TU Delft

Delft University of Technology

A data-driven traffic modeling for analyzing the impacts of a freight departure time shift policy

Nadi, Ali; Sharma, Salil; van Lint, J. W.C.; Tavasszy, Lóránt; Snelder, Maaike

DOI
10.1016/j.tra.2022.05.008

Publication date
2022

Document Version
Final published version

Published in
Transportation Research Part A: Policy and Practice

Citation (APA)
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. https://doi.org/10.1016/j.tra.2022.05.008

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

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Transportation Research Part A 161 (2022) 130–150

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

A data-driven traffic modeling for analyzing the impacts of a freight departure time shift policy

Ali Nadi [a,*], Salil Sharma [a], J.W.C. van Lint [a], Lóránt Tavasszy [a], Maaike Snelder [a,b]

[a] Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands
[b] TNO, Anna van Buerenplein 1, 2595 DA The Hague, The Netherlands

ARTICLE INFO

Keywords:
Freight departure time shifts
Freight transport policy
Predictive departure time advice
Data-driven traffic modelling
Graph convolutional deep neural network

ABSTRACT

This paper proposes a data-driven transport modeling framework to assess the impact of freight departure time shift policies. 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.

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 modeling 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?

* Corresponding author.
E-mail address: a.nadinajafabadi@tudelft.nl (A. Nadi).

https://doi.org/10.1016/j.tra.2022.05.008
Received 22 July 2020; Received in revised form 7 January 2022; Accepted 12 May 2022
Available online 23 May 2022
0965-8564/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Candidate modeling 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 modeling study of an optimized FDTS scheme linked to overall road network traffic conditions.

The remainder of this paper is organized as follows: Section 2 provides a general overview of the existing studies about departure time shift policies and methods. Section 3 presents the model for short-term traffic prediction and the data-driven decision support system for FDTS. Section 4 denotes results, tests the predictive capabilities of the model, presents the designed scenarios, and discusses policy implications. Finally, section 5 offers the conclusions and recommendations of the paper.

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.

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 m. A subset of these loop detectors can distinguish vehicle categories based on their lengths. These data are available at a resolution level of 1 min time periods.

We collected time series of volumes and speeds for five motorways near the port of Rotterdam. Table 1 and Fig. 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.

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 Table1 equals the number of sensors divided by 3). We also aggregated in time to 5 min 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.

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 Fig. 2, and consists of two main modules: (1) a data-driven traffic forecasting model and (2) a predictive departure time control model.

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.

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 $X^{t+\Delta t}$.

$$X^{t+\Delta t} \cong f(X^t, X^{t-\Delta t}, X^{t-2\Delta t}, \ldots, X^{t-d\Delta t}, C^t, C^{t-\Delta t}, \ldots, C^{t-d\Delta t} \mid G) \tag{1}$$

Fig. 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

Table 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

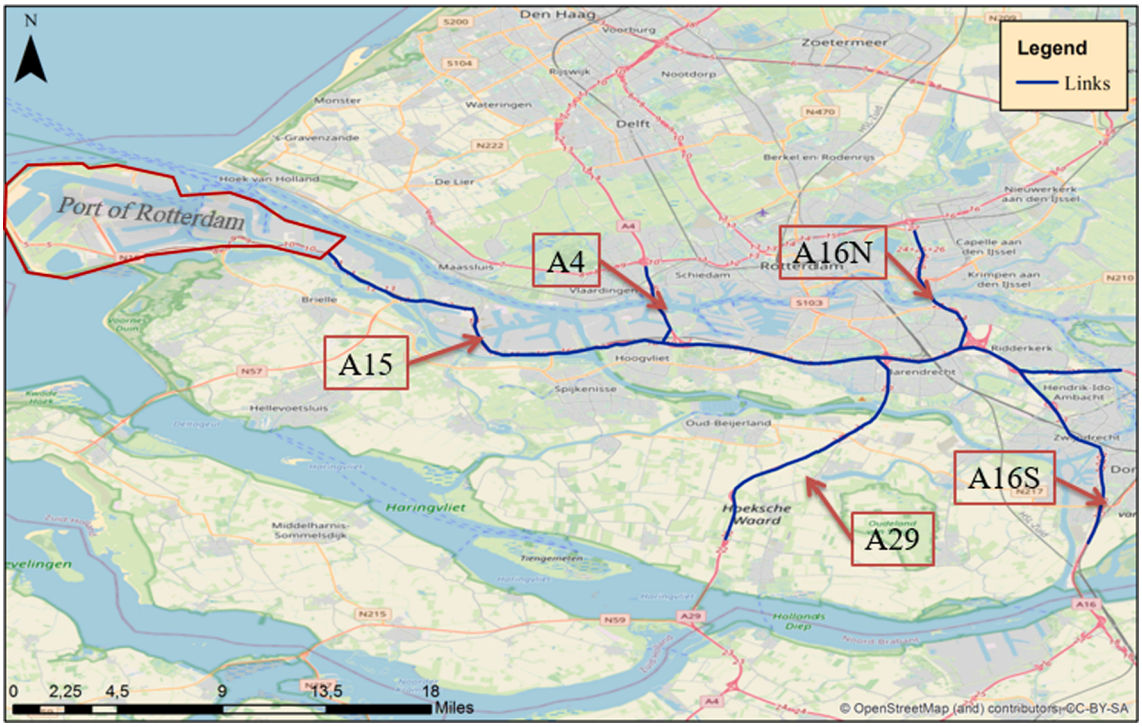

Fig. 1. The road network that provides accessibility for the Port of Rotterdam and Rotterdam city.

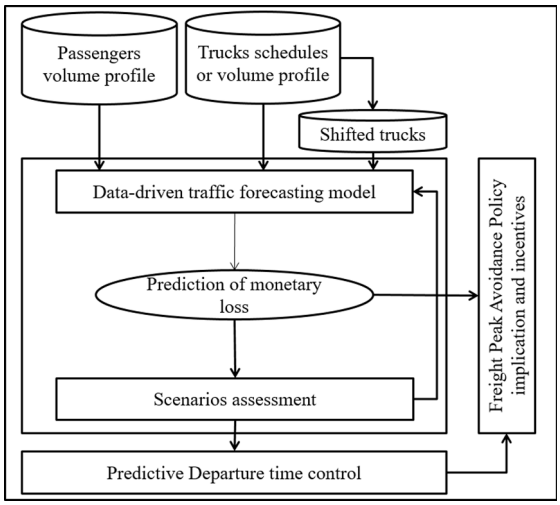

Fig. 2. Data-driven decision support framework for freight departure time and peak avoidance policy.

input and output to each other and predict the desired target variable.

