Vrije Universiteit Amsterdam Centrum Wiskunde & Informatica

Master Thesis

Analysis and Short-Term Forecasting of Traffic Intensity: Exploring the Impact of Road Maintenance

Author: Eline A. Belt

1st supervisor:	Prof. dr. Rob van der Mei	
daily supervisors:	Dr. Elenna Dugundji, Dr. Thomas Koch	(CWI, MIT)
2nd reader:	Dr. René Bekker	

A thesis submitted in fulfillment of the requirements for the VU Master of Science degree in Business Analytics

July 12, 2023

To my son Adam, your presence has added depth and purpose to my academic journey, making every hurdle worth overcoming; this thesis is dedicated to you with love.

Preface

This thesis has been written for the completion of the Master of Science degree in Business Analytics at the Vrije Universiteit (VU) Amsterdam. The research for this thesis was conducted during an internship at Centrum Wiskunde & Informatica (CWI) in the Stochastics department. This research was part of the FTMAAS project and was done for Rijkswaterstaat, it studies the impact of maintenance on roads and whether the traffic intensity can be accurately predicted during these periods of maintenance.

Acknowledgements

First of all, I would like to thank the company supervisors from CWI, dr. Elenna Dugundji and dr. Thomas Koch for their guidance, expertise and advice throughout my master project. I would also like to thank prof. dr. Rob van der Mei and dr. René Bekker for being my first supervisor and the second reader, respectively.

I would also like to thank my husband, Jawad, and my parents for their support throughout this project. Without them, I could have never finished it.

Abstract

Context. Road maintenance is an important part of maintaining the quality of public roads. Regular maintenance is important to ensure that the roads are safe, it can improve traffic flow, it will increase the lifespan of the road, and there are economic benefits as well.

Goal. The first goal of this research is to study the impact that road maintenance has on traffic intensity, with a focus on the impact it has on freight traffic. The second is to forecast the traffic intensity during maintenance. And the third is to implement a model that incorporates a graph neural network because it can utilise the data from multiple sensors at once.

Method. To study the impact of road maintenance on traffic intensity, hypothesis tests were performed. To predict the traffic intensity, a seasonal naive (baseline) model, Holt-Winters exponential smoothing, a Spatial Temporal Graph Neural Network (GNN), and a Transformer were implemented.

Results. Hypothesis tests confirmed that there was a significant change in traffic intensity during maintenance compared to the traffic intensity before maintenance. Comparing the models based on their RMSE and MAE showed that XGBoost had the lowest errors for the 5-minute forecast for passenger traffic and for the hour-ahead forecast for both types of traffic. The GNN had the lowest errors for the 5-minute forecast for freight traffic.

Conclusions. This study concludes that there is a significant change in traffic intensity during maintenance. On the detour there is a significant increase in both types of traffic and on the advisory route there is a significant increase in freight traffic. Rijkswaterstaat can use this information to plan these advisory routes to facilitate freight traffic. The forecasts of the XGBoost models showed that they can accurately predict traffic intensity, even during maintenance. The GNN had lower errors than the other models, except for XGBoost. However, it had lower errors than XGBoost for the 5-minute forecast for freight traffic.

Contents

1

Introduction

1.1 Context

Road maintenance is an important part of maintaining the quality of public roads. Most people only consciously experience the nuisance they cause, but they do not realise all the benefits that maintenance can bring. Regular maintenance of roads is important for several reasons:

Safety

Regular maintenance to repair roads ensures that they are safe for use by vehicles. This can help prevent accidents caused by potholes and other types of damage.

Smooth Traffic Flow

Maintenance of roads can improve traffic flow by repairing damages, this can help reduce congestion and delays on the road, making it easier for people to reach their destination.

Increased Lifespan

Regular maintenance can extend the lifespan of the road, preventing it from deteriorating and requiring more extensive repairs later on. This can also help save money in the long run by avoiding more expensive maintenance projects.

Economic Benefits

Roads are essential for the transportation of goods and services, and maintenance can help keep the road network in good condition. This can help support (local) businesses by providing reliable access to customers and suppliers.

Overall, road works are important because they help maintain and improve the road network, making it safer and more efficient for everyone who uses it.

1.2 Problem Statement

Rijkswaterstaat has to perform road maintenance to ensure the quality of roads and to keep up with the demand of road users. However, they would like to perform maintenance with minimal disruption to the traffic flow. Unfortunately, it is not possible to have no disruption to traffic when road maintenance is performed, but by understanding the impact that these road works have on different types of traffic, it will be possible to decrease its impact and improve the experience that road users have with maintenance.

1.3 Research Goals and Questions

This thesis aims to research the impact that road maintenance has on traffic intensity, with a specific focus on the impact of maintenance on freight traffic. Another goal of this thesis is to predict the traffic intensity during maintenance and research whether models are robust to changes in traffic flow because of the maintenance. The final goal is to determine whether a model that incorporates a graph neural network can utilise the data from multiple sensors to enhance forecasting accuracy in comparison to models that rely on data from only one sensor. For these goals, the following research questions were made:

- Does road maintenance have a significant impact on the intensity of freight traffic?
- Can models accurately forecast the traffic intensity during maintenance?
- Can a model that incorporates a graph neural network and utilises data from multiple sensors at once forecast traffic intensity better than other models that use data from only one sensor?

1.4 Business Info

1.4.1 Rijkswaterstaat

Rijkswaterstaat is the institution responsible for maintaining and developing roads and waterways in the Netherlands. Rijkswaterstaat schedules road maintenance and gives advice about alternative routes in case of road works. The department ‘West-Nederland

Noord' (WNN) of Rijkswaterstaat is responsible for the construction and maintenance of highways in the province of Noord-Holland. Until 2030 WNN is facing a lot of highway maintenance, and they will also renovate and replace multiple bridges, tunnels, roads, sluices and the largest pumping station in Europe in IJmuiden. They want to obtain insight into the impact of maintenance on the traffic intensity on the road network with a focus on minimizing its impact on freight traffic.

1.4.2 Centrum Wiskunde & Informatica

Centrum Wiskunde & Informatica (CWI) is the national research institute for mathematics and computer science in the Netherlands. Their main goal is to generate new ideas that have positive impacts on society, the economy, and various scientific fields. In the coming years, their focus is on the following four areas of research: Algorithms, Data Intelligent systems, Cryptography Security, and Quantum Computing.

1.5 Thesis Outline

This paper is set up in the following way: in Section 2, literature related to this paper will be discussed. Then in Section 3 the highway maintenance project will be discussed and the traffic intensity data will be analysed. The methods that will be applied in this paper are explained in Section 4 and their implementation details and results are shown in Section 5 and Section 6, respectively. Then the paper will conclude a discussion of the findings in Section 7 and with a conclusion in Section 8.

2

Literature Review

In this section, literature about traffic forecasting will be discussed. Various types of models have been used for traffic forecasting purposes. Traditionally statistical (time series) models have been used. However, in recent years, different types of neural networks have also been applied to this task and are showing promising results. Additionally, hybrid models that combine statistical models and neural networks have been explored. Some studies have utilized other models, such as XGBoost (eXtreme Gradient Boosting) and clustering methods for forecasting.

2.1 Statistical Models

For time series data, exponential smoothing methods are well-known and have been extensively applied to different types of time series data (4, 10). There are multiple types of exponential smoothing methods. Specifically, for time series that have multiple seasonal components, Holt-Winters exponential smoothing is most often used. Holt-Winters exponential smoothing can be used as a baseline model for more complex models, as papers often compare their newly proposed models to Holt-Winters (22, 29).

2.2 XGBoost

XGBoost is an algorithm that uses gradient-boosted decision trees for machine learning tasks. It has been successfully applied to traffic prediction by different researchers (1, 9).

In (18) XGBoost is used to predict hourly traffic intensity. This study also investigates the effect of different types of regularization on the performance of XGBoost. The performance of the XGBoost model was compared with other models; a Support Vector Machine, k-Nearest Neighbors, Decision Trees, Random Forest, Gradient Boosting Decision Trees, Fully Connected Deep Neural Network, and a Long Short Term Memory network. XGBoost shows the best performance according to multiple performance measures, as well as being able to better predict traffic intensity during days that have different traffic patterns, such as during holidays or extreme weather conditions.

In (37), a hybrid model that combines XGBoost and Harris Hawk optimization is proposed for a multi-step prediction model. The paper concludes that this model has high accuracy and stability when applied to a dataset.

2.3 Neural Networks

Various types of neural networks can be used for traffic intensity forecasting. In this literature review, two types, namely graph neural networks and transformers will be discussed.

2.3.1 Graph Neural Networks

Graph neural networks (GNNs) are used in traffic intensity prediction because they can leverage the spatial information provided by the locations of sensors on the road network, which can increase the forecasting capabilities of these networks compared to other types of neural networks. Most Graph Neural Networks consist of multiple layers, one of which is a graph neural network, the other layers are other types of neural networks, often recurrent neural networks (5, 11, 23) or convolutional neural networks (8, 17, 33) are used in the network architecture as well. Often, a multilayer perceptron is used as the final layer to output the predictions.

For instance, in (28) a so-called spatial temporal graph neural network (STGNN) is proposed to predict the future traffic flow. The network consists of a spatial graph neural network layer to model the spatial dependencies in the data, a layer incorporating gated recurrent units to model the short-term temporal dependencies, a transformer layer to model long-term temporal dependencies, and a multi-layer feed-forward network as the final layer for outputting predictions. To incorporate the spatial information into the graph neural network, the pair-wise relations between sensors are modeled using a relation matrix based on latent positional representations of the sensors. This STGNN outperforms the baseline models on two well-known traffic datasets, the METR-LA dataset, and the PEMS-BAY dataset.

In another study, (36) proposes a Spatio-Temporal Graph Convolutional Network (STGCN) which consists of two spatio-temporal convolutional blocks (ST-Conv blocks) and a multi-layer perceptron as the output layer. These ST-Conv blocks contain temporal gated convolution layers that derive the most useful temporal features from the data and a spatial graph convolution layer to obtain spatial features. In this paper, the spatial information is modeled using a weighted adjacency matrix, based on the distance between sensors in the network.

A model called Graph WaveNet is proposed in (34) a, which is inspired by WaveNet (24). The authors utilize stacked dilated convolutions from WaveNet, allowing the Graph WaveNet to efficiently process long temporal sequences. Additionally, they introduce a self-adaptive adjacency matrix, which is better at representing spatial data compared to a static adjacency matrix. Based on this paper, (25) proposes modifications to Graph WaveNet to enhance its performance.

2.3.2 Transformers

Transformer models are well-known for their successful application in Natural Language Processing. Due to their ability to process sequential data, they can also be applied to time series data.

In (3), a Traffic Transformer is proposed. This model utilizes a new type of positional encoding to capture the temporal dependencies in time series data. Both the encoder and decoder of the Traffic Transformer include a graph convolutional network (GCN) block to model spatial dependencies in the data.

In (15) trafficBERT is proposed, this model is based on the BERT (bidirectional encoder representations from transformers) language model. Instead of using NLP embeddings, trafficBERT incorporates a weekday embedding. The model is pre-trained using are large dataset to enhance its predictive abilities for traffic flows. The study demonstrates that trafficBERT outperforms baseline models, like ARIMA on benchmark datasets METR-LA, PeMS-L, and PeMS-Bay. This study not only highlights the effectiveness of transformer models for predicting traffic intensity but also shows the benefits of transfer learning in improving the forecasting ability of a model.

