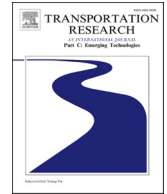




ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Transportation Research Part C

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

Short-term prediction of outbound truck traffic from the exchange of information in logistics hubs: A case study for the port of Rotterdam

Ali Nadi^{a,*}, Salil Sharma^a, Maaïke Snelder^{a,b}, Taoufik Bakri^b, Hans van Lint^a, Lóránt Tavasszy^a

^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:

Short-term prediction
Truck traffic
Container transport
Port community system
Artificial neural network

ABSTRACT

Short-term traffic prediction is an important 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 paper 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. The findings are relevant for traffic management agencies to help improve the efficiency and reliability of transport networks, in particular around major freight hubs.

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

* Corresponding author.

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

<https://doi.org/10.1016/j.trc.2021.103111>

Received 5 March 2020; Received in revised form 1 February 2021; Accepted 27 March 2021

Available online 18 April 2021

0968-090X/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

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, carriers, 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 paper is organized as follows: [Section 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 3](#) presents the data collection outlines with which we describe the data characteristics and preparation challenges. [Section 4](#) describes the ANN-based methodology we proposed to model the data. While [Section 5](#) denotes results and tests the prediction capability from a simulation of the model with some designed scenarios, [Section 6](#) at the end, offers concluding insights.

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.

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 the 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 15–60 min 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).

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

Table 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 |

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 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 modeling approach and the results.

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. [Fig. 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.

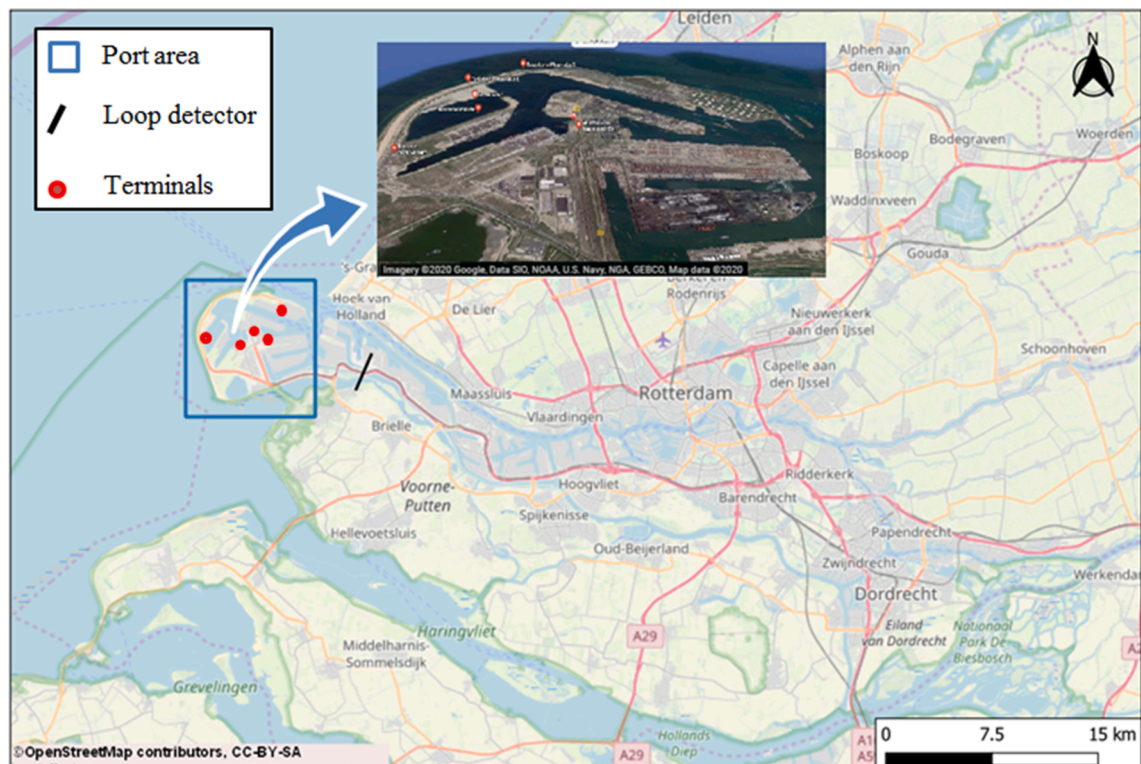


Fig. 1. Terminal and main loop-detector location considered in the port of Rotterdam.

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.

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 m 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 Fig. 1). These data were available at a resolution level of 1 min 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. Fig. 2 shows hourly containers that are ready to be picked up and Fig. 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, we provide our approach that we used to handle missing data before the neural networks were applied.

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.

Fig. 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 from 13:00 to 19:00. By comparison, container pickup schedules show that containers are mostly scheduled to be picked up from 11:00 to 17:00. 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 and 16:00. This is due to the working shift change at terminals, which leads to fewer truck departures during these hours.

Fig. 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.