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 Fig. 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.

4.1.1. Spatial and temporal dependencies

We model the spatial dependencies by using graph convolution, message passing, and neighborhood aggregation techniques. Graph

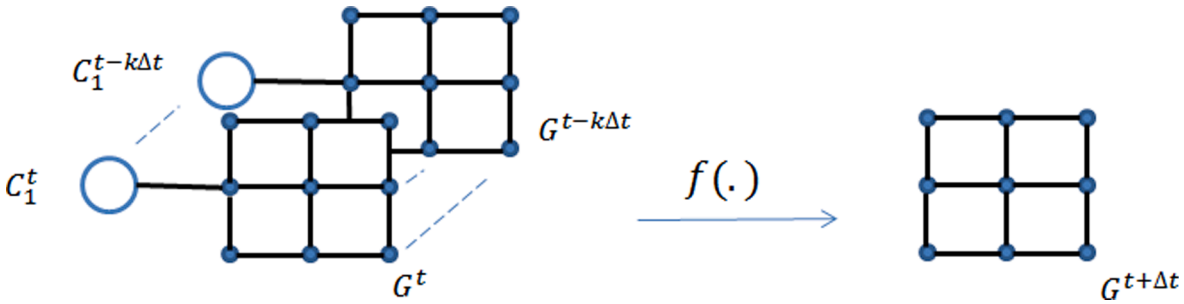

Fig. 3. An example road network with one centroid in a graph representation.

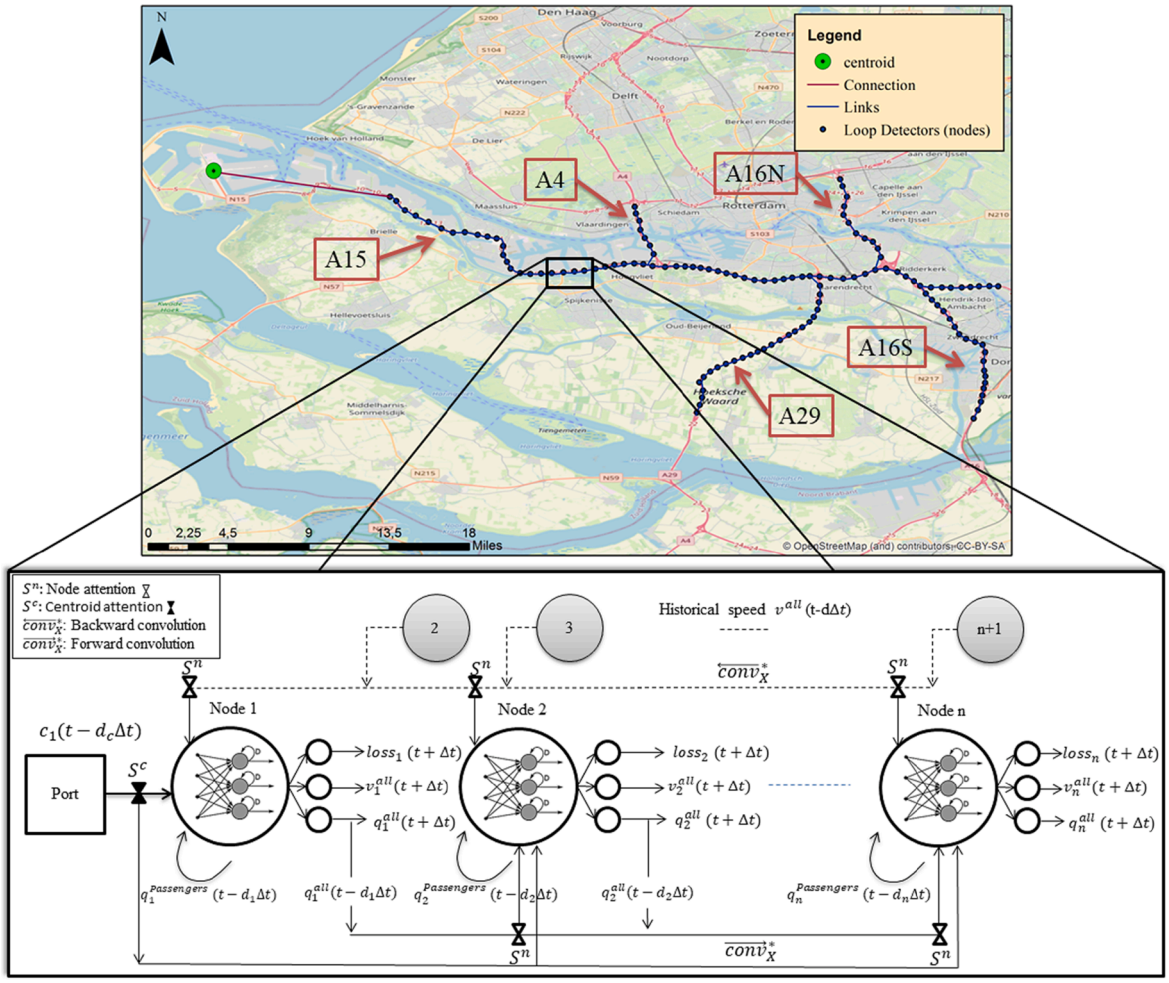

Fig. 4. Graphical representation of the structure of the proposed graph-based modular convolutional neural network model.

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).

$$conv^* = \overline{D}^{-0.5}\widetilde{A}\overline{D}^{-0.5}\overline{X}W \tag{2}$$

$$D_{ii} = \sum_j \widetilde{A}_{ij} \tag{3}$$

$$\widetilde{A} = A + I \tag{4}$$

where D is the diagonal degree matrix, I is the identity matrix and $\widetilde{A}$ 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 (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 (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.

$$S_{ij}^{n} = \exp\left(-\frac{1}{2}\left(\frac{p_{ij}}{\sigma_i^n}\right)^2 \right) \tag{5}$$

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) 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 $\widetilde{A}$ (element-wise) and then will be fed to the convolution operator for aggregation of input features. Equation (6) shows graph convolution operator with node attention on $\overline{X}$ input features.

$$conv_X^* = \overline{D}^{-0.5}\left(S^n \otimes \widetilde{A}\right)\overline{D}^{-0.5}\overline{X}W_X \tag{6}$$

Where W_x is learnable parameter for convolution operator on $\times$ 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 Fig. 4).