2.4 Hybrid Models

There is a wide range of papers discussing different types of hybrid models, (2, 12, 27, 32), A few notable examples are highlighted in this section.

In (21) NN-ARIMA is proposed, in which a Multi-Layer Perceptron (MLP) is used to identify the pattern of the traffic flow and ARIMA is used to process the residuals of the MLP to identify location-specific traffic features.

Another hybrid approach is used by (35), where an ARIMA-BPNN (Back Propagation Neural Network) optimized using Simulated Annealing (SA) is proposed. This model improves traffic prediction accuracy by leveraging the linearity of ARIMA and the non-linearity of the BPNN.

Statistical methods and transformers can also be combined into a single model. An example of this is ETSformer: Exponential Smoothing Transformers for Time-series Forecasting (31), which builds upon the basic transformer model by incorporating exponential smoothing attention (ESA) and frequency attention (FA) instead of the self-attention mechanism used in the original transformer model.

3

Case Study

3.1 Road Maintenance

In May of 2021, Rijkswaterstaat scheduled maintenance on the A4 Highway between interchanges 'De Hoek' and 'Burgerveen', the location can be seen in Figure 3.1. This maintenance was part of 'Groot variabel onderhoud 2021' in the region West Nederland Noord. As part of the maintenance the top layer of asphalt, and in some areas also the intermediate layer, will be replaced. Other maintenance activities are repairing the crash barriers, vehicle detection loops, and the road verge.

The maintenance is divided into two phases, which will be explained in Section 3.1.1 and 3.1.2 below. Phase 1 took place during the weekend, starting on Friday the 7th of May 21:00 and ending on Monday the 10th of May 05:00. Phase 2 took place after that, starting when phase 1 ended and lasted until Monday the 31st of May 05:00.

3.1.1 Phase 1

During Phase 1 of the maintenance project, a 2-0 system is applied to interchange Burgerveen, which means that in one direction of the highway 2 lanes are open, and in the other direction no lanes are open. In this case, no lanes will be open on the A4 starting from the interchange Burgerveen in the direction of Den Haag/Rotterdam. The traffic going in this direction will be directed to the A44 where advisory routes are indicated. For traffic coming from Den Haag/Rotterdam, 2 lanes are open, but they will be redirected from the main carriageway left (HRL) to the main carriageway right (HRR) and back to the HRL after interchange Burgerveen. This is shown in Figure 3.2.

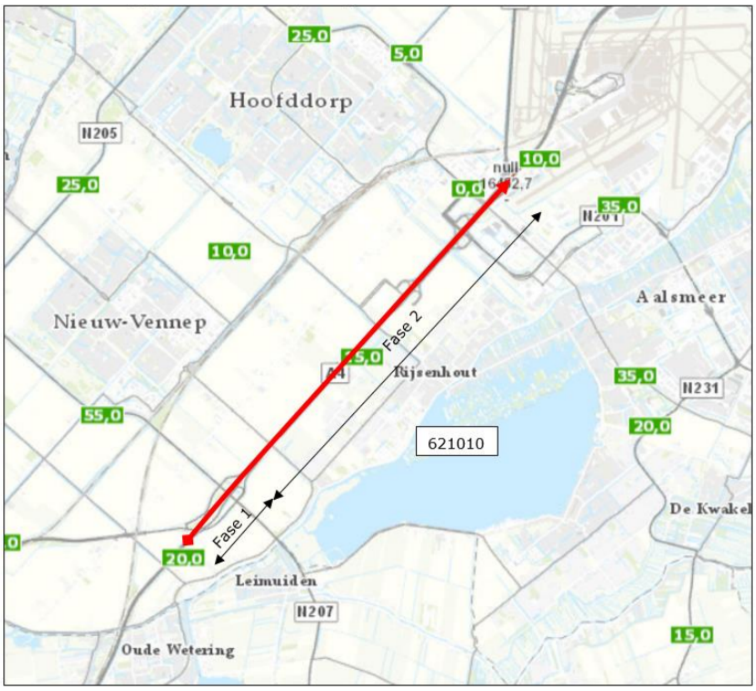
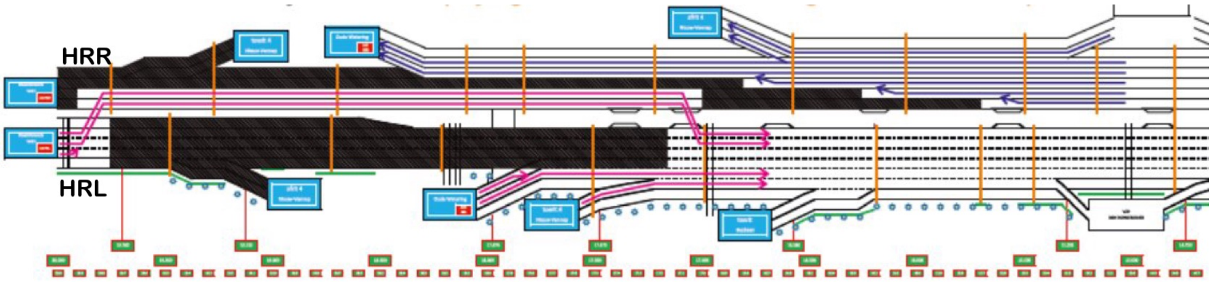


Figure 3.1: Map showing the location of Phase 1 and Phase 2 of the maintenance.

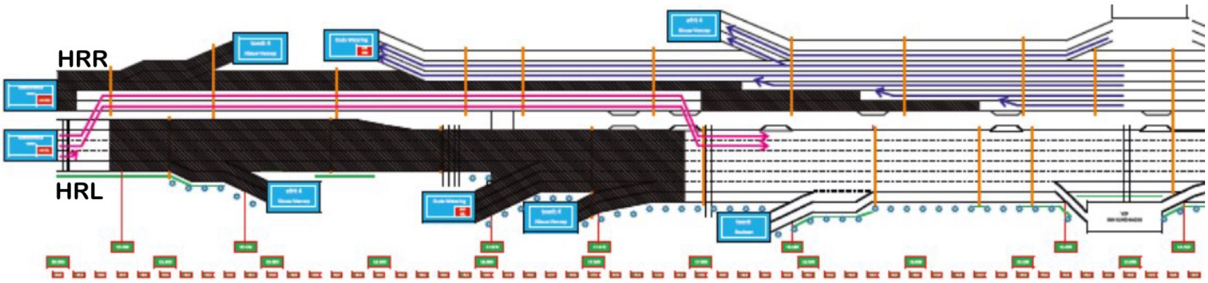
Phase 1 is split into two sub-phases, Phase 1a and Phase 1b. Phase 1a lasts for the duration of Phase 1, except for the night of Saturday May 8th 21:00 until Sunday May 9th 09:00, when Phase 1b is in effect. During Phase 1a the situation is as described above and traffic on the A44 from Den Haag/Wassenaar can continue as normal onto the A4 HRR to Amsterdam and the same holds for traffic entering the highway using the on-ramp (toerit Nieuw Vennep) at the interchange, this is shown in Figure 3.2a. During Phase 1b, shown in Figure 3.2b, the situation is again as described above, however, now the traffic coming from the A44 from Den Haag/Wassenaar and traffic wanting to enter the A4 using the on-ramp (toerit Nieuw Vennep) at the interchange cannot go onto the A4, they have to use a detour instead, which is shown in Figure A.1 in Appendix A.

3.1.2 Phase 2

During Phase 2, a 6-2 system is applied between interchanges Burgerveen and de Hoek. On the HRR, towards Den Haag, four lanes are open to traffic that normally drives on the HRR towards HRR, but two lanes will be used for traffic going in the opposite direction towards Amsterdam. On the other carriageway, the HRL towards Amsterdam, two lanes will be open to traffic during the day. The other lanes on the HRL are closed for maintenance.



(a) Phase 1a



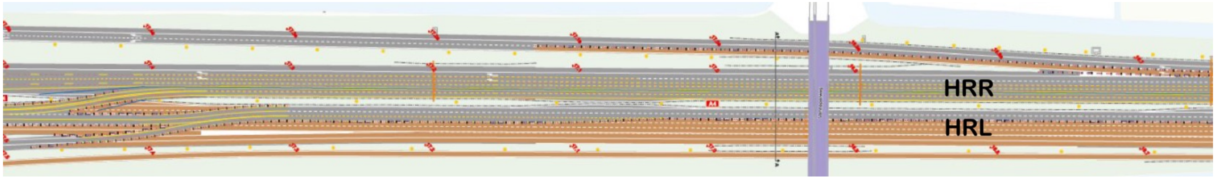
(b) Phase 1b

Figure 3.2: 2-0 system during Phase 1 of the maintenance.

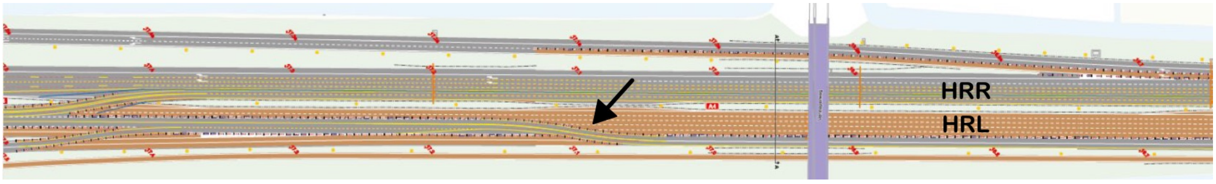
However, at night, only one lane will be open for traffic. During phase 2 detours are not necessary, but there are advisory routes traffic can follow using the A2, A12, and A20.

Because the HRR consists of five lanes and now has to facilitate six lanes, four of the lanes are smaller than usual and no freight traffic is allowed on those. Freight traffic going towards Den Haag have to use the two rightmost (slowest) lanes on the HRR and freight traffic going towards Amsterdam has to use the two lanes that are open on the other carriageway, the HRL. The maximum speed for all lanes is decreased from 100 km/h to 70 km/h.

Phase 2 also has two sub-phases, however no specific dates were given for these phases. There is a minor difference between Phase 2a and Phase 2b, which is a change of the open lanes on the HRL. In phase 2a, traffic is directed towards the two leftmost lanes, and in phase 2b, traffic is directed to the two rightmost lanes to facilitate the maintenance of different parts of the carriageway. Phase 2a and Phase 2b are shown in Figure 3.3, where the arrow in Figure 3.3b indicates the change in open lanes.



(a) Phase 2a



(b) Phase 2b

Figure 3.3: 6-2 system during Phase 2 of the maintenance.

3.2 Data Analysis

3.2.1 Data Acquisition

The data concerning the traffic intensity of Dutch roads and highways can be downloaded from the Dexter application[1] by Nationaal Dataportaal Wegverkeer (NDW)[2]. Sensors are used to measure the traffic intensity and other traffic information. From Dexter, sensor locations can be selected for which the traffic data can be downloaded.

For this case study, 52 sensor locations were chosen along the A4, A44, and A5 because the roadworks took place on the A4, and the A44 and A5 are directly connected to the A4 highway. And 15 locations were chosen along the N205, N207, N208, and A9 as these roads were used in a detour during Phase 1 of the maintenance.