Finally, Fig. 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

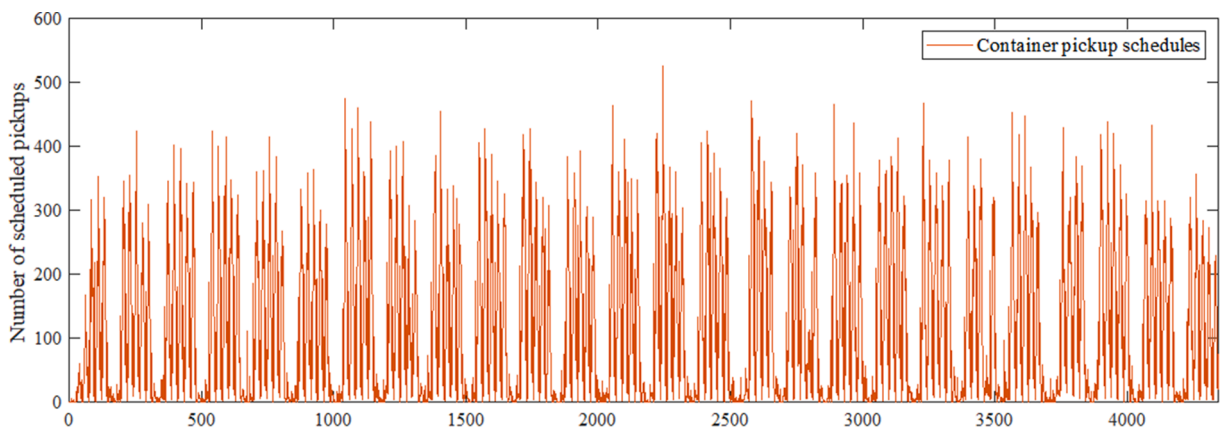


Fig. 2. Hourly containers ready to be picked up at the port of Rotterdam.

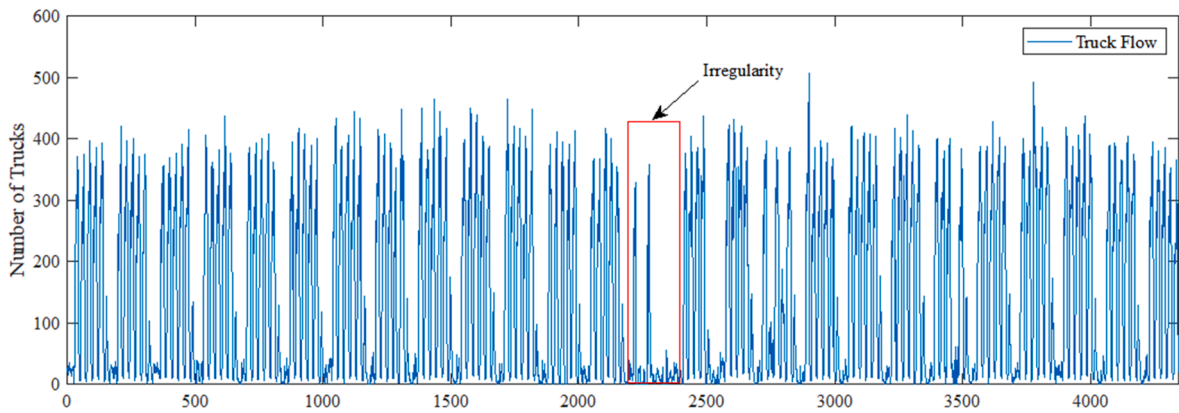


Fig. 3. Hourly truck traffic counts observed at the loop-detector on the A15.

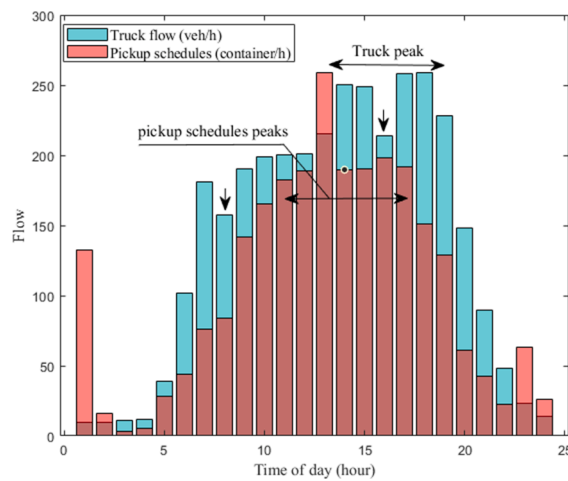


Fig. 4. Number of containers and number of trucks throughout the hours of the day.

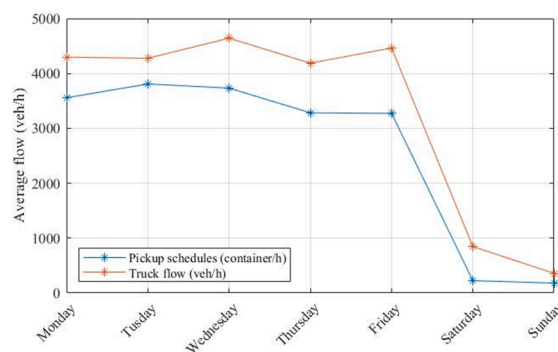


Fig. 5. Comparison of truck flows and scheduled container pickups during the week.

irregularity is also observable in Fig. 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.