In addition to the input features $\overline{X}$ 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 (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.

$$E_{mi} = \frac{e^{c_m}}{\sum_{j=1}^{M} e^{c_j}} \quad \forall i \in N \tag{7}$$

Equation (8) shows the graph convolution operator with centroid attention on $\overline{C}$ input features.

$$conv_C^* = (\overline{E}S^c \otimes \overline{O})\overline{C}W_c \tag{8}$$

Where $\overline{E}$ is a matrix of decision gates with elements of E_{mi}, $\overline{O}$ 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, $\overline{C}$ 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.

$$f(\overline{X}, \overline{C}|G) = g\left(conv_X^* \times \theta_1 + b_1, conv_C^* \times \theta_2 + b_2\right) \tag{9}$$

Equation (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 $f(\overline{X}, \overline{C}|G)$ is the output of the layer given input $\overline{X}$ and $\overline{C}$ on graph G.

For the temporal dependency, we considered a delay matrix $d \in \mathbb{N}^{N\times(N+n)}$ (see equation (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.

4.1.2. 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 Fig. 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.

Fig. 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.

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.

$$X^0 = conv_c^*(C(t - d_c\Delta t)) \tag{10}$$

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 $\overrightarrow{conv}_X^*$,

$$X_j^1 = \overrightarrow{conv}_X^*\left(q_j^{all}(t - k\Delta t)\right)\forall j \to i \tag{11}$$

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

$$X_i^2 = conv_X^*(q_i^p(t - d_i\Delta t)) = \overrightarrow{conv}_X^*\left(q_i^{all}(t - d_i\Delta t) - q_i^{truck}(t - d_i\Delta t)\right) \tag{12}$$

- 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 $\overleftarrow{c}\ onv_X^*$.

$$X_j^4 = \overleftarrow{c}\ onv_X^*\left(v_j(t - k\Delta t)\right)\forall i \to j \tag{13}$$

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 Fig. 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 (14) to (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, 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 100 km/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 (14) to (22) show how we calculate the monetary loss for each node i at time t in this multi-class system.

$$tp_{i,t} = \frac{q_{i,t}^{Trucks}}{q_{i,t}^{all}} \qquad \forall i \in V \quad , \quad \forall t \in T \tag{14}$$

$$q_{i,t}^{Passengers} = \left(1 - tp_{i,t}\right) q_{i,t}^{all} \tag{15}$$

$$Vh_{i,t}^{Truck} = \frac{q_{i,t}^{Trucks} l}{v_{i,t}^{Trucks}} \tag{16}$$

$$Vh_{i,t}^{Passengers} = \frac{q_{i,t}^{Passengers} l}{v_{i,t}^{all}} \tag{17}$$

$$Vh_{free}^{Trucks} = \frac{q_{i,t}^{Trucks} l}{v_{free}^{Trucks}} \tag{18}$$

$$Vh_{free}^{Passengers} = \frac{q_{i,t}^{Passengers} l}{v_{free}^{Passengers}} \tag{19}$$

$$VLH_{i,t}^{Trucks} = \begin{cases} Vh_{i,t}^{Trucks} - Vh_{free}^{Trucks} & Vh_{i,t}^{Trucks} > Vh_{free}^{Trucks} \\ 0 & otherwise \end{cases} \tag{20}$$

$$VLH_{i,t}^{Passengers} = \begin{cases} Vh_{i,t}^{Passengers} - Vh_{free}^{Passengers} & Vh_{i,t}^{Passengers} > Vh_{free}^{Passengers} \\ 0 & otherwise \end{cases} \tag{21}$$

$$Loss_{i,t} = VoT^{\ Passengers} \cdot VLH_{i,t}^{Passengers} + VoT^{\ Trucks} \cdot VLH_{i,t}^{Trucks} \tag{22}$$

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 (14) and (15)). For each timestamp and node in the road network graph, equations (16) and (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 (18) and (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 (20) and (21)). Finally, the *Loss in equation* (22) 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.

4.1.3. 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) to (9) 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). Fig. 5 shows how the model's parameters are estimated in this bi-level parameter estimation.

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

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{23}$$

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

4.1.4. 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.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2} \tag{24}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\|t_i - y_i\| \tag{25}$$