3.2.2 Technical Exclusions

Of these 67 locations, multiple locations were marked as exclusions in the dataset due to anomalies in the measurements. One sensor is excluded because it's a deviating location since it is in the middle of a 'weefvak'[3]. Another is located on a hard shoulder (vluchtstrook, in Dutch) and has deviating traffic patterns because this part of the road is only opened when there's a lot of traffic on the road (usually during rush hour), that's why this location is removed from the dataset as well. Three more locations were excluded because of

[1] https://dexter.ndwcloud.nu/
[2] https://www.ndw.nu/
[3] https://nl.wikipedia.org/wiki/Weefvak

damages to the sensors after roadworks, which left the sensors with no measurements for the intensity and speed.

3.2.3 Data Description

The dataset contains multiple features which are either characteristics of the sensor, or measurements made by the sensor. Table 3.1 shows the features that were used in this case study. Of these features, the most important ones are '*start_meetperiode*', '*gem_intensiteit*', '*voertuigcategorie*', '*start_locatie_latitude*', and '*start_locatie_longitude*'. These will be discussed more in-depth.

Table 3.1: Features from the dataset used for the analysis in this case study.

Column Name	Description
id_meetlocatie	ID of the sensor
start_meetperiode	Start time of measurements
eind_meetperiode	End time of measurements
incomplete_waarnemingen_intensiteit	Number of incomplete intensity measurements
incomplete_waarnemingen_snelheid	Number of incomplete speed measurements
waarnemingen_intensiteit	Number of intensity observations (vehicles)
waarnemingen_snelheid	Number of vehicles speed was observed for
data_error_snelheid	Indicator of error in speed measurement
data_error_intensiteit	Indicator of error in intensity measurement
gem_intensiteit	Average intensity
gem_snelheid	Average speed
totaal_aantal_rijstroken	Total number of lanes
nauwkeurigheid	Accuracy of sensor measurements
voertuigcategorie	Vehicle category
start_locatie_latitude	Latitude of sensor location
start_locatie_longitude	Longitude of sensor location
naam_meetlocatie	Description of sensor location (including road name and mile marker)

The most important feature is *gem_intensiteit* because this will be used as explanatory variable in the forecasting models. *gem_intensiteit* is the measured traffic intensity, expressed in vehicles per hour. *gem_snelheid* is the average speed that the vehicles were driving. *start_meetperiode* indicates the date and time that a measurement was taken, and because the granularity of the dataset is 5 minutes, *eind_meetperiode* is always 5 minutes later than *start_ meetperiode*. Another important feature is *voertuigcategorie*, which indicates the vehicle category the measurements were taken for. The vehicle category differs per sensor, some sensors measured only three different vehicle types, which are shown in Table 3.2, while others measured five different vehicle types, shown in Table 3.3. The two features *start_ locatie_ latitude* and *start_ locatie_ longitude* indicate the latitude and longitude of the sensor location, which are important features for the GNN to create a graph out of the sensor locations.

Table 3.2: Three different vehicle categories used by sensors.

	Vehicle Category
1	length ≤ 5.6
2	$5.6 <$ length ≤ 12.2
3	length > 12.2

Table 3.3: Five different vehicle categories used by sensors.

	Vehicle Category
1	$1.85 <$ length ≤ 2.4
2	$2.4 <$ length ≤ 5.6
3	$5.6 <$ length ≤ 11.5
4	$11.5 <$ length ≤ 12.2
5	length > 12.2

3.2.4 Missing Data

For this project, it is important that the sensors measure the traffic intensity separately for different vehicle categories so that the impact of the roadworks on freight traffic and passenger traffic can be analysed. However, it was found that for some sensors no vehicle categories were recorded. Because of this, these sensors were removed from the dataset.

From visual inspection, it was found that two sensors had a considerable period of time when no intensity was measured, Since these dates are important for the analysis, these two sensors were removed from the dataset as well.

Some values for speed and intensity were missing as well for other sensors, but as these missing values were just incidental occurrences, the values were interpolated using a linear interpolation method.

3.2.5 Data Aggregation

The dataset contains different types of sensors, some sensors measure multiple lanes, but others only measure one lane so there are multiple sensors in one location. The measurements from these multiple sensors at one location are added together, resulting in a total of 46 sensors. The locations of these sensors are shown in Figure 3.4.

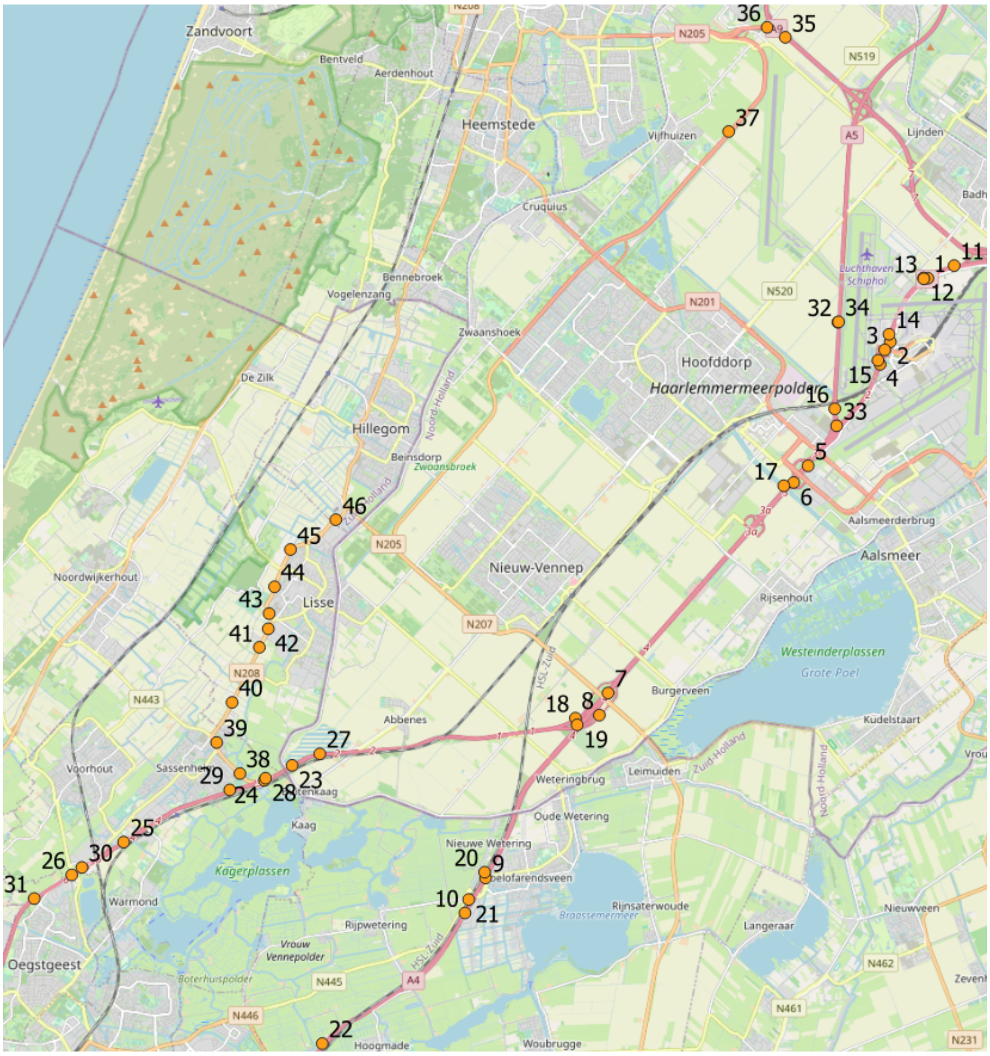


Figure 3.4: Map showing the locations of the 46 sensors.

Because the focus of this paper is on the impact of maintenance on freight traffic, the different vehicle categories have to be aggregated in a way that there will be two categories: freight and passenger traffic. As shown before in Tables 3.2 and 3.3, there were multiple vehicle categories that the sensors measured. To simplify these multiple vehicle categories into two categories, it was decided to put vehicles that measured less than or equal to 5.6

meters into the passenger category, and vehicles that are longer than 5.6 meters into the freight category.

3.2.6 Data Preparation

To prepare the dataset for training and testing the performance of different models, it was split into a training, validation, and test set. 70% of the data was used for training, 10% for validation, and 20% was used for testing. The dates and times that are included in these splits are shown in Table 3.4. From this table we can see that Phase 1 of the maintenance is included in the training set, as well as four days of Phase 2. This means that the models are able to learn from regular traffic patterns in April 2021, as well as from changes in traffic intensity during maintenance.

Table 3.4: Overview the dates and times that are included in the training, validation, and test set.

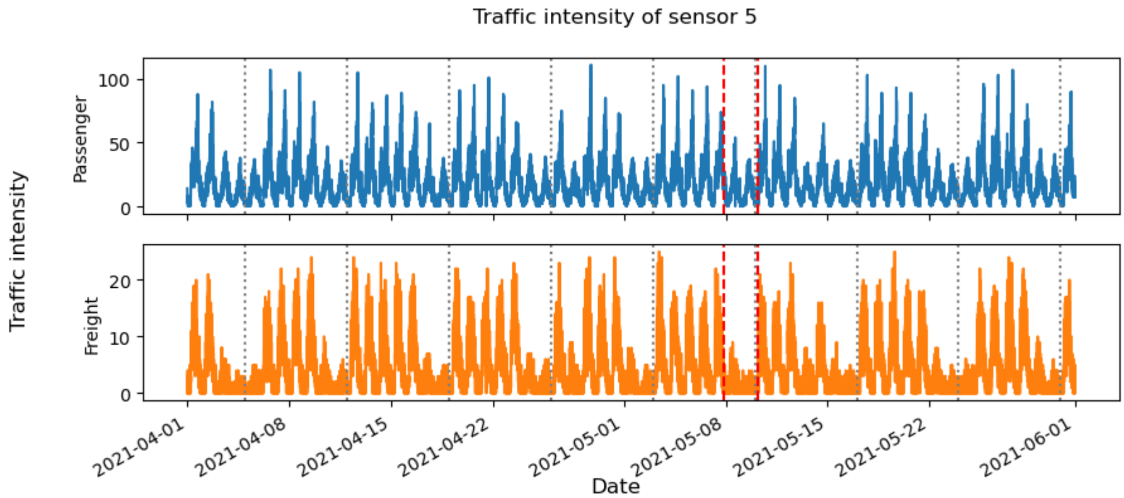
	Start date and time	End date and time
Training set	01-04-2021 00:00	13-05-2021 16:45
Validation set	13-05-2021 16:50	19-05-2021 19:05
Test set	19-05-2021 19:10	31-05-2021 23:55

3.3 Impact of Road Maintenance

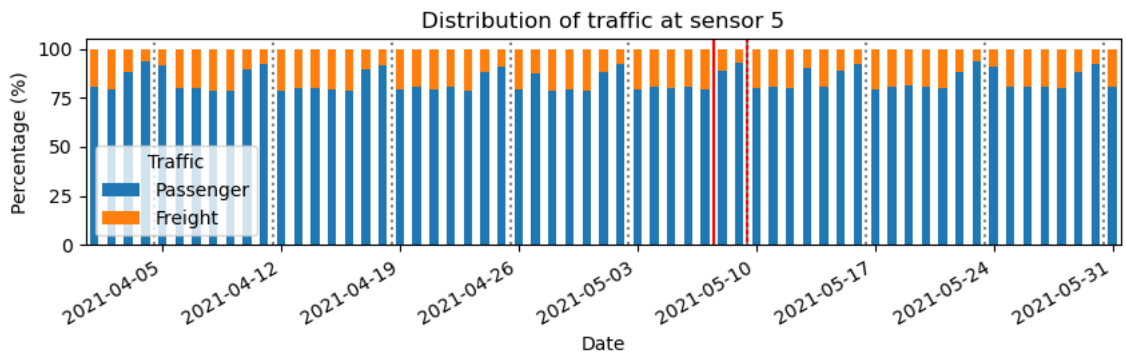
3.3.1 Analysis of Traffic Intensity

To analyse the impact that the maintenance on the A4 had on the traffic intensity, sensors at some key points were selected. These are sensors 5, 8, 20, 27, and 38. These were selected because sensor 5 is on the HRL near interchange De Hoek where Phase 2 took place, sensor 8 is on interchange Burgerveen where Phase 1 took place, sensor 20 is located on the A4 south of interchange Burgerveen, sensor 27 is located on the A44 further southwest from sensor 18, and sensor 38 is located on the N208 at the start of the detour that was used during Phase 1b. These locations are highlighted in Figure A.3 in Appendix A. For these sensors, we will look at the traffic intensity in April and May of 2021, and we will also examine the distribution of freight and passenger traffic during these months.