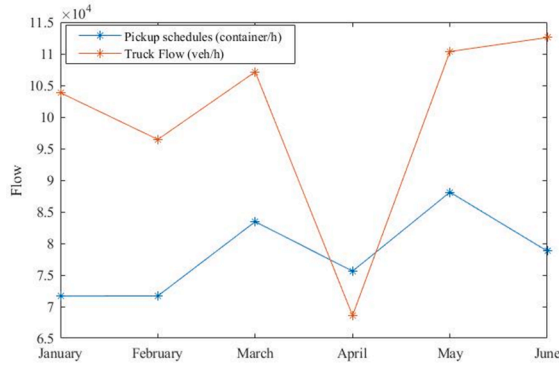


Fig. 6. Comparison of the truck flows and scheduled container pickups in the first six months of 2017.

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. Fig. 7 shows the building blocks of this 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:

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

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, 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.

4.1. 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 application in many disciplines, especially in modeling 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

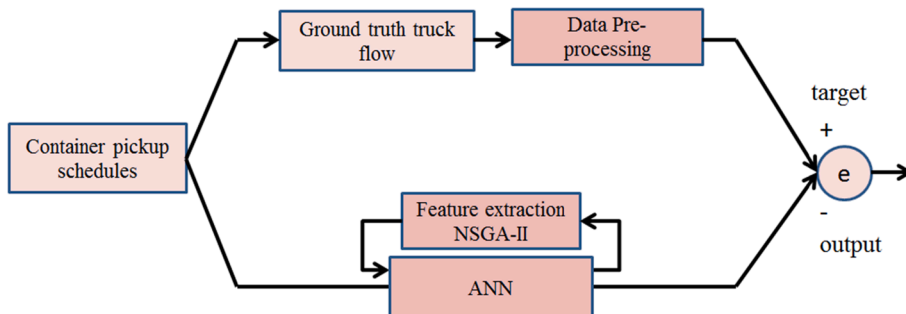


Fig. 7. Building blocks of our truck traffic prediction model.

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 the 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 4.2.1 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, see Eqn 2) is the most commonly used error function for estimating the model weights.

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

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, see Eqn 3), mean absolute error (MAE, see Eqn 4), mean absolute percentage error (MAPE, see Eqn 5), and the probability of absolute percentage error (PAPE, see Eqn 6) as the indicators.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (t_i - y_i)^2} \quad (3)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |t_i - y_i| \quad (4)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{t_i - y_i}{t_i} \right| \quad (5)$$

$$PAPE = Pr\left(\left| \frac{t_i - y_i}{t_i} \right| < 10\%\right) \quad (6)$$

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.

4.2. 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).

4.2.1. 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 (Eqn 8) and neurons (Eqn 9) is used in a way that the performance of the model (Eqn 7) does not decrease significantly.

$$z_1 = \frac{1}{N} \sum_i^N (t_i - y_i)^2 \quad (7)$$

$$z_2 = |\vec{X}| \quad (8)$$

$$z_3 = |\vec{W}| \quad (9)$$

where $|\vec{X}|$ and $|\vec{W}|$ 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 (Eqn 10).

$$z = \min(z_1 + \alpha z_2 + \beta z_3) \quad (10)$$

However one has to tune α and β correctly to get accurate results. To cope with this problem, the Akaike information criterion (AIC, see Eqn 11) is used as defined by Hurvich and Tsai (1993). Bayesian Information Criterion (BIC, see Eqn 12) 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.

$$AIC_c = n \log(MSE) + \frac{n+m}{1 - \frac{(m+2)}{n}} \quad (11)$$

$$BIC = n \log(MSE) + m \log(n) \quad (12)$$

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).

4.2.2. 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.

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.

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 Fig. 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. Fig. 8 shows the approximation of the Pareto front using the NSGA-II algorithm.

From the AIC point of view, Table 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 neurons 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 h 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.

5.2. In-sample validation

Fig. 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.

Fig. 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.

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. Fig. 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 $Pr(\mu - RMSE \leq e \leq \mu + RMSE) \approx 0.68$. 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.

All indicators show the high accuracy of the model despite the irregularities in the truck flow data. Figs. 11 and 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 Fig. 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.

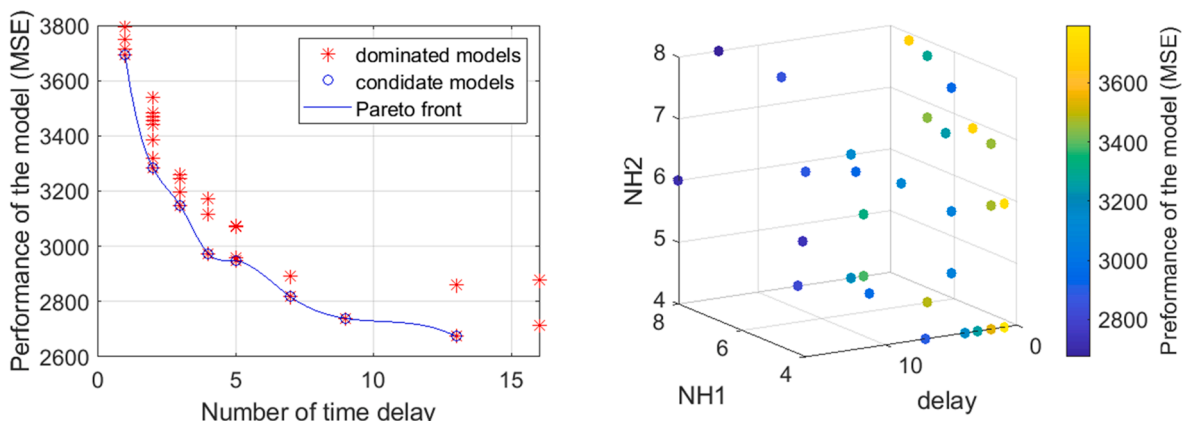


Fig. 8. Approximation of the models' Pareto frontier using the NSGA-II algorithm.