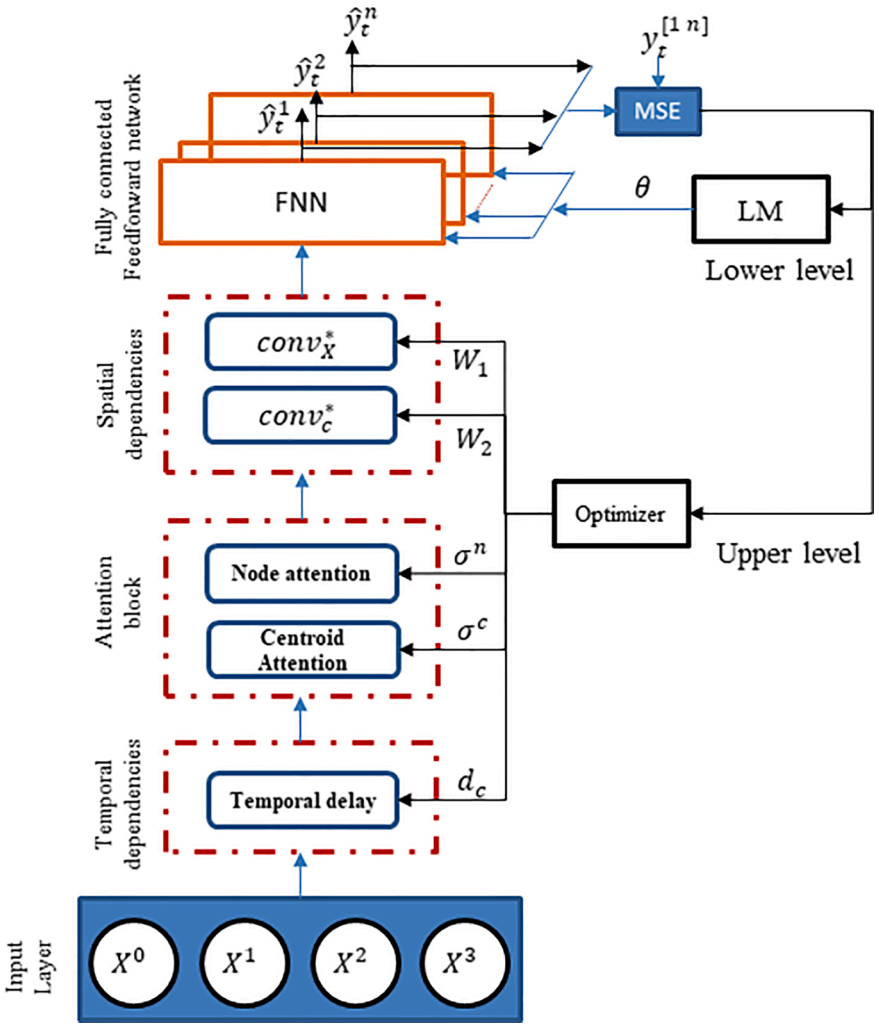

Fig. 5. High level presentation of network structure and parameter estimation procedure.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{t_i - y_i}{t_i}\right| \tag{26}$$

$$PAPE = Pr\left(\left|\frac{t_i - y_i}{t_i}\right| < 5\%\right) \tag{27}$$

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.

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 (see Fig. 6) 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.

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 $T = \{1, 2, \cdots, t\}$ intervals each has 30 min of span.
2. Assume $C_{K\times T}$ be a matrix of container flow generated by K={1,2,..,k} truck generating centroids.

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1t} \\ \vdots & \ddots & \vdots \\ c_{k1} & \cdots & c_{kt} \end{bmatrix}_{K\times T} \tag{28}$$

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 (10)) 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

$$C_k^{average} = \frac{1}{|T|}\sum_{t=1}^{T}\left(C_k^{\max} - C_t\right) \tag{29}$$

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 $M = \{1, 2, \cdots, m\}$ 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.

Given that $Y^m_{K\times T}$ is the shift matrix where its elements y^m_{ij} 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.

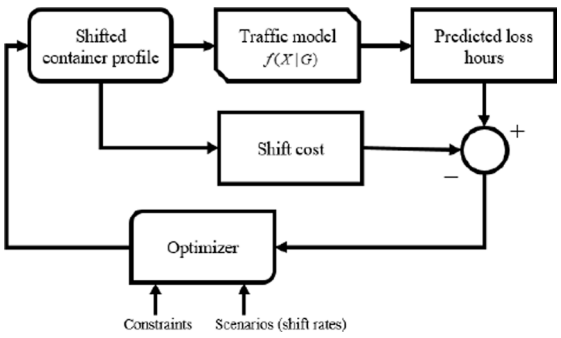

Fig. 6. A Graphical representation of departure time control procedure.

$$\widehat{C}_{K\times T} = C_{K\times T} \otimes (O_{K\times T} - \lambda_{K\times T}) + C_{K\times T} \otimes \lambda_{K\times T} \times Y^{m}_{K\times T} \tag{30}$$

Where O is an all-ones matrix, $\widehat{C}_{K\times T}$ 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 (10), we have.

$$\widehat{X}^{0} = \sigma^{c}(\widehat{C}(t - d_c\Delta t)) \tag{31}$$

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

$$\min \sum_{n=1}^{N}\sum_{t=1}^{T} f_{nt}\left(\widehat{X}^{0}|G\right) \tag{32}$$

$$\min \sum_{m=1}^{M}\sum_{i=1}^{T}\sum_{j=1}^{T} d_{ij}y^{m}_{ij} \tag{33}$$

s.t.

$$\sum_{i=1}^{T}\sum_{j}^{T} Y^{m}_{ij} = 1 \quad \forall m \in M \tag{34}$$

$$\widehat{C}_{kt} \leqslant C^{\max}_{k} \qquad \forall t \in T \tag{35}$$

$$d_{ij} \leqslant D \qquad \forall i,j \in T \tag{36}$$

$$Y^{m}_{ij} \in \{0,1\} \tag{37}$$

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 (34) ensures that all the groups of containers are assigned to one interval. Equation (35) indicates that the number of containers generated at centroid k at time t i.e $\widehat{C}_{kt}$ does not exceed the capacity of the interval t after applying the shift. Finally, Equation (36) is the time shift constraint that controls the maximum shift (D = 2 h) that can happen for all groups of containers. We used the NSGA-II algorithm to solve this multi-objective assignment problem.

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 min 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.

5.1. 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. Fig. 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.

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 Fig. 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.

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 Fig. 9. Also, visual inspection at this level of detail confirms a very satisfactory result.

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

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 2 illustrates the facts about the performance of the model under several indicators introduced in section 4.1.3. Table 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.

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.

5.2. 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. Fig. 11 shows that the model is sensitive to changes in demand, with the expected signs of impacts.