First we will look at the traffic intensity for sensor 5. This is shown in Figure 3.5a, where the blue line shows the traffic intensity for passenger traffic and the orange line shows that of freight traffic. The grey lines indicate the start of a week, starting on a Monday, and the red dotted lines are located at the start and end of Phase 1. The traffic intensity to the right of the rightmost red dotted line is measured during Phase 2. For freight traffic there does not seem to be a change in traffic intensity during Phase 1, however, it does look like the traffic intensity decreases during the first week of Phase 2. For passenger traffic, it does look like there is a slight increase in traffic intensity during Phase 1, and similar to freight traffic, there seems to be a slight decrease in traffic intensity during Phase 2. From Figure 3.5b we can see that the traffic intensity does not appear to change during maintenance as compared to the distribution before maintenance.



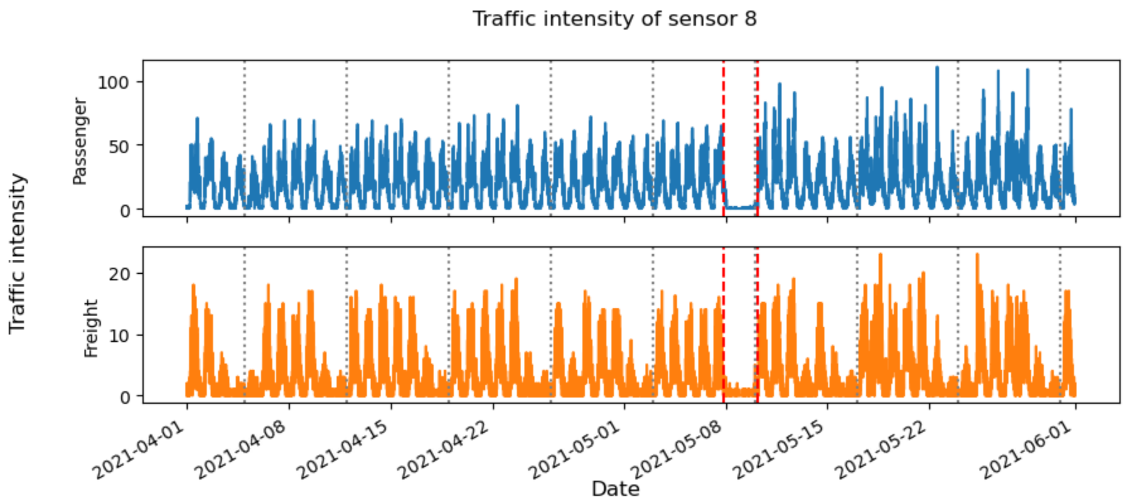
(a) Traffic intensity at sensor 5.



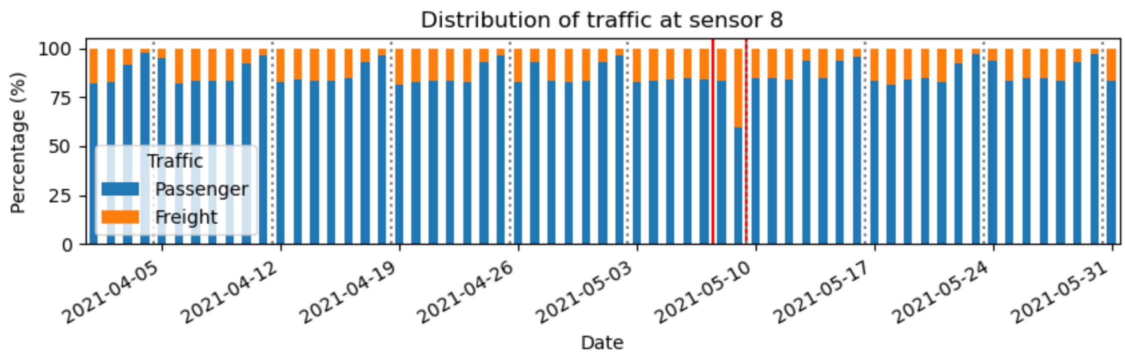
(b) Distribution of traffic at sensor 5.

Figure 3.5: Traffic intensity and distribution of traffic at sensor 5.

For sensor 8, shown in Figure 3.6 we can see a clear decrease in traffic intensity during Phase 1, especially for passenger traffic. We can also see this change in the distribution of traffic, comparatively, passenger traffic decreases more than freight traffic. This decrease in traffic makes a lot of sense because sensor 8 is located on interchange Burgerveen, where less traffic can pass through during Phase 1. During Phase 2, it seems that the traffic intensity increases during Phase 2, but there does not appear to be a change in distribution during this phase.



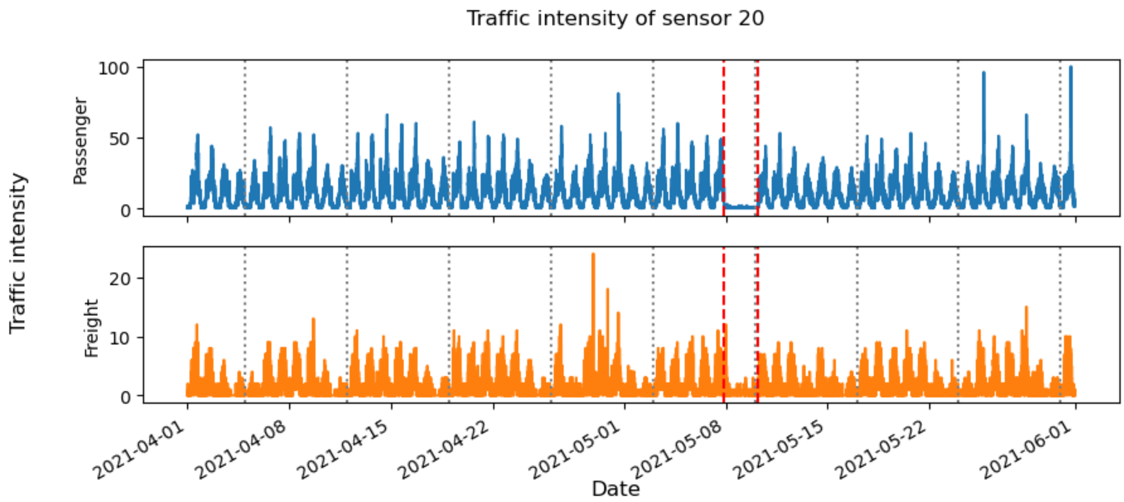
(a) Traffic intensity at sensor 8.



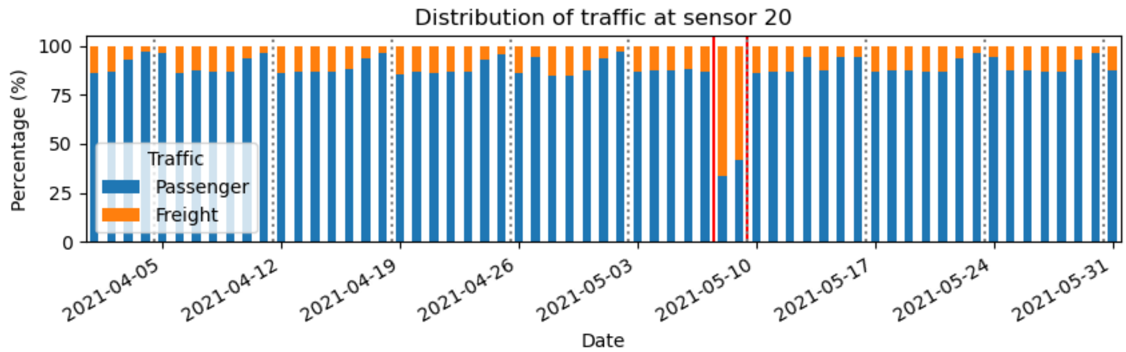
(b) Distribution of traffic at sensor 8.

Figure 3.6: Traffic intensity and distribution of traffic at sensor 8.

Similar to sensor 8, sensor 20 shows a clear decrease in traffic intensity during Phase 1, this can be seen in Figure 3.7a. This makes sense because it is located on the A4 south of interchange Burgerveen, where traffic coming from the A4 cannot go. Also similar to sensor 8, there is a shift in the distribution of traffic during Phase 1, which can be seen in Figure 3.7b. Comparatively, passenger traffic decreases more than freight traffic. During Phase 2, it looks like there is a slight decrease in traffic intensity during the first two weeks, but in the third week of Phase 2, it seems to go back to the level it was before maintenance, even with some peaks for passenger traffic.



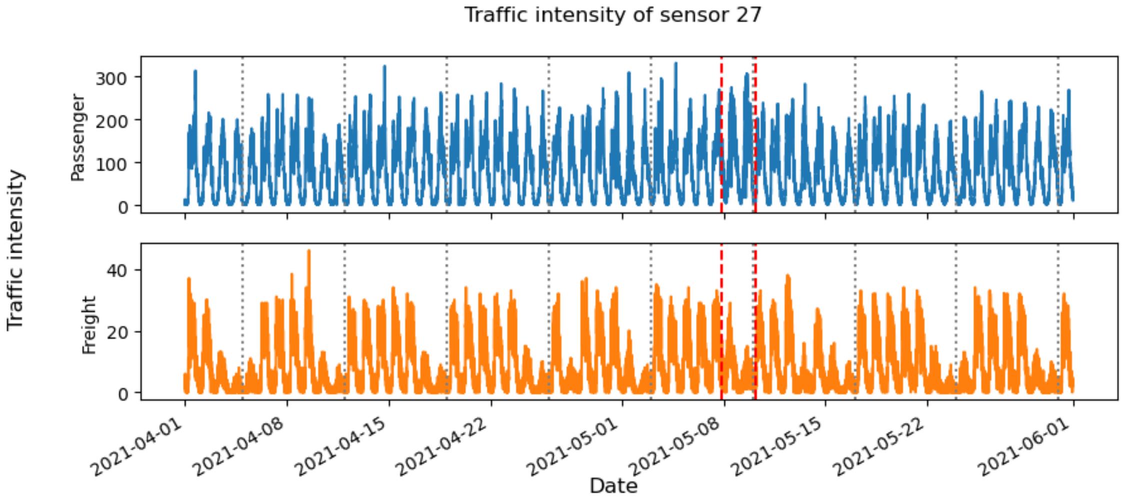
(a) Traffic intensity at sensor 20.