Table 2
Comparison of candidate models using their BIC and 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 |

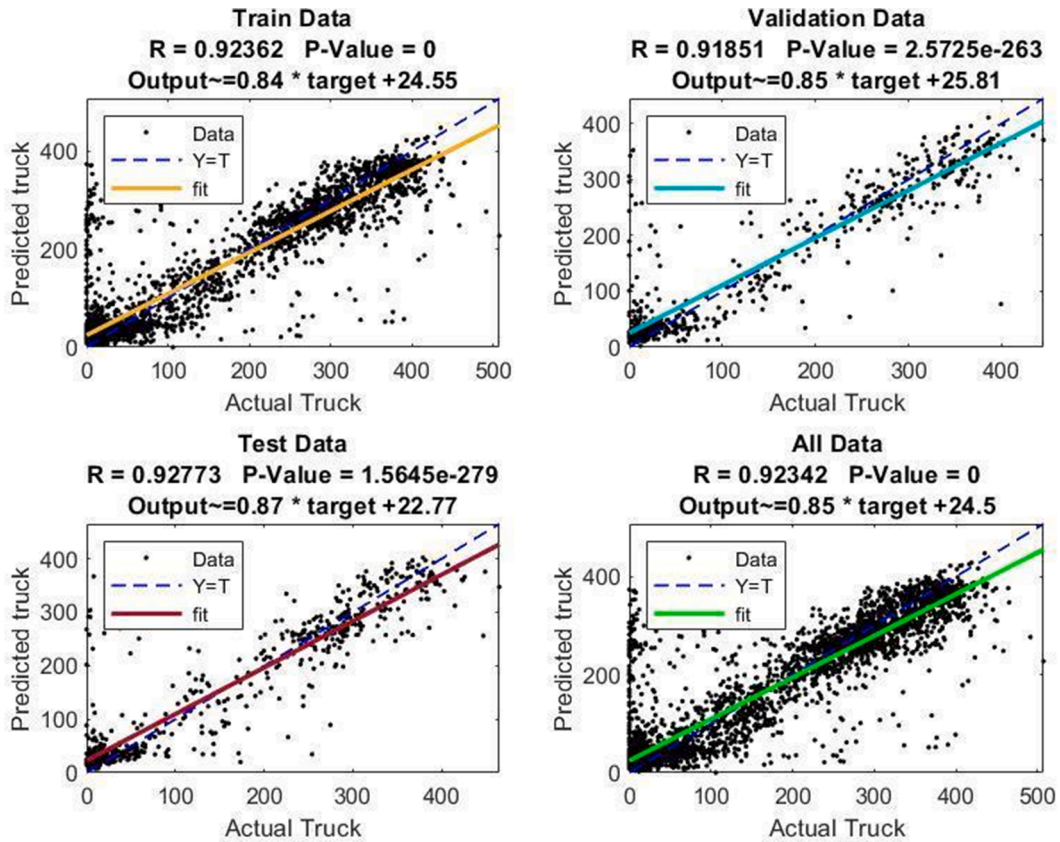


Fig. 9. Correlation coefficient and linear fit between modeled and observed volumes.

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. Fig. 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.

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

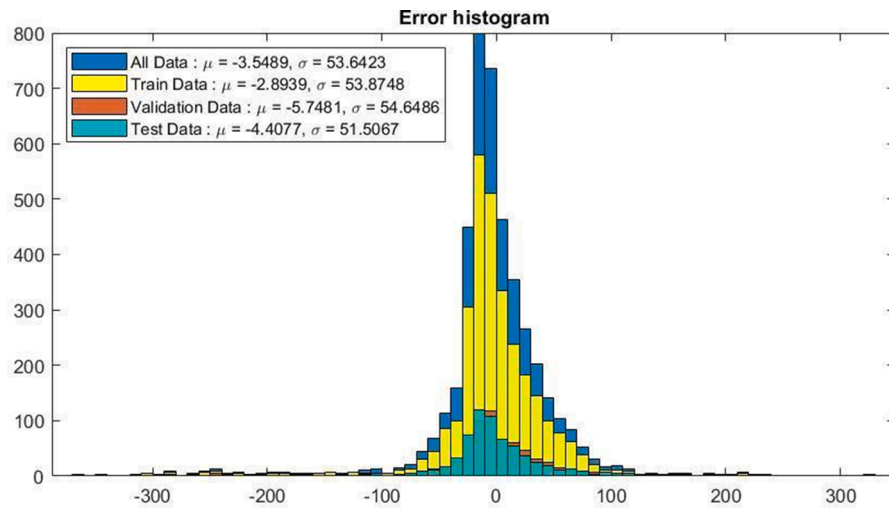


Fig. 10. Error histogram of the model for the three splits of the data.