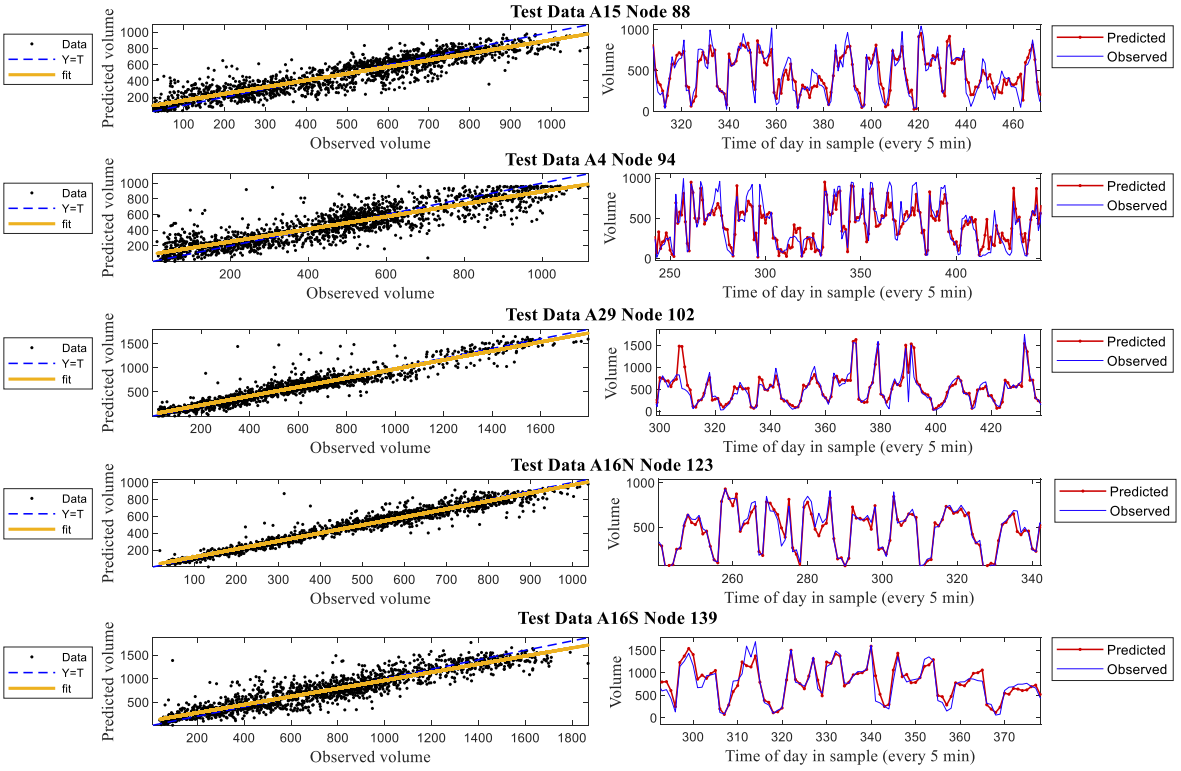

Fig. 7. Linear fit between predicted and observed volumes.

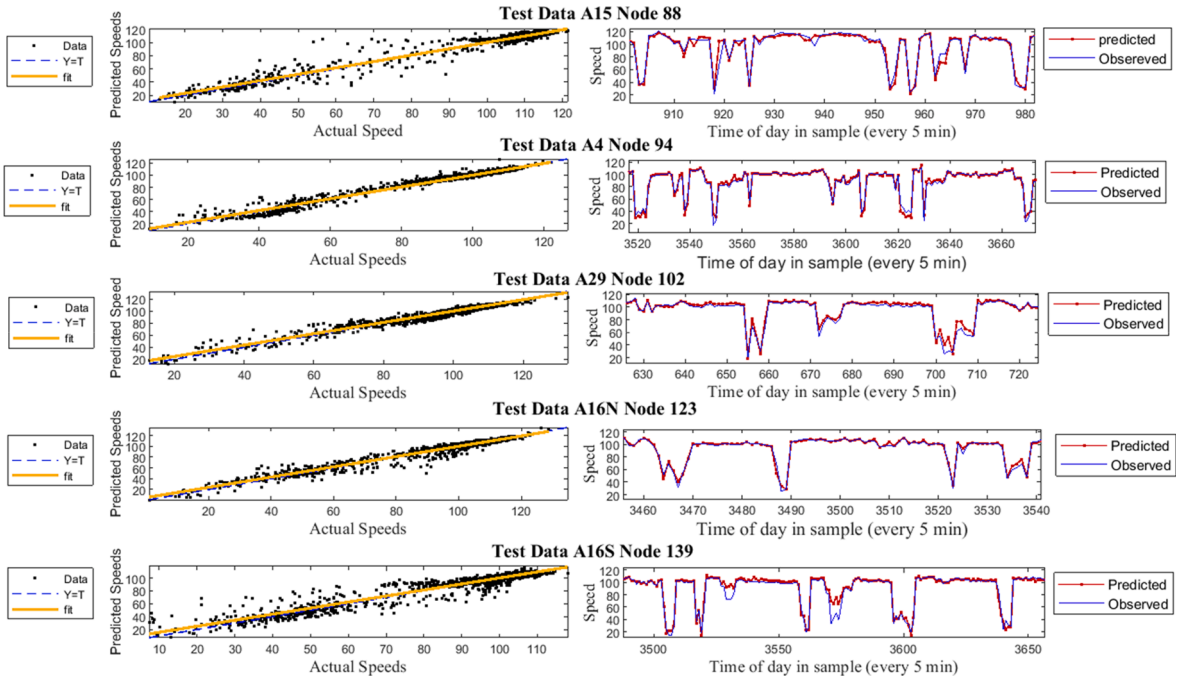

Fig. 8. Linear fit between observed and predicted speed.

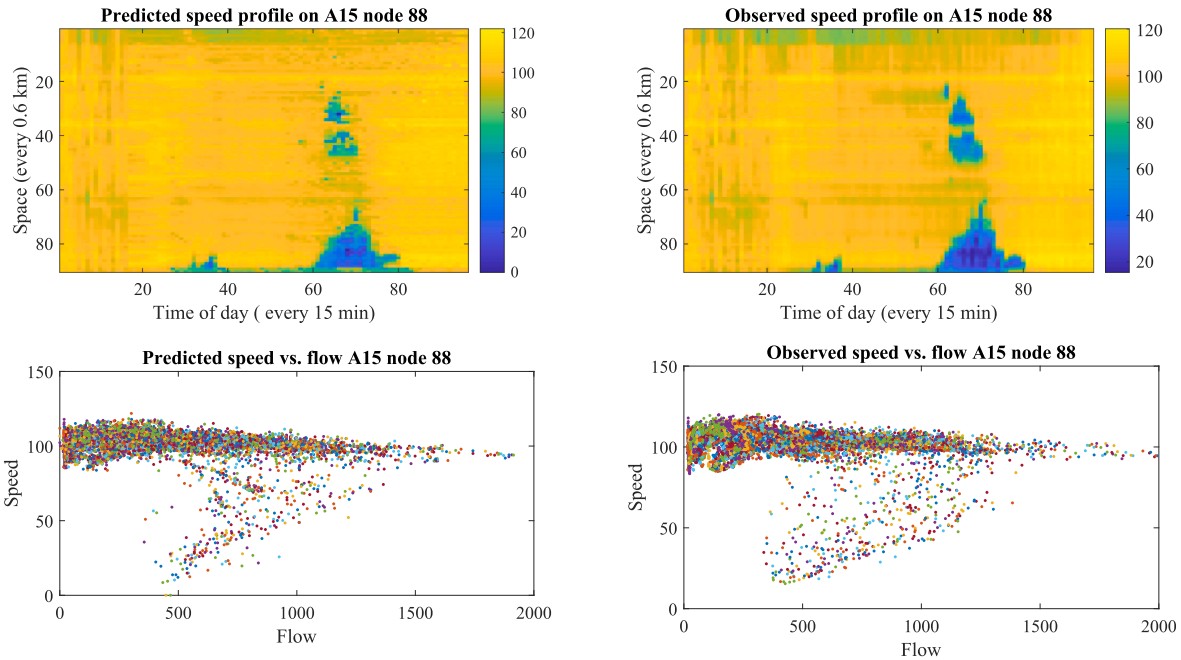

Fig. 9. Observed versus predicted space–time speed profile and speed-flow fundamental diagram.