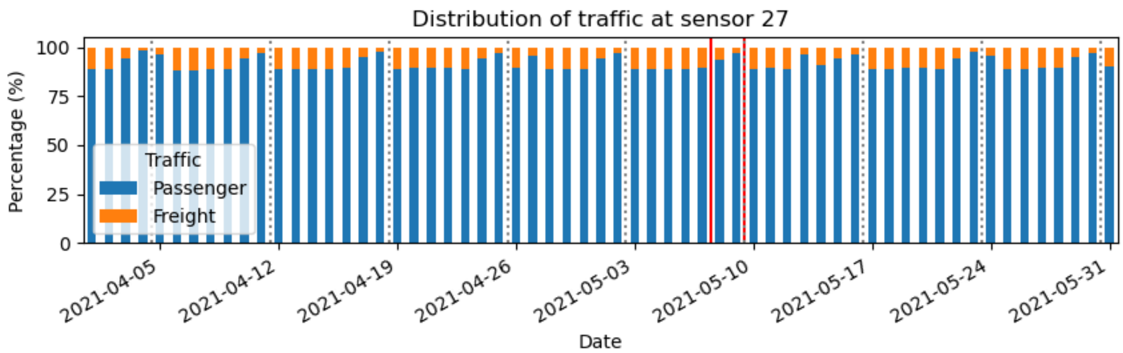
(b) Distribution of traffic at sensor 20.

Figure 3.7: Traffic intensity and distribution of traffic at sensor 20.

During Phase 1, sensor 27 shows an increase in both passenger and freight traffic, this is shown in Figure 3.8a. This increase is expected because traffic from the A4 going southbound towards Den Haag is redirected onto the A44, where sensor 27 is located. During Phase 2, there does not seem to be a big difference between traffic then and before maintenance. The distribution of traffic also does not change noticeably during Phase 1 nor during Phase 2, which can be seen in Figure 3.8b.



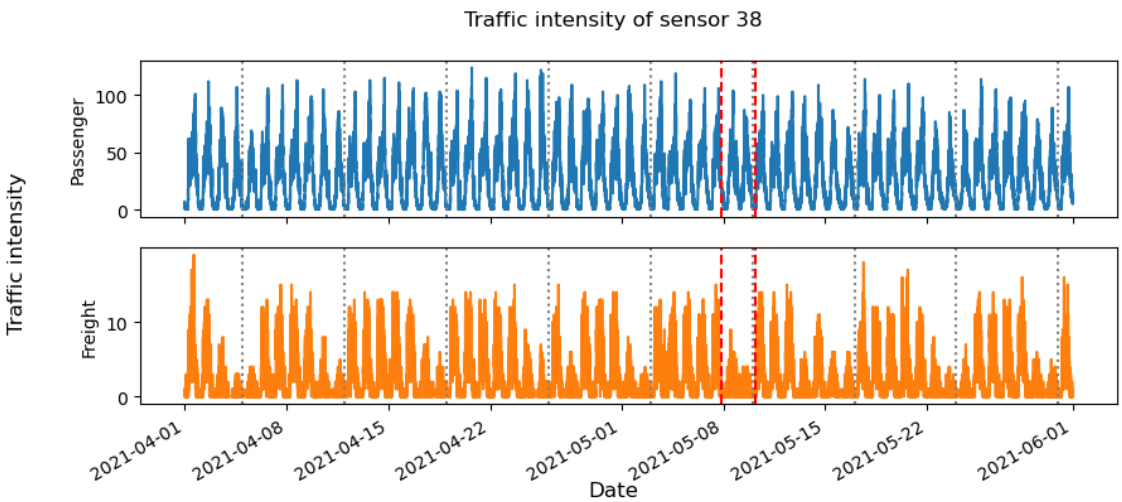
(a) Traffic intensity at sensor 27.



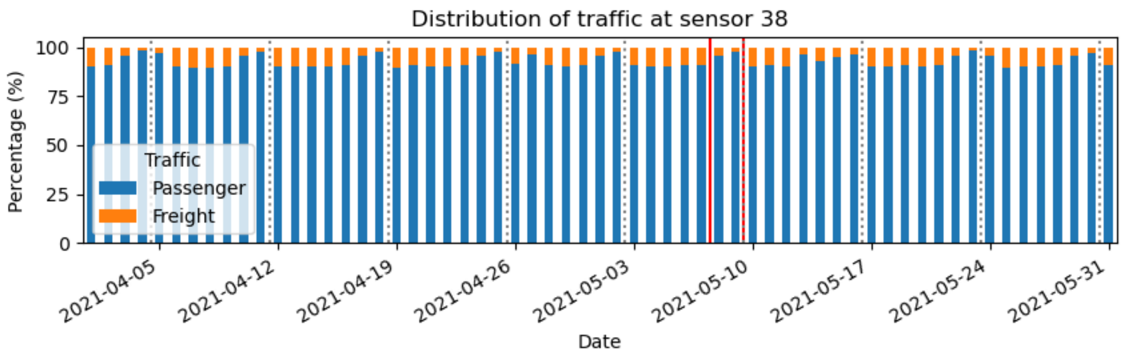
(b) Distribution of traffic at sensor 27.

Figure 3.8: Traffic intensity and distribution of traffic at sensor 27.

Finally, we will look at sensor 38, which is located on the detour that is in place during Phase 1b. Figure 3.9a shows that for passenger traffic, there does not seem to be a change in traffic intensity during Phase 1, but for freight traffic, there seems to be a slight increase in traffic. During Phase 2, it seems that there might be a decrease in traffic intensity for both passenger and freight traffic. The distribution of traffic, shown in Figure 3.9b also does not seem to change for sensor 38 when comparing the distribution before and during maintenance.



(a) Traffic intensity at sensor 38.



(b) Distribution of traffic at sensor 38.

Figure 3.9: Traffic intensity and distribution of traffic at sensor 38.

3.3.2 Hypothesis Testing

To test whether the road maintenance had any impact on traffic intensity, hypothesis testing will be performed. Because the traffic intensity does not have a normal distribution, the Mann-Whitney U (MWU) test will be used to compare the traffic intensity before and during road maintenance. The MWU test assumes that sample $(X_1, ..., X_n)$ comes from distribution X and that sample $Y_1, ..., Y_m$ comes from distribution Y.

Another test that will be performed is the two-sample t-test, which compares the mean of two samples. This test can be used as the means of the two samples should follow normal distributions. In large samples, this is true due to the central limit theorem, despite the samples not having a normal distribution themselves (19). The significance level for both tests will be set to $\alpha = 0.05$.

For these hypothesis tests, two sets of hypotheses will be used. The hypothesis set that is used for testing the hypothesis that there is less traffic during maintenance than before maintenance is shown in (H1) and the hypothesis set for testing the hypothesis that there is more traffic during maintenance is shown in (H2).

$$\begin{aligned} &H_{0,\text{MWU}} : X \geq Y,\ H_{1,\text{MWU}} : X < Y \\ &H_{0,\text{t-test}} : \bar{X} \geq \bar{Y},\ H_{1,\text{t-test}} : \bar{X} < \bar{Y} \end{aligned} \tag{H1}$$

$$\begin{aligned} &H_{0,\text{MWU}} : X \leq Y,\ H_{1,\text{MWU}} : X > Y \\ &H_{0,\text{t-test}} : \bar{X} \leq \bar{Y},\ H_{1,\text{t-test}} : \bar{X} > \bar{Y} \end{aligned} \tag{H2}$$

Phase 1

To test the impact of maintenance during Phase 1, the traffic intensity of the weekends in April 2021 will be compared to the traffic intensity during Phase 1. So in the hypotheses, X is the distribution of the traffic intensity during the weekends before Phase 1 with mean $\bar{X}$, and Y is the distribution of traffic intensity during Phase 1 with mean $\bar{Y}$. For the sensors in the detour during Phase 1b (sensor 35-46), X is the distribution of traffic intensity during the weekends from Saturday 21:00 until Sunday 09:00, and Y is the traffic intensity during Phase 1b, and $\bar{X}$ and $\bar{Y}$ are their respective means.

Hypothesis set (H1) is used for sensors 5, 6, 7, 8, 19, 20, 21, and 22 to test the hypothesis that these sensors measured less traffic during Phase 1 than before Phase 1. These sensors are shown in Appendix A in Figure A.4a and are indicated with a blue dot. The reasons these hypotheses are chosen for the sensors are listed below:

- Sensors 5 and 6: because there are only two lanes going onto the A4 during phase 1, possibly leading to a lower traffic intensity at these two sensors further down the A4.
- Sensors 7, 8: because they are located on interchange Burgerveen.
- Sensor 19: because it is on the HRL, which is closed.
- Sensors 20, 21, and 22: because all traffic from interchange Burgerveen in that direction was redirected to the A44, likely causing lower traffic intensity at these three sensors.

Hypothesis set (H2) is used for sensors 18, 27-31, and 35-46. These sensors are shown in Appendix A in Figure A.4a and are indicated with a red dot. The reasoning for choosing the hypothesis that there is more traffic at these locations during Phase 1 than before is listed below:

- Sensors 18: because it is located on the road connecting the A4 and A44, where all cars going in the direction of Den Haag are redirected to.
- 27-31: because they are located on the A44 where cars going in the direction of Den Haag are redirected to.
- Sensors 35-46: because they are located on the detour which is used at night when the A4 at interchange Burgerveen is closed entirely.

In this analysis, we will only consider the sensors that have rejected the null hypothesis for both tests, as this is a clear indication that the traffic intensity differs before and during maintenance. From Table A.1 in Appendix A, we can see that the null hypothesis is rejected in favour of the alternative hypothesis, mean that there is indeed less traffic during Phase 1 at sensor 7 for freight traffic, an at sensors 8, 19, 20, 21, and 22 for both types of traffic. And there is more traffic at sensors 18, 27-31, 38-45 for both types of traffic, and at sensors 35, 37, and 46 for passenger traffic. Figure A.4b in Appendix A visualises these results.

Phase 2

For Phase 2, we will compare the distribution of the traffic intensity from April 1st 00:00 until May 7th 21:00, X, to the traffic intensity during Phase 2, Y. Again, we will use the two hypothesis sets as specified in (H1) and (H2).

For sensors 5, 6, 7, and 8 the hypothesis is that there will be less traffic during Phase 2 than there was before maintenance because these sensors 5 and 6 are located on the HRL where only two lanes are open. Hypothesis set (H1) will be used to test this hypothesis for these sensors and they are shown in Appendix A in Figure A.5a with a blue dot.

The other sensors that hypothesis testing was performed for during Phase 2 were sensors 17, 18, 19, 23-31, 35, 36, and 37. For these sensors, hypothesis set (H2) was used to test whether there was more traffic during Phase 2 than before maintenance. These sensors were chosen because sensors 17, 18, and 19 are on the HRR where now more lanes are open than usual, thus it is likely that more traffic will be measured there. Sensors 23-31 were chosen because they are located on the A44, which is part of the advisory route during Phase 2 and sensors 35-37 are also on this advisory route. The sensors are shown in Appendix A in Figure A.5a with red dots.

From Table A.2 in Appendix A, we can see that the null hypothesis is rejected in favour of the alternative hypothesis for sensors 7, 8, 23, and 35-37 for both types of traffic, for sensors 5, 24-28, 30, and 31 for freight traffic, and for sensors 17, 18, and 19 for passenger traffic. This means that there is less traffic at sensor 5 for freight traffic and at sensors 7 and 8 for both types of traffic during Phase 2. There is more passenger traffic at sensors 17, 18, and 19 but not more freight traffic. However, sensors 24-28, 30, and 31 measure more freight traffic, but not more passenger traffic. This could indicate that freight traffic does make use of advisory routes, while passenger traffic does not as sensors 24-28, 30, and 31 are located on the advisory route. Figure A.5b in Appendix A visualises these results.