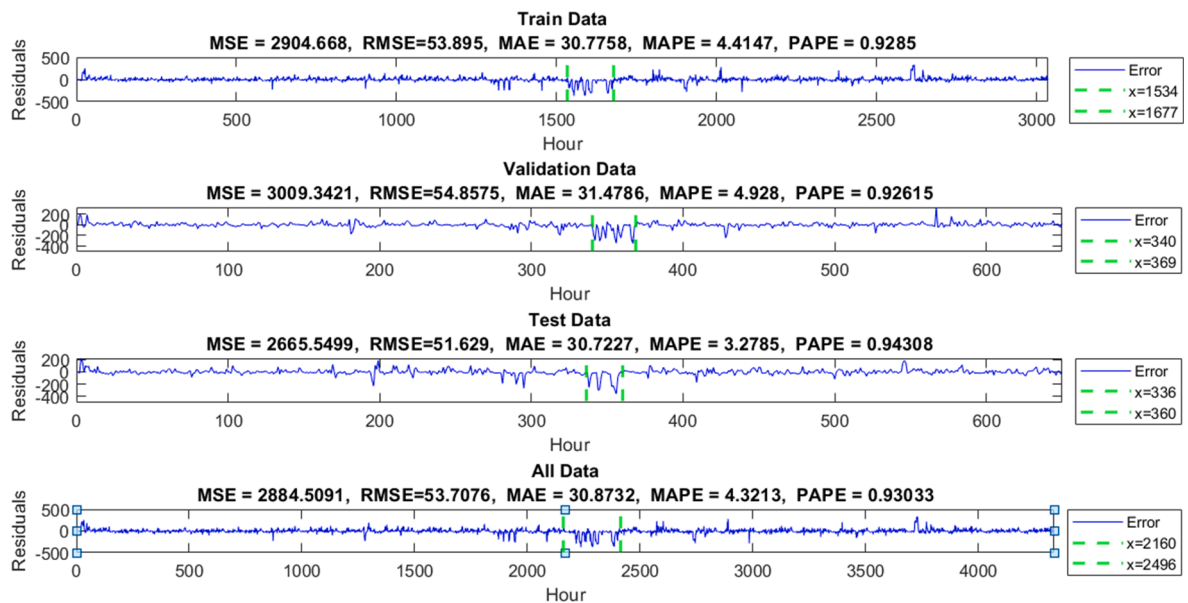


Fig. 11. Models residual and performance indicators for three splits of the data.

trial and error). This comparison is based on the predictions on the out-of-sample validation set.

The results in Table 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.

To compare for accuracy, Fig. 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.

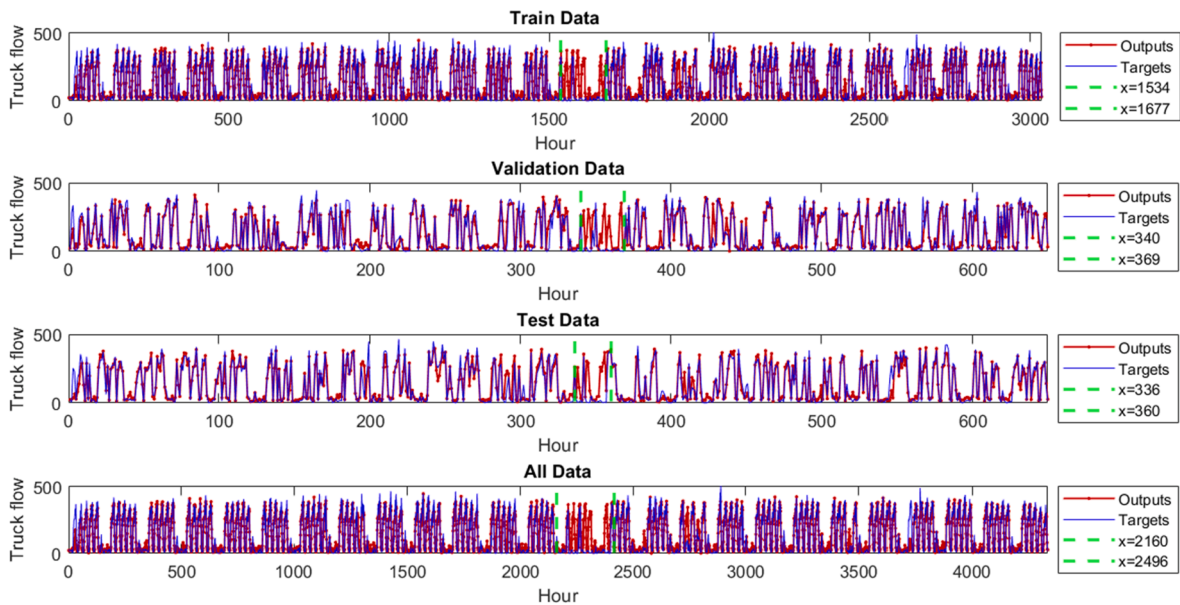


Fig. 12. Comparison of model fit with observed values for three splits of the data.