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

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

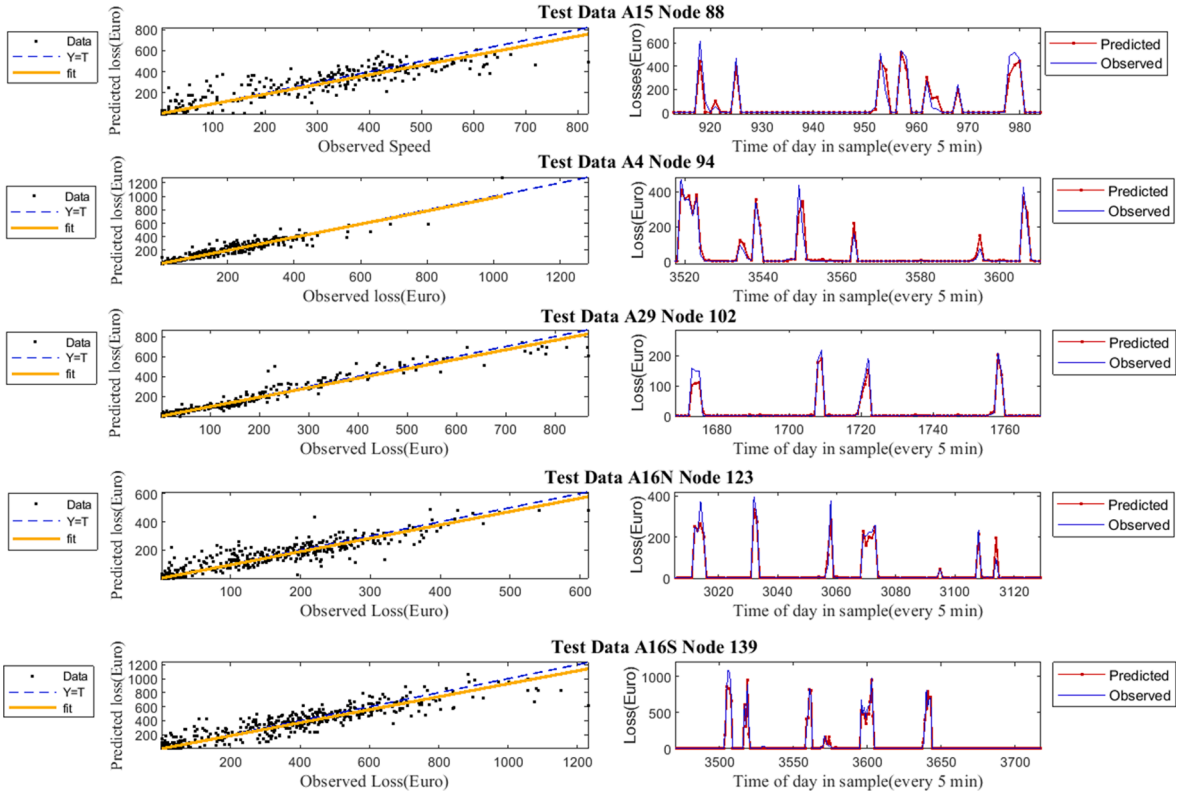

Fig. 10. Correlation coefficient for loss prediction.

Table 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	$y \approx 0.93x + 6.18$
		losses	0.081	40.90	6.40	1.77	1.5%	99%	0.95	$y \approx 0.95x + 0.61$
A16S	139	volume	5.66	14179.95	119.08	86.33	1.6%	99%	0.95	$y \approx 0.91x + 66.13$
		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	$y \approx 0.92x + 1.85$

earlier or later time. The time shift operator ranges from 30 min to 3 h across all scenarios. The gain is calculated based on the differences between the base case and the scenario losses. Fig. 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.

We can see from Fig. 12 that these time shift scenarios may lead to negative or positive gains for various days. The distribution of gains is differing from 30 min to 3 h 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

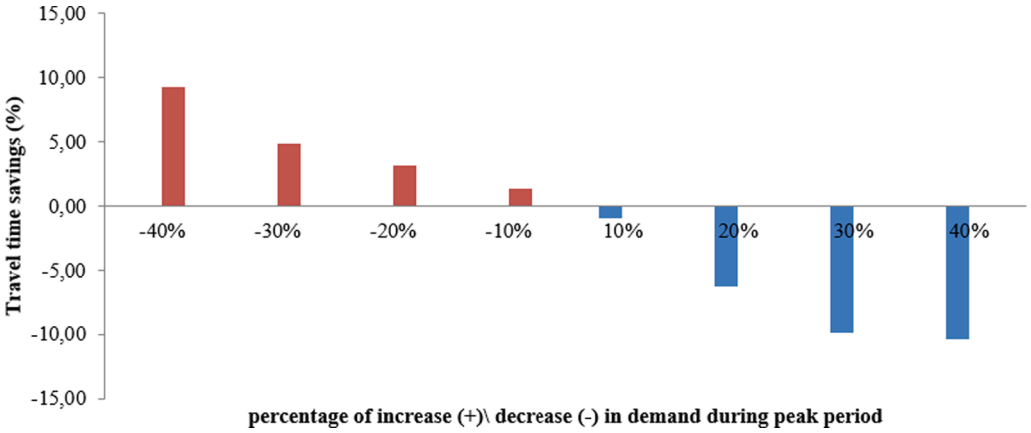

Fig. 11. Sensitivity of the model towards increasing or decreasing demand during the peak period.