4

Methodology

4.1 Seasonal Naive

For the baseline model, a seasonal naive model is used. This model uses the value of the previous season as a forecast for the current time point (13). The forecast for time t ($\hat{y}_t$) is given by the following equation:

$$\hat{y}_t = y_{t-m}, \tag{4.1}$$

where y_{t-m} is the actual value at time $t-m$ and m is the number of time steps in a season.

4.2 Holt-Winters Exponential Smoothing

Holt-Winters Exponential Smoothing, also called triple exponential smoothing, is a type of exponential smoothing which includes a seasonal component (30). Since there are different types of trends, this model has two variations; one with an additive and another with a multiplicative seasonal component. The additive model is best applied when the seasonal component remains more or less constant over time, meaning that the amplitude of the seasonal fluctuations stays constant, while the multiplicative model is better for times series that have a seasonal component which varies over time (13).

The equations for the additive model are shown in Equations 4.2-4.5 below.

$$\hat{y}_{t+h} = \ell_t + hb_t + s_{t+h-m} \tag{4.2}$$

$$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) \tag{4.3}$$

$$b_t = \beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1} \tag{4.4}$$

$$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}. \tag{4.5}$$

Equations 4.6-4.9 show the equations that are used for the multiplicative model.

$$\hat{y}_{t+h} = (\ell_t + hb_t)s_{t+h-m} \tag{4.6}$$

$$\ell_t = \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1}) \tag{4.7}$$

$$b_t = \beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1} \tag{4.8}$$

$$s_t = \gamma \frac{y_t}{\ell_{t-1} + b_{t-1}} + (1-\gamma)s_{t-m}. \tag{4.9}$$

For both the additive and the multiplicative model m is the number of time steps in a season, h is the number of time steps that the forecast lies in the future, and α, β, and γ are smoothing constants. Equations 4.2 and 4.6 are the forecast and ℓ_t, b_t, and s_t represent the level, trend, and seasonal component of the forecast, respectively.

4.3 XGBoost

XGBoost stands for eXtreme Gradient Boosting, and it is a tree boosting system which combines multiple regression trees into one algorithm. This combination is called a tree ensemble model, and it has as advantage that multiple models that perform average by themselves can have great performance when their forecasts are combined.

The objective of XGBoost is to minimize the following equation (6).

$$\begin{aligned} \mathcal{L}^{(t)} &= \sum_{i=1}^{n} l(y_i, \hat{y}_i^{t-1} + f_t(\mathbf{x}_i)) + \Omega(f_t) \\ \text{where } \Omega(f) &= \gamma T + \frac{1}{2}\lambda\|w\|^2 \end{aligned} \tag{4.10}$$

Here $\hat{y}_i^{(t)}$ is the prediction of the i^{th} instance at the t^{th} iteration and l is a loss function that calculates the error between prediction $\hat{y}_i$ and actual value y_i. Ω is a regularisation function that penalises the complexity of a single regression tree model, denoted by f_t. In Equation 4.10, the regression tree f_t that most improves the model is greedily added to it.

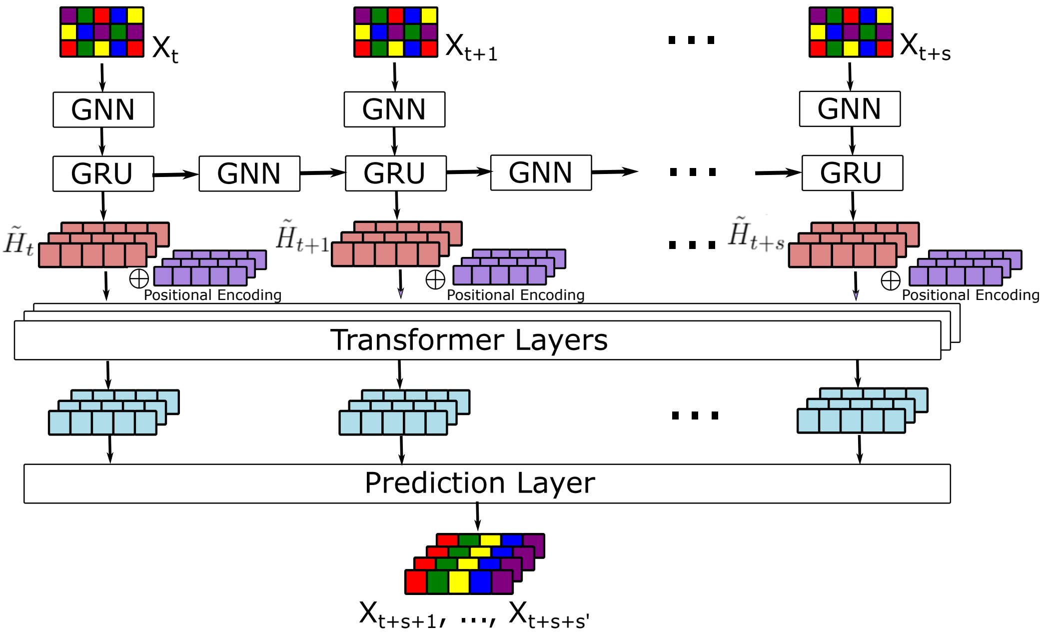


Figure 4.1: Spatial Temporal Graph Neural Network framework, adapted from Wang et al. (28)

4.3.1 SHapley Additive exPlanations (SHAP)

SHAP is a framework for interpreting predictions made by complex models (20). This framework introduces so-called SHAP values to measure feature importance for a model.

4.4 Graph Neural Network

As discussed in Section 2.3.1, there are many different types of graph neural networks. The GNN that will be implemented for this paper, is called a Spatial Temporal Graph Neural Network (STGNN) (28). The framework for this STGNN consists of 4 layers, a spatial graph neural network layer to capture spatial information, a Gated Recurrent Unit (GRU) layer to capture local temporal dependencies, a transformer layer to capture global temporal dependencies, and as final layer a multi-layer feedforward network to output predictions. This framework is shown in Figure 4.1.

For the STGNN, the road network must be modelled using a directed graph G, which can be defined in the following way: $G = (V, E)$, where $V = \{v_1, ..., v_N\}$ is the set of nodes, which represent the N sensors which are chosen to be analysed, and $E = \{e_1, ..., e_M\}$ is the set of edges, which represent the roads that connect the different sensors. There will only be an edge between two sensors if they are direct neighbours of each other.

4.4.1 Graph Neural Network Layer

The GNN layer (16) tries to capture the spatial information from the directed graph G, it does so by using the input matrix $X_t \in \mathbb{R}^{N \times d_{\text{in}}}$ in Equation 4.11, where N is the number of sensors in G and d_{in} is the number of features used in the model.

$$\begin{aligned} \text{GNN}(X_t) &= X_{\text{out},t} = \text{ReLU}(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}X_t W) \\ \text{with } \tilde{A} &= A + I_N \\ \tilde{D}_{ii} &= \sum_j \tilde{A}_{ij} \\ \tilde{D}_{ij} &= 0 \quad \text{if } i \neq j \end{aligned} \tag{4.11}$$

Here, A is the adjacency matrix, which can be constructed in multiple ways as shown in Section 4.4.5, and I_N is the identity matrix of size N.

4.4.2 Gated Recurrent Unit Layer

The output of the GNN layer is used as input for a Gated Recurrent Unit layer to model the local temporal dependency. This layer applies the GRU (7) to each of the inputs, obtained from the GNN layer, separately. The GRU operation at time t for node v_i can be seen in Equation 4.12 below.

$$\begin{aligned} z_t &= \sigma_z(W_z\tilde{X}_{\text{out},t}[i,:] + U_z\tilde{H}_{t-1}[i,:] + b_z), \\ r_t &= \sigma_r(W_r\tilde{X}_{\text{out},t}[i,:] + U_r\tilde{H}_{t-1}[i,:] + b_r), \\ \hat{H}_t[i,:] &= \tanh\left(W_h\tilde{X}_{\text{out},t}[i,:] + U_h(r_t \odot U_h\tilde{H}_{t-1}[i,:]) + b_h\right), \\ H_t[i,:] &= (1 - z_t) \odot \tilde{H}_{t-1}[i,:] + z_t \odot \hat{H}_t[i,;] \\ \tilde{H}_t &= \text{GNN}(H_t) \end{aligned} \tag{4.12}$$

where σ is the sigmoid function, $\odot$ is the element-wise multiplication, and the matrices $W_z, W_r, W_h, U_z, U_r, U_h$ are the parameters to be learned. $H_t[i,:]$ is the hidden representation of the current time step and is also the output of the GRU layer.

4.4.3 Transformer Layer

The transformer layer models the global temporal dependency. This layer is applied to the output sequence $(H_1[i,:], ..., H_T[i,:])$ from the GRU layer. The details of a transformer model are explained in Section 4.5.

in Equation 4.11 Because the transformer layer needs input which is arranged in sequences per sensor/node, the output of the GRU layer needs to be transformed. To do this, the output of the GRU layer will be stacked row-wise, such that $H^{v_i} = (H_1[i,:], ..., H_T[i,:]) \in \mathbb{R}^{T \times d_{\text{in}}}$. The positional encoding e_t will then be added to matrices H^{v_i}, such that $H_t'^{v_i}[i,:] = H_t^{v_i}[i,:] + e_t$. This matrix H'^{v_i} can then be used as input for the transformer layer. The output of the transformer layer is $H_{\text{out}}^{v_i} \in \mathbb{R}^{T \times d}$.

4.4.4 Prediction Layer

To make the final predictions, a multi-layer feed-forward network is used. This network uses the output $\{H_{\text{out}}^{v_i} | v_i \in V\}$ of the transformer layer to make these forecasts.

4.4.5 Adjacency matrix

4.4.5.1 Physical distance with neighbours

A very intuitive way to model the adjacency matrix is by using a distance-based matrix, namely a weighted adjacency matrix. The closer two sensors are to each other, the bigger their connection is in the adjacency matrix (36). The entries for the weighted adjacency matrix can be determined by Equation 4.13.

$$A_{ij} = \begin{cases} \exp(-\frac{d_{ij}^2}{\sigma^2}), & \text{if } i \neq j \text{ and } \exp(-\frac{d_{ij}^2}{\sigma^2}) \geq \epsilon \\ 0, & \text{otherwise} \end{cases} \tag{4.13}$$

d_{ij} represents the distance between sensor i and sensor j, and σ^2 and ϵ are thresholds to be set.

4.4.5.2 Correlation matrix

Another way to model the adjacency matrix is by using a similarity-based matrix (14). This adjacency matrix uses the Pearson correlation between two sensors, and its entries can be calculated as shown in Equation 4.14.

$$\begin{aligned} A_{ij} &= \text{corr}(v_i, v_j) \\ \text{with } \text{corr}(v_i, v_j) &= \frac{\sum_{t=1}^{T}(v_{i,t} - \bar{v}_i)(v_{j,t} - \bar{v}_j)}{\sqrt{\sum\sum_{t=1}^{T}(v_{i,t} - \bar{v}_i)^2 \sum\sum_{t=1}^{T}(v_{j,t} - \bar{v}_j)^2}} \end{aligned} \tag{4.14}$$

where $v_{i,t}$ is a measurement for sensor i at time t, and $\bar{v}_i$ is the mean of the measurements for sensor i.

