi, A., Yorke-Smith, N., Snelder, M., van Lint, J.W.C., Tavasszy, L. (2022). Data-Driven Preference-Based Routing and Scheduling for Activity-Based Freight Transport Modelling – submitted to a journal.

7. Mohammed, R.A., Nadi, A., Tavasszy, L., de Bok, M. (2022). A data fusion approach to identify distribution chain segments in freight shipment databases. *Transportation Research Record*, accepted – in press .

Peer-reviewed Conference papers

1- Nadi, A., Snelder, M., Tavasszy, L., Sharma, S., & Van Lint, H. (2018). Truck identification on freeways using Bluetooth data analysis, 26-31 May 2019. IIT Bombay.

2- Nadi, A., Van Lint, H., Tavasszy, L., & Snelder, M. (2020, September). Identifying tour structures in freight transport by mining of large trip databases. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1-7). IEEE.

3- Nadi, A., Nugteren, A., Snelder, M., van Lint, J.W.C., Rezaei, J. (2022). Truck Arrival Shift: An Advisory-Based Time Slot Management System to Mitigate Waiting Time at Container Terminal gates. *Transportation Research Board*.

4- Nadi, A., Yorke-Smith, N., Snelder, M., van Lint, J.W.C., Tavasszy, L. (2022). A data-driven routing and scheduling for Activity-based freight transport modelling. *Extended abstract submitted for presentation at the 11th Triennial Symposium on Transportation Analysis conference (TRISTAN XI) June 19-25, 2022, Mauritius Island.*

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 275 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.

Nadi Najafabadi, A., *Data-Driven Modelling of Routing and Scheduling in Freight Transport,* T2022/14, October 2022, TRAIL Thesis Series, the Netherlands

Heuvel, J. van den, *Mind Your Passenger! The passenger capacity of platforms at railway stations in the Netherlands,* T2022/13, October 2022, TRAIL Thesis Series, the Netherlands

Haas, M. de, *Longitudinal Studies in Travel Behaviour Research,* T2022/12, October 2022, TRAIL Thesis Series, the Netherlands

Dixit, M., *Transit Performance Assessment and Route Choice Modelling Using Smart Card Data,* T2022/11, October 2022, TRAIL Thesis Series, the Netherlands

Du, Z., *Cooperative Control of Autonomous Multi-Vessel Systems for Floating Object Manipulation,* T2022/10, September 2022, TRAIL Thesis Series, the Netherlands

Larsen, R.B., *Real-time Co-planning in Synchromodal Transport Networks using Model Predictive Control,* T2022/9, September 2022, TRAIL Thesis Series, the Netherlands

Zeinaly, Y., *Model-based Control of Large-scale Baggage Handling Systems: Leveraging the theory of linear positive systems for robust scalable control design,* T2022/8, June 2022, TRAIL Thesis Series, the Netherlands

Fahim, P.B.M., *The Future of Ports in the Physical Internet,* T2022/7, May 2022, TRAIL Thesis Series, the Netherlands

Huang, B., *Assessing Reference Dependence in Travel Choice Behaviour,* T2022/6, May 2022, TRAIL Thesis Series, the Netherlands

Reggiani, G., *A Multiscale View on Bikeability of Urban Networks*, T2022/5, May 2022, TRAIL Thesis Series, the Netherlands

Paul, J., *Online Grocery Operations in Omni-channel Retailing: opportunities and challenges,* T2022/4, March 2022, TRAIL Thesis Series, the Netherlands

Liu, M., *Cooperative Urban Driving Strategies at Signalized Intersections,* T2022/3, January 2022, TRAIL Thesis Series, the Netherlands

Feng, Y., *Pedestrian Wayfinding and Evacuation in Virtual Reality*, T2022/2, January 2022, TRAIL Thesis Series, the Netherlands

Scheepmaker, G.M., *Energy-efficient Train Timetabling,* T2022/1, January 2022, TRAIL Thesis Series, the Netherlands

Bhoopalam, A., *Truck Platooning: planning and behaviour,* T2021/32, December 2021, TRAIL Thesis Series, the Netherlands

Hartleb, J., *Public Transport and Passengers: optimization models that consider travel demand,* T2021/31, TRAIL Thesis Series, the Netherlands

Azadeh, K., *Robotized Warehouses: design and performance analysis,* T2021/30, TRAIL Thesis Series, the Netherlands

Chen, N., *Coordination Strategies of Connected and Automated Vehicles near On-ramp Bottlenecks on Motorways*, T2021/29, December 2021, TRAIL Thesis Series, the Netherlands

Onstein, A.T.C., *Factors influencing Physical Distribution Structure Design*, T2021/28, December 2021, TRAIL Thesis Series, the Netherlands

Olde Kalter, M.-J. T., *Dynamics in Mode Choice Behaviour*, T2021/27, November 2021, TRAIL Thesis Series, the Netherlands

Los, J., *Solving Large-Scale Dynamic Collaborative Vehicle Routing Problems: an Auction-Based Multi-Agent Approach,* T2021/26, November 2021, TRAIL Thesis Series, the Netherlands

Khakdaman, M., *On the Demand for Flexible and Responsive Freight Transportation Services,* T2021/25, September 2021, TRAIL Thesis Series, the Netherlands

Wierbos, M.J., *Macroscopic Characteristics of Bicycle Traffic Flow: a bird's-eye view of cycling,* T2021/24, September 2021, TRAIL Thesis Series, the Netherlands

Qu, W., *Synchronization Control of Perturbed Passenger and Freight Operations,* T2021/23, July 2021, TRAIL Thesis Series, the Netherlands

Nguyen, T.T., *Highway Traffic Congestion Patterns: Feature Extraction and Pattern Retrieval,* T2021/22, July 2021, TRAIL Thesis Series, the Netherlands