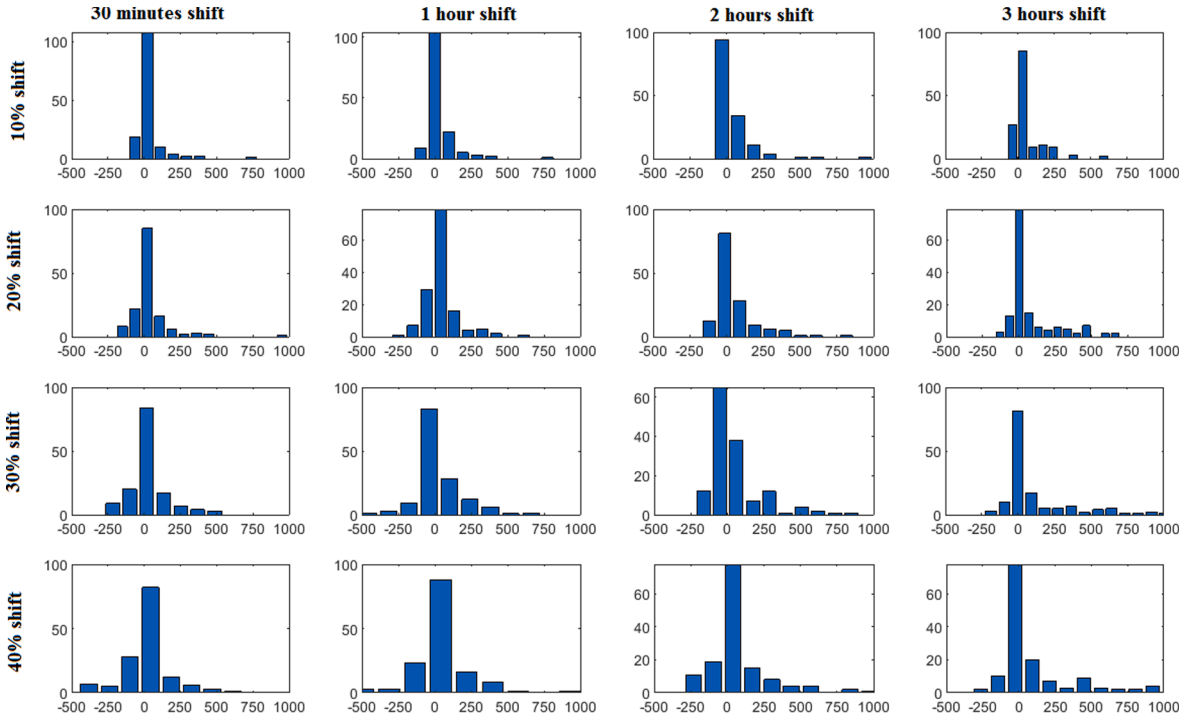

Fig. 12. Histogram of gains for different scenarios on A15.

negative for a given day, it could be positive on other motorways. For instance, Fig. 13 points to a particular day on A15 with nonrecurring congestion. This congestion lasted almost 5 h 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.

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, Fig. 14 shows the

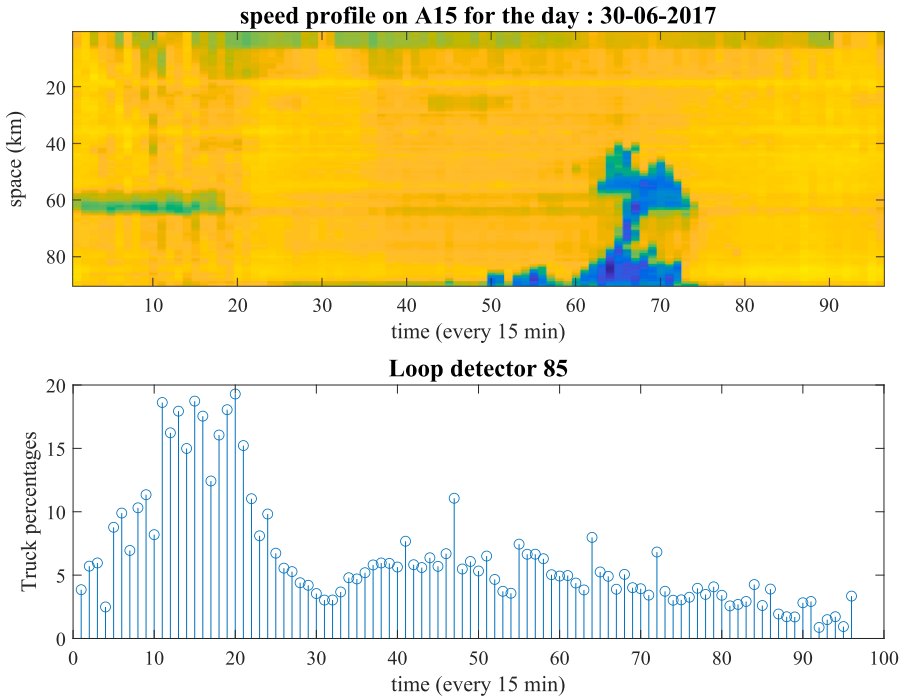

Fig. 13. changes in truck percentages due to nonrecurring congestion on A15.

Pareto frontier and the optimum solution regarding the time shift constraint.

Table 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 3 also indicates the average improvement of travel time during peak period on every 5 alternative paths after applying departure time shift policy.

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. Fig. 15 shows the predicted social benefits based on the recommended departure time for any of the shifting scenarios.

We calculate the social benefits for the afternoon peak and off-peak period (from 12:00 to 21:00). Table 4 shows that between 23,144 to 69,528 Euro can return to the carriers, depending on the percentage of participants, where the government can pay between

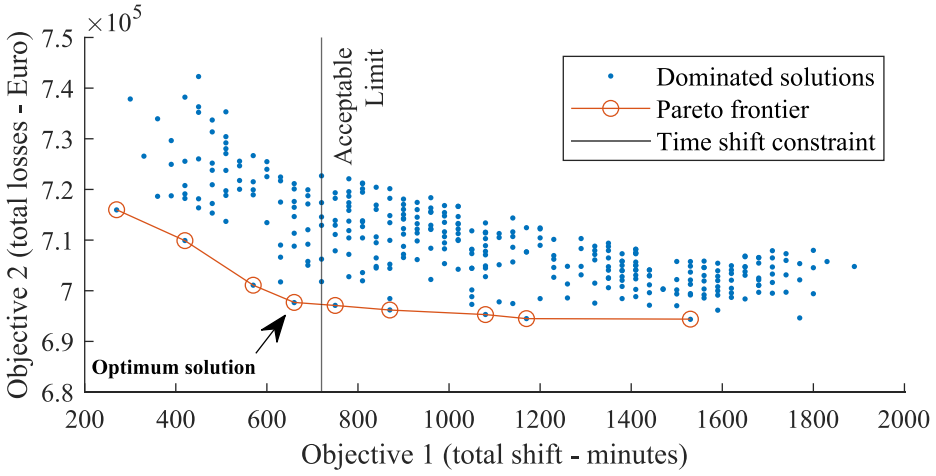

Fig. 14. Pareto frontier of the optimized FDTS for the 10% participation rate.