4.5 Transformer

A transformer (26) consists of an encoder and a decoder. The encoder receives a sequence $(x_1, ..., x_n)$ as an input and maps it to a continuous representation $z = (z_1, ..., z_n)$. The decoder uses the sequence z to generate an output $(y_1, ..., y_m)$ elementwise.

4.5.1 Encoder and Decoder

The encoder consists of a stack of M layers that are identical. Each of these layers contain a multi-head self-attention sublayer and a fully connected feed-forward network as a sublayer. Both of these sublayers have a residual connection from the input before the sublayer to a layer normalization, which means the output for each sublayer is a normalized version of the output of the sublayer and the residual added together.

in Equation 4.11The decoder consists of a stack of M identical layers as well. The decoder layer uses the same sublayers as the encoder layer, however, the decoder has another sublayer added before the other two sublayers. This new sublayer uses a masked multi-head attention mechanism on the output of the encoder stack. This masking ensures sure that the predictions of the model for a certain input position cannot depend on information after this input. The decoder also uses residual connections from the input before the sublayer to a layer normalization in the same way that is used for the encoder.

4.5.2 Multi-Head Attention

The multi-head attention layer uses queries, keys, and values. The keys have dimension d_k, and the values have dimension d_v. Then the attention function is defined as

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}V\right) \tag{4.15}$$

where $Q, K \in \mathbb{R}^{T \times d_k}$ and $V \in \mathbb{R}^{T \times d_v}$ are the queries, keys, and values for all the nodes, respectively.

$$\begin{aligned} \text{MultiHead}(Q, K, V) &= Concat(\text{head}_1, ..., \text{head}_S)W^O; \\ \text{head}_s &= \text{Attention}_s(QW_s^Q, KW_s^K, VW_s^V) \end{aligned} \tag{4.16}$$

where W_s^Q, W_s^K, W_s^V, and W^O are matrices to be learned.

4.5.3 Positional Encoding

Because the transformer does not utilise recurrence or convolution, it has to obtain information about the order of the sequence in another manner. This is done by using positional encoding and adding it to the input embeddings. These positional encodings can be determined as shown in Equation 4.17.

$$e_t = \begin{cases} \sin(t/10000^{2i/d_{\text{model}}}), & \text{if } t = 0, 2, 4, ... \\ cos(t/10000^{2i/d_{\text{model}}}), & \text{otherwise,} \end{cases} \tag{4.17}$$

where d_{model} is the dimension of the output of the transformer model.

4.6 Error Measures

In this paper, two error measures will be used to compare the performance of the different algorithms.

4.6.1 Root Mean Square Error)

The Root Mean Square Error (RMSE) can be calculated using the following equation,

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}. \tag{4.18}$$

The RMSE will be more affected by predictions which are much worse than others since the difference between the forecast and the actual value is squared.

4.6.2 Mean Absolute Error

The Mean Absolute Error (MAE) uses the absolute value of the difference between the forecast and the actual value to determine the error, this is shown in Equation 4.19

$$\text{MAE} = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n} \tag{4.19}$$

5

Implementation Details

5.1 Seasonal Naive

For the seasonal naive model, the seasonal period was chosen to be 24 hours for both the 5 minutes ahead forecast and the hour ahead forecast. This means that the forecast for a certain time point would be the value of the traffic intensity 24 hours before that time point. Since the dataset has an aggregation level of 5 minutes, the seasonal value is $m = 12 * 24 = 288$, filling this in in Equation 4.1 gives the following equation for the seasonal naive model:

$$\hat{y}_t = y_{t-288}. \tag{5.1}$$

5.2 Holt-Winters

Just like for the seasonal naive model, the seasonal period was chosen to be 24 hours, so $m = 288$ and since the time series has a seasonal component that is more or less constant, the additive model is used for both the 5-minute and the hour-ahead forecast. The model is occasionally updated with data from past times that a forecast was made for, to increase the forecasting accuracy.

5.2.1 5-Minute Forecast

The 5-minute forecast is a 1-step ahead forecast, this means that in Equations 4.2-4.5, h will be equal to 1, and as mentioned in Section 5.1, $m = 288$, and the values for the smoothing constants are shown in Table B.1 for freight traffic and in Table B.2 for passenger traffic.

5.2.2 Hour-Ahead Forecast

The hour-ahead forecast is 12 time steps ahead, this means that in Equations 4.2-4.5, h will be equal to 12, and equal to the 5-minute forecast, $m = 288$ and the smoothing constants are shown in Table B.1 and B.2 for freight and passenger traffic, respectively, as well since the values of the smoothing constants are the same for both the 5 minute and the hour ahead forecast.

5.3 XGBoost

The standard hyperparameters for the XGBoost model were used, but different combinations of features were tested.

5.3.1 Features

The same types of features were created for the 5 minute and the hour ahead forecasts. There are three types of features that were created, namely time features, moving average features and average intensity features.

The time features are the same for the 5 minute and the hour ahead forecasts. They are derived from the date and time that the intensity measurement was taken on, and they are 'hour', 'day of week', and 'day of month'.

5.3.2 5-Minute Forecast

The moving average features are based on a simple moving average of a specified number of intensity measurements. These features can be determined using Equation 5.2 for a forecast for time $t + 1$.

$$\text{SMA}_k = \frac{\sum_{i=1}^{k} x_t}{k} \tag{5.2}$$

where x_t is the value of the traffic intensity at time t. These moving average features were created for times up to an hour before $t + 1$, so for $k = 1, ..., 12$.

The average intensity features are just the average intensity, but then from previous time periods. If the intensity for time t is x_t, then the average intensity features that were created were $x_{t-1}, ..., x_{t-12}$, i.e. the intensities from 5 minutes before t, up to an hour before t.

5.3.3 Hour-Ahead Forecast

The moving average features for the hour-ahead forecast are defined in the same way as for the 5-minute forecast, only now the forecast is made for time $t + 12$. These moving average features were created for times going from one hour up to two hours before t, so for $k = 1, ..., 12$.

The average intensity features are defined in the same manner as for the 5-minute forecast. For the forecast at time $t + 12$, the average intensity features that were created were $x_t, ..., x_{t-12}$, i.e. the intensities from time t, up to an hour before t.

5.3.4 Final XGBoost Models

Using the SHAP values, the best features were selected from an XGBoost model that contained all features. Then, to find the best model, six different models were created. Starting with a model with only the best feature, up to a model containing the top six features.

The features that were used for the 5-minute forecasts are shown in Table B.3, and Table B.4 for freight and passenger traffic, respectively. And the features that were used for the hour-ahead forecasts are shown in Table B.5, and Table B.6 for freight and passenger traffic, respectively.

5.4 GNN

5.4.1 Features

For the implementation of the STGNN, the features that were chosen are the hour, weekday, and month of the start of the measuring period, the average intensity, and the average speed. The hour, weekday, and month were transformed into cyclical features using the following equations:

$$\begin{aligned} x_{\sin} &= \sin\left(\tfrac{2\pi x}{\max(x)}\right) \\ x_{\cos} &= \cos\left(\tfrac{2\pi x}{\max(x)}\right), \end{aligned} \tag{5.3}$$

where x represents the vector that contains the cyclical feature. In total eight features were used, and the number of sensors that were used as input is 46, which means the feature matrices are $X_t \in \mathbb{R}^{46\times 8}$. For the transformer layer, $d_{\text{model}} = 8$ is used, the number of features in the model.

The different learning rates that were tried for the STGNN were 0.01, 0.001, 0.0001, and 0.00001. Together with trying out two different types of adjacency matrices, there were 8 different setups for both the 5-minute and the hour-ahead forecasts.

5.4.2 5-Minute Forecast

The input to the GNN for the 5-minute forecast will have a window size of 6 time steps (half an hour) for freight traffic and a window size of 12 time steps (an hour) for passenger traffic to predict the next time step. As adjacency matrix, the distance-based matrix was used for freight traffic, and the correlation matrix was used for passenger traffic. The learning rates that were used were 0.00001 and 0.0001 for freight and passenger traffic, respectively.

5.4.3 Hour-Ahead Forecast

The input to the GNN for the hour-ahead forecast has a window size of 12, meaning that the traffic intensities $x_t, ..., x_{t-11}$ will be used to create a forecast for time $t+12$. For both freight and passenger traffic, the correlation matrix was used as adjacency matrix and both use a learning rate of 0.0001.

5.5 Transformer

For the transformer model, it was decided to use $M = 8$ for the encoder and decoder stacks. For the positional encoding, $d_model = 512$ like in the original paper (26). For the encoder sequence lengths, multiple values were tried out, namely $1, 3, 6, 9, 12, 24, 288$. And the learning rates that were tried were 0.01, 0.001, 0.0001, and 0.00001.

5.5.1 5-Minute Forecast

For the 5-minute forecast, the decoder (output) sequence length was set to 1. The parameters that were chosen for each sensor for freight traffic are shown in Table B.7, and those that were chosen for passenger traffic are shown in Table B.8.

5.5.2 Hour-Ahead Forecast

The decoder sequence length for the hour-ahead forecast was set to 12. The parameters that were chosen for freight traffic are shown in Table B.9, and those that were used for passenger traffic are shown in Table B.10.

6

Results

The mean errors for each of the models are shown in Table 6.1 and Table 6.2 for the 5-minute forecast and the hour-ahead forecast, respectively. From these tables, we can see that the model with the lowest error for freight traffic is XGBoost for both the 5-minute and the hour-ahead forecast, however, the GNN is a close second for the 5-minute forecast. For passenger traffic, the best model is the GNN for the 5-minute forecast and XGBoost for the hour-ahead forecast. The full tables showing the errors per sensor are shown in Appendix C in Tables C.1-C.8 and Figures C.1-C.8 visualise these errors for each of the sensors.

Table 6.1: Overview of the mean of the errors for the 5-minute forecasts.

	Metric	Baseline	Holt-Winters	XGBoost	GNN	Transformer
Freight	RMSE	5.0097	6.5639	**2.5015**	2.5228	4.8510
	MAE	3.2346	4.8537	**1.7162**	1.7237	4.1466
Passenger	RMSE	19.8530	46.1597	9.9588	**9.2536**	28.7138
	MAE	11.7526	37.5546	6.6968	**6.2266**	24.8073

Table 6.2: Overview of the mean of the errors for the hour-ahead forecasts.

	Metric	Baseline	Holt-Winters	XGBoost	GNN	Transformer
Freight	RMSE	5.4061	6.5064	**2.7721**	3.2765	3.2443
	MAE	3.5984	4.7952	**1.8815**	2.2411	2.5253
Passenger	RMSE	24.4048	45.7531	**12.8267**	14.1388	27.9945
	MAE	16.1847	37.0973	**8.6306**	9.4969	24.2121

6.1 Final Model

In Appendix C we can see that each sensor has a different model which works best for freight and passenger traffic, and for the 5-minute and hour-ahead forecasts. However, since it is most useful to have only one model make forecasts for all of the sensors, the model that had the best overall performance was chosen as the final model that will be applied to all sensors. This model is XGBoost because it had the lowest mean errors for freight for the 5-minute forecasts and for freight and passenger traffic for the hour-ahead forecasts, so for three out of four scenarios.

6.2 Visualisations

To create some more insight into the results, some visualisations will be made for the sensors that were selected in Section 3.3, sensors 5, 8, 20, 27, and 38. The locations of these sensors are shown in Figure A.3 in Appendix A. For these sensors, we will compare the forecasts that XGBoost made to the real traffic intensity and we will look at the uncertainty of the forecasts.