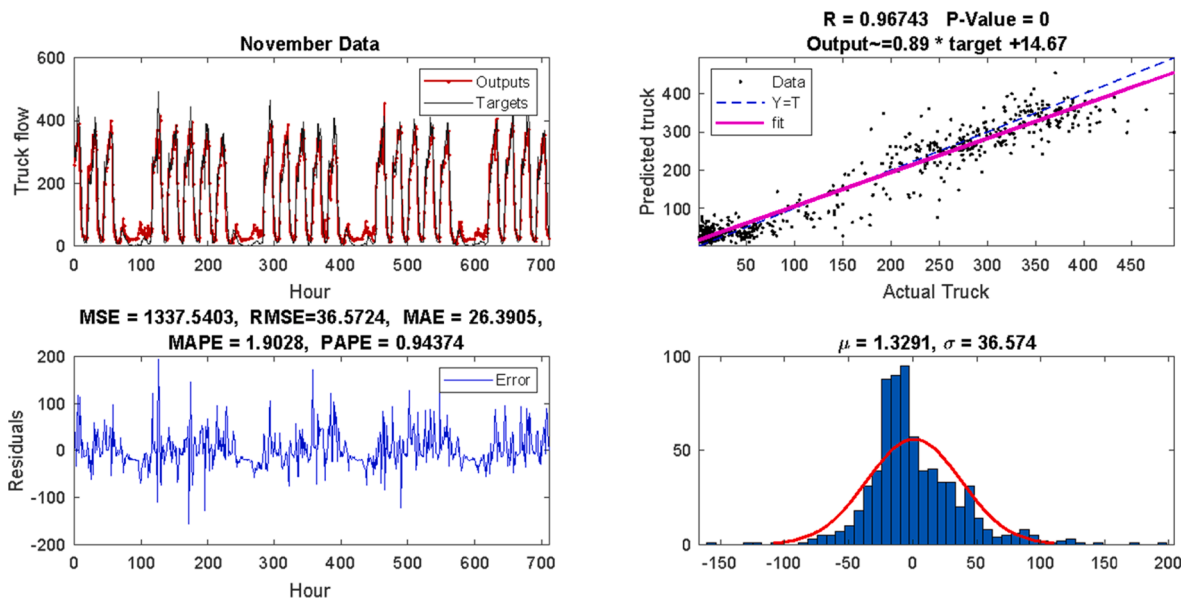


Fig. 13. Out-of-sample truck traffic forecast for November 2017.

5.5. Feature importance

As we saw in Table 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 Eq. (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. Fig. 15 shows the average score of each delay.

We can see, from Fig. 15, that $C_t, C_{t-1}, C_{t-2}, C_{t-9}$ are the most important features. Although C_t is in the same timestamp as T_t (see Eq. (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 3.3.

Table 3
Prediction temporal resolutions and methods comparison.

| Method Horizon | HA | | | LSE | | | BPNN | | | MLP-AFE | | | LSTM | | |
|-------------------|-------|------|------|-------|-------|------|------|------|------|---------|------|------|------|------|------|
| | 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 |

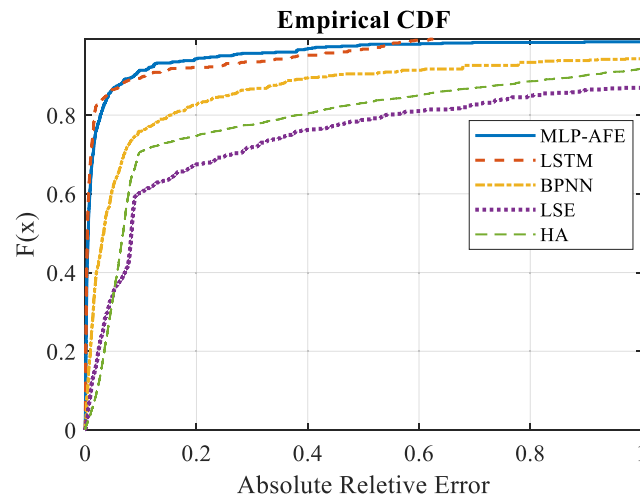


Fig. 14. methods comparison in 1-hour resolution with CDF of absolute relative errors.

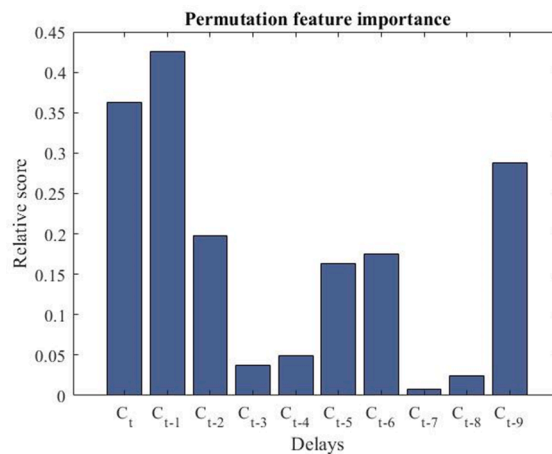


Fig. 15. Global explanation of the relative importance of each delay on truck volume predictions.

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.

Table 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 |

5.6.1. 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 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 Fig. 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.

5.6.2. 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. Fig. 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.

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.

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.