Table 3
Travel time improvement during peak hours for optimized departure time shift.

Scenarios	Recommended modification in departure time				Links travel time improvement				
	±30 min	±1h	±1.5 h	±2h	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 %

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.

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 Fig. 16, participants approximately experience travel time savings ranging from 4 to 10 min 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.

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.

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

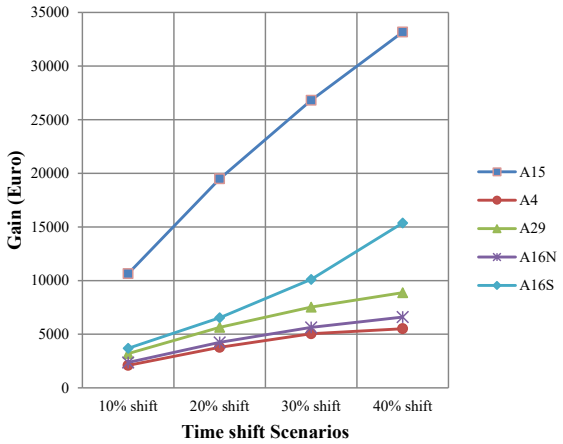

Fig. 15. Predicted social benefits after application of DTS policy.

Table 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	10,644	2258	3440	2549	4253	23,144 (3.2%)	7.3
20 %	6375	19,485	3776	5642	4238	6537	39,678 (5.5%)	6.2
30 %	9562	26,816	5048	7529	5637	10,112	55,141 (7.6%)	5.8
40 %	12,750	33,170	5511	8876	6598	15,373	69,528 (9.6%)	5.5

* Total monetary loss for the base-case scenario is 720,842 Euro.

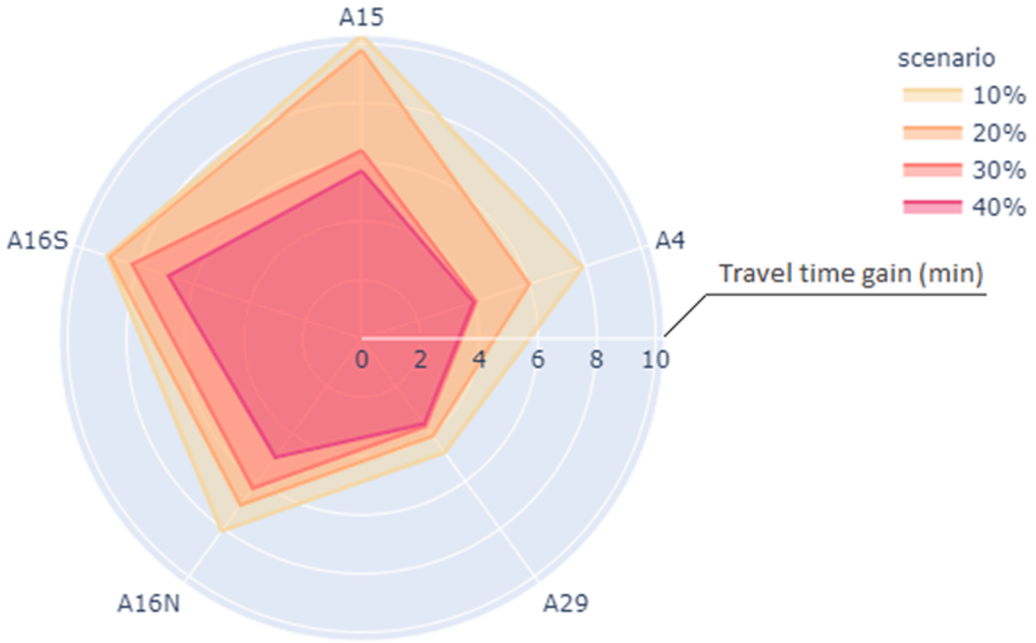

Fig. 16. Travel time savings for shifted trucks traveling during off-peak.

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 modeling 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 modeling 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 modeling 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.

CRediT authorship contribution statement

Ali Nadi: Conceptualization, Investigation, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Salil Sharma:** Conceptualization, Investigation, Formal analysis, Data curation. **J.W.C. van Lint:** Conceptualization, Writing – review & editing, Supervision. **Lóránt Tavasszy:** Conceptualization, Writing – review & editing, Supervision. **Maaike Snelder:** Conceptualization, Project administration, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was funded by the Netherlands Organization for Scientific Research (NWO) under project "ToGRIP-Grip on Freight Trips" with grant number 628.009.001, and supported by TKI Dinalog, Commit2data, Port of Rotterdam, SmartPort, Portbase, TLN, Deltalinqs, Rijkswaterstaat, and TNO. The authors would like to thank Portbase and NDW for providing us with port community system data and traffic data, respectively. We sincerely appreciate anonymous reviewers for all the thoughtful reviews and constructive remarks that helped us improve our paper.

References

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.
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.
Bliemer, M.C., van Amelsfort, D.H., 2010. Rewarding instead of charging road users: a model case study investigating effects on traffic conditions. European TransportTrasporti Europei 23–40.
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.
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 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.
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.
Hagan, M.T., Menhaj, M.B., 1994. Training feedforward networks with the Marquardt algorithm. IEEE transactions on Neural Networks 5, 989–993.
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., 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.
KIPF, T. N. & WELLING, M. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.
Kleff, A., Bräuer, C., Schulz, F., Buchhold, V., Baum, M., Wagner, D., 2017. Time-Dependent Route Planning for Truck Drivers. International Conference on Computational Logistics Springer, 110–126.

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.
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.
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.
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.
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.
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.
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.
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.
TLN 2019. Economische Wegwijzer 2019 nader verklaard The Dutch Association for Transport and Logistics.
Treiber, M., Helbing, D., 2003. An adaptive smoothing method for traffic state identification from incomplete information. Springer, Interface and Transport Dynamics.
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.
Lint, V.A.N., 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.
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.
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.
Yoshii, T., Ajisawa, S., Kuwahara, M., 1998. Impacts on traffic congestion by switching routes and shifting departure time of trips. 5th World Congress on Intelligent Transport Systems Proceedings.
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*.
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.