6.2.1 Real vs. Forecasted Traffic Intensity

6.2.1.1 5-minute Forecast

For sensor 5, the forecasted and real traffic intensity for passenger and freight traffic are shown in Figure 6.1. From this figure we can see that the 5-minute forecast follows the actual traffic intensity quit well, especially for passenger traffic. For freight traffic, when the traffic intensity oscillates, the forecast seems to stay in the middle of these fluctuations.

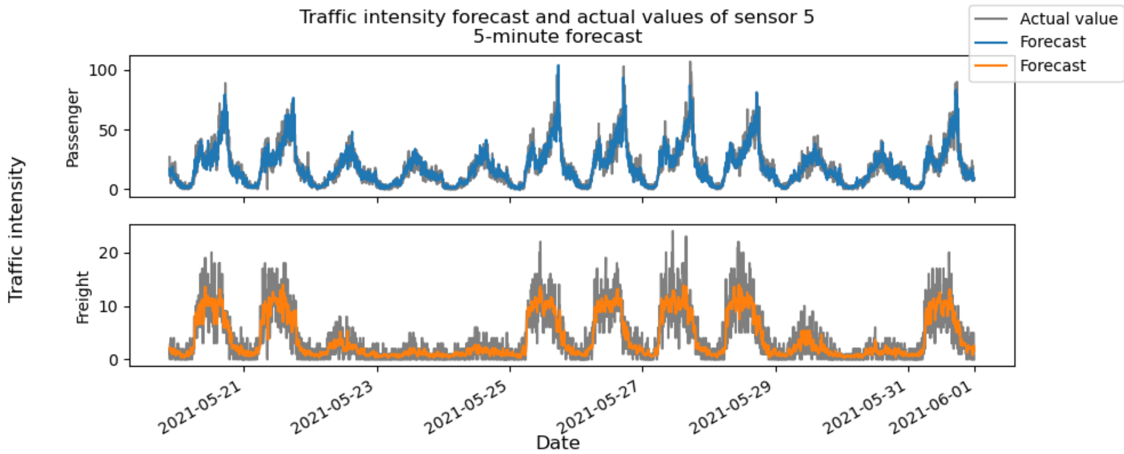


Figure 6.1: Real vs. 5-minute forecast of traffic intensity for sensor 5.

Figure 6.2 shows the 5-minute forecast of the traffic intensity for sensor 8, where we can see that the forecasted traffic intensity follows the general pattern of the real traffic intensity quite well. But for passenger traffic, the model is not able to predict the highest peaks and for freight traffic, again, the forecasts stay in the middle when the traffic pattern fluctuates.

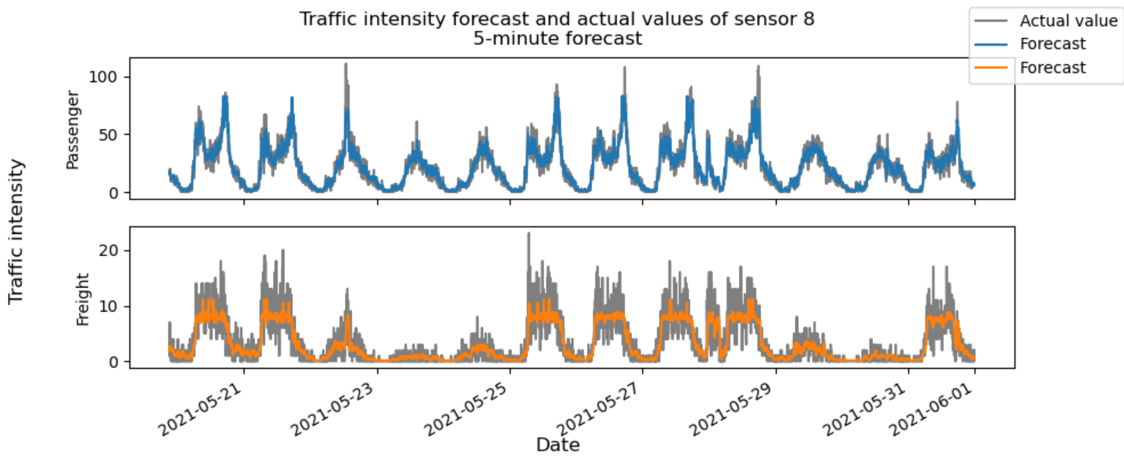


Figure 6.2: Real vs. 5-minute forecast of traffic intensity for sensor 8.

The 5-minute forecasts for sensor 20, shown in Figure 6.3, shows the same patterns as the forecasts for sensor 5 and 8. The forecast for passenger traffic follows the general traffic pattern nicely, except for a few peaks, where the forecast is not as high as the real traffic intensity. For freight traffic, the same pattern shows as well, when the actual traffic intensity fluctuates a lot, the forecasts stays in the middle of these fluctuations, but the forecast does follow the general traffic pattern.

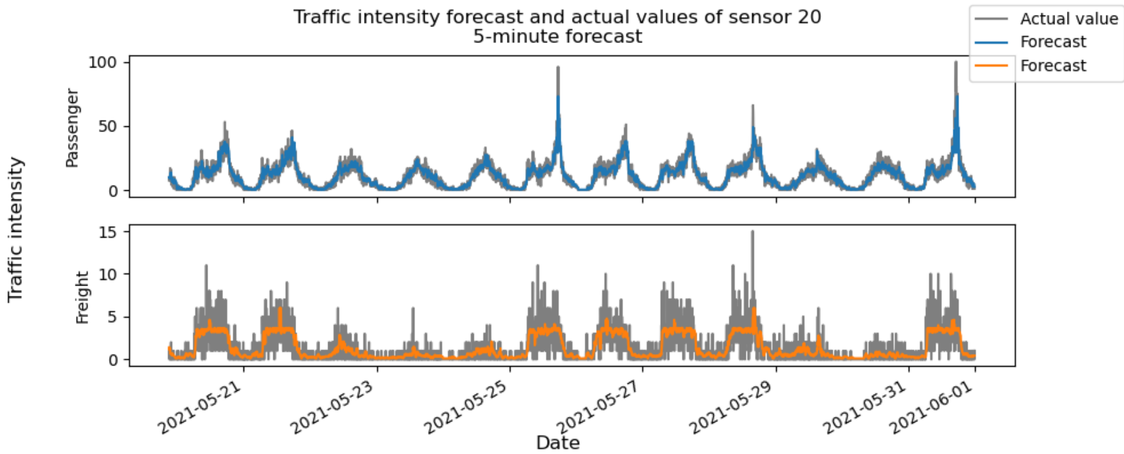


Figure 6.3: Real vs. 5-minute forecast of traffic intensity for sensor 20.

For sensors 27 and 38, which can be seen in Figures 6.4 and 6.5, the forecasts show the same patterns as for the other sensors. The forecasts follow the traffic pattern, but have some difficulty with peaks. And when the traffic pattern fluctuates, the forecasts stay in the middle of these fluctuations.

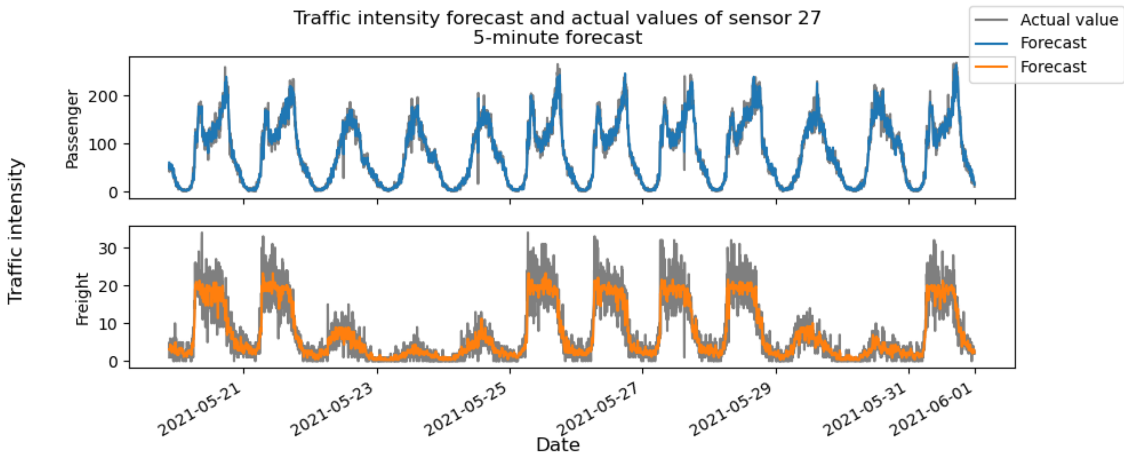


Figure 6.4: Real vs. 5-minute forecast of traffic intensity for sensor 27.

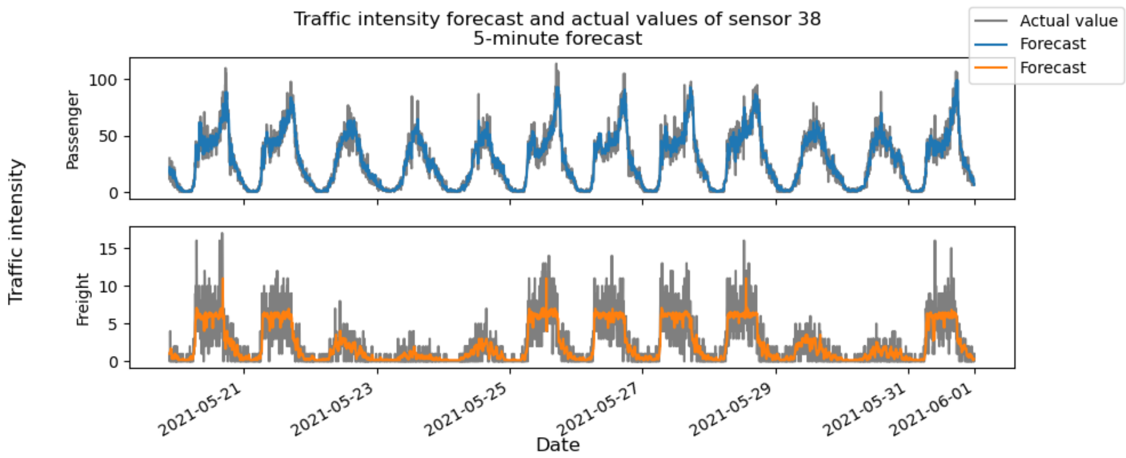


Figure 6.5: Real vs. 5-minute forecast of traffic intensity for sensor 38.

6.2.1.2 Hour-Ahead Forecast

For the hour-ahead forecasts, we can see the same patterns in the forecasts for all the selected sensors. This can be seen in Figures 6.6, 6.7, 6.8, 6.9 and 6.10 for sensors 5, 8, 20, 27, and 38, respectively. First we will look at the forecasts for passenger traffic. As for the 5-minute forecasts, the hour-ahead forecasts for passenger traffic follow the general traffic pattern quite well, but the model has some trouble forecasting when there are high peaks in traffic intensity. For freight traffic, the forecasts are actually leading on the actual traffic intensity. Unlike the 5-minute forecasts, the hour-ahead forecasts seem to be able to better capture the fluctuations in the traffic patterns. Similar to the hour-ahead forecasts of passenger traffic, the model has some difficulty forecasting sudden increases in traffic intensity, and the forecasts are lower than the actual peaks in traffic intensity.