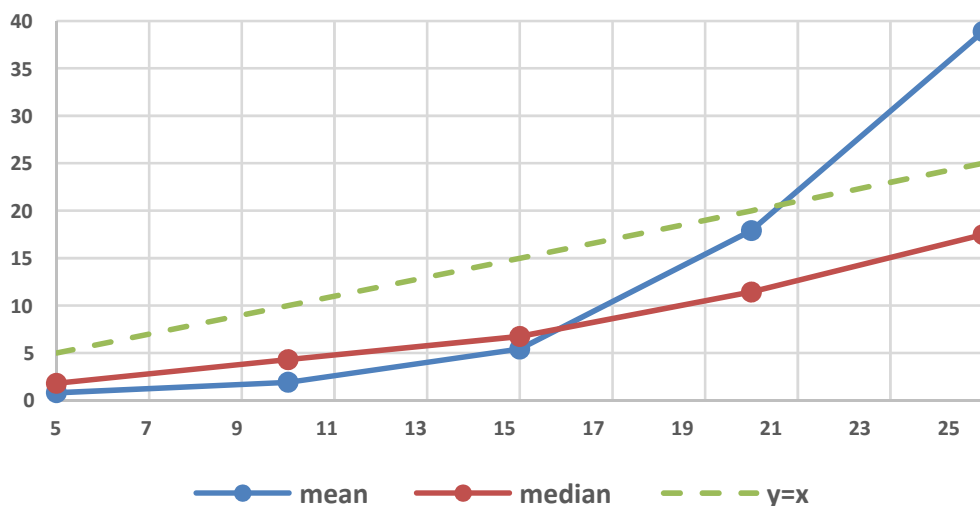


Fig. 16. Mean and median trends in the forecasted outbound truck in a growth scenario.

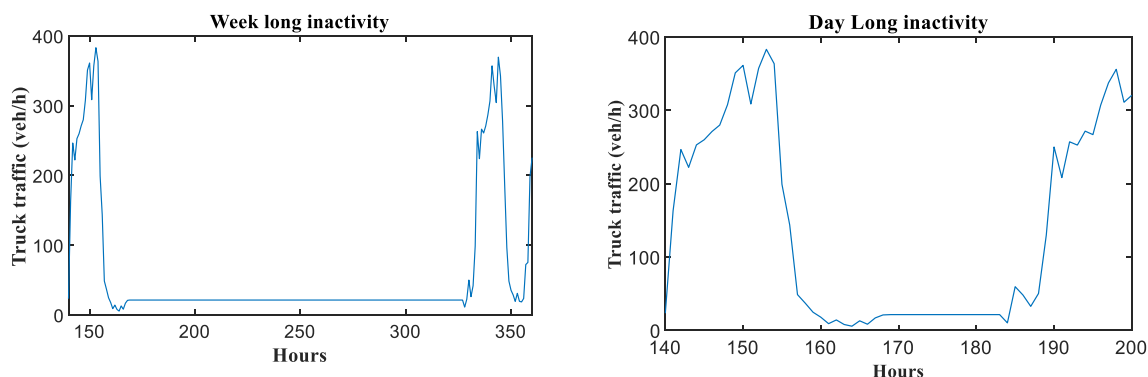


Fig. 17. Forecasted truck volume (a) weeklong inactivity and (b) day-long inactivity at seaport terminals.

- 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.

CRediT authorship contribution statement

Ali Nadi: Conceptualization, Investigation, Methodology, Software, Formal analysis, Writing - original draft, Writing - review & editing, Visualization. **Salil Sharma:** Conceptualization, Investigation, Software, Formal analysis, Writing - original draft. **Maaïke Snelder:** Conceptualization, Project administration, Writing - review & editing, Supervision. **Taoufik Bakri:** Conceptualization, Data curation. **Hans van Lint:** Conceptualization, Writing - review & editing, Supervision. **Lóránt Tavasszy:** Conceptualization, 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.

Acknowledgments

This work was supported by the Netherlands Organization for Scientific Research (NWO), TKI Dinalog, Commit2data, Port of Rotterdam, SmartPort, Portbase, TLN, Deltalinqs, Rijkswaterstaat, and TNO under project “ToGRIP-Grip on Freight Trips“. The authors would like to thank Portbase and NDW for providing us with port community system data and traffic data, respectively.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.trc.2021.103111>.

References

Al-Deek, H.M., 2002. Use of vessel freight data to forecast heavy truck movements at seaports. *Transp. Res. Rec.* 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. *Transp. Res. Rec.* 1719, 1–9.
- Banister, D., Stead, D., 2004. Impact of information and communications technology on transport. *Transp. Rev.* 24, 611–632.
- 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. *Transp. Res. Part C: Emerg. Technol.* 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. *Transp. Res. Part C: Emerg. Technol.* 118.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 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. *J. Adv. Transport.* 27, 227–249.
- di Febbraro, A., Sacco, N., Saeednia, M., 2016. An agent-based framework for cooperative planning of intermodal freight transport chains. *Transp. Res. Part C: Emerg. Technol.* 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. *Transp. Res. Part C: Emerg. Technol.* 108, 12–28.
- Fu, R., Zhang, Z., Li, L., 2016. Using LSTM and GRU neural network methods for traffic flow prediction. In: 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), 2016. IEEE, pp. 324–328.
- Giannopoulos, 2004. The application of information and communication technologies in transport. *Eur. J. Oper. Res.* 152, 302–320.
- Gonzalez-Calderon, C.A., Holguín-Veras, J., 2017. Entropy-based freight tour synthesis and the role of traffic count sampling. *Transp. Res. Part E: Logistics Transp. Rev.*
- Hassoun, M.H., 1995. *Fundamentals of Artificial Neural Networks*. MIT press.
- Heilig, L., Voß, S., 2017. Information systems in seaports: a categorization and overview. *Inf. Technol. Manage.* 18, 179–201.
- 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.
- Holguín-Veras, J., Patil, G.R., 2008. A multicommodity integrated freight origin–destination synthesis model. *Networks Spatial Econ.* 8, 309–326.
- Hong, W.C., 2011. Traffic flow forecasting by seasonal SVR with chaotic simulated annealing algorithm. *Neurocomputing* 74, 2096–2107.
- Huang, B., Buckley, B., Kechadi, T.-M., 2010. Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications. *Expert. Syst. Appl.* 37, 3638–3646.
- Hurvich, C.M., Tsai, C.L., 1993. A corrected Akaike information criterion for vector autoregressive model selection. *J. Time Ser. Anal.* 14, 271–279.
- Kim, I.Y., de Weck, O., 2006. Adaptive weighted sum method for multiobjective optimization: a new method for Pareto front generation. *Struct. Multidiscip. Optim.* 31, 105–116.
- 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. *Transp. Res. Rec.* 1873, 35–44.
- Kumar, S.V., 2017. Traffic flow prediction using Kalman filtering technique. *Procedia Eng.* 187, 582–587.
- Kumar, S.V., Vanajakshi, L., 2015. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *Eur. Transp. Res. Rev.* 7, 21.
- Lana, I., del Ser, J., Velez, M., Vlahogianni, E.I., 2018. Road traffic forecasting: recent advances and new challenges. *IEEE Intell. Transp. Syst. Mag.* 10, 93–109.
- Li, B., Tan, K. W., Tran, K.T., 2016. Traffic simulation model for port planning and congestion prevention. In: *Proceedings of the 2016 Winter Simulation Conference*, 2016. IEEE Press, pp. 2382–2393.
- Liedtke, G., 2009. Principles of micro-behavior commodity transport modeling. *Transp. Res. Part E: Logist. Transp. Rev.* 45, 795–809.
- Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.-Y., 2014. Traffic flow prediction with big data: a deep learning approach. *IEEE Trans. Intell. Transp. Syst.* 16, 865–873.
- 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. *Transp. Res. Part C: Emerg. Technol.* 63, 182–194.
- Nadi Najafabadi, A., Sharma, S., Snelder, M., Tavasszy, L., Van Lint, J., 2019. Analyzing the role of seaport operations in generating inbound/outbound truck traffic demand and its implications on traffic system. In: *TRAIL PhD Congress 2019*, 2019.
- Nicholson, H., Swann, C., 1974. The prediction of traffic flow volumes based on spectral analysis. *Transp. Res.* 8, 533–538.
- Polson, N.G., Sokolov, V.O., 2017. Deep learning for short-term traffic flow prediction. *Transp. Res. Part C: Emerg. Technol.* 79, 1–17.
- Poonia, P., Jain, V., Kumar, A., 2018. Short term traffic flow prediction methodologies: a review. *Mody University Int. J. Computing Eng. Res.* 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. *J. Operat. Res. Soc.* 46, 1090–1101.
- PORT-OF-ROTTERDAM, 2019. Continuing container growth pushes throughput at port of Rotterdam to a new high.
- Rajabzadeh, Y., Rezaie, A.H., Amindavar, H., 2017. Short-term traffic flow prediction using time-varying Vasicek model. *Transp. Res. Part C: Emerg. Technol.* 74, 168–181.
- Sánchez-Díaz, I., Holguín-Veras, J., Ban, X.J., 2015. A time-dependent freight tour synthesis model. *Transp. Res. Part B: Methodol.* 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. *Transp. Res. Rec.* 1906, 113–120.
- Tian, Y., Pan, L., 2015. Predicting short-term traffic flow by long short-term memory recurrent neural network. In: 2015 IEEE international conference on smart city/SocialCom/SustainCom (SmartCity), 2015. IEEE, pp. 153–158.
- Van Hinsbergen, C., Van Lint, J., Sanders, F., 2007. Short term traffic prediction models. In: *Proceedings of the 14th World Congress on Intelligent Transport Systems (ITS)*, Held Beijing, OCTOBER 2007, 2007.
- van Lint, J., van Hinsbergen, C., 2012. Short-term traffic and travel time prediction models. *Artificial Intelligence Appl. Critical Transp. Issues* 22, 22–41.
- Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2014. Short-term traffic forecasting: where we are and where we're going. *Transp. Res. Part C: Emerg. Technol.* 43, 3–19.
- Williams, B.M., Hoel, L.A., 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. *J. Transp. Eng.* 129, 664–672.
- Wisetjindawat, W., Yamamoto, K., Marchal, F., 2012. A commodity distribution model for a multi-agent freight system. *Procedia-Social Behav. Sci.* 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. *Transp. Res. Part C: Emerg. Technol.* 90, 166–180.
- Xie, Y., Huynh, N., 2010. Kernel-based machine learning methods for modeling daily truck volume at seaport terminals. In: 51st Annual Transportation Research Forum, Arlington, Virginia, March 11–13, 2010, 2010. Transportation Research Forum.
- Xu, D., Wei, C., Peng, P., Xuan, Q., Guo, H., 2020. GE-GAN: A novel deep learning framework for road traffic state estimation. *Transp. Res. Part C: Emerg. Technol.* 117, 102635.
- Zhang, Y., Ye, Z., 2008. Short-term traffic flow forecasting using fuzzy logic system methods. *J. Intelligent Transp. Syst.* 12, 102–112.
- 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 Trans. Intell. Transp. Syst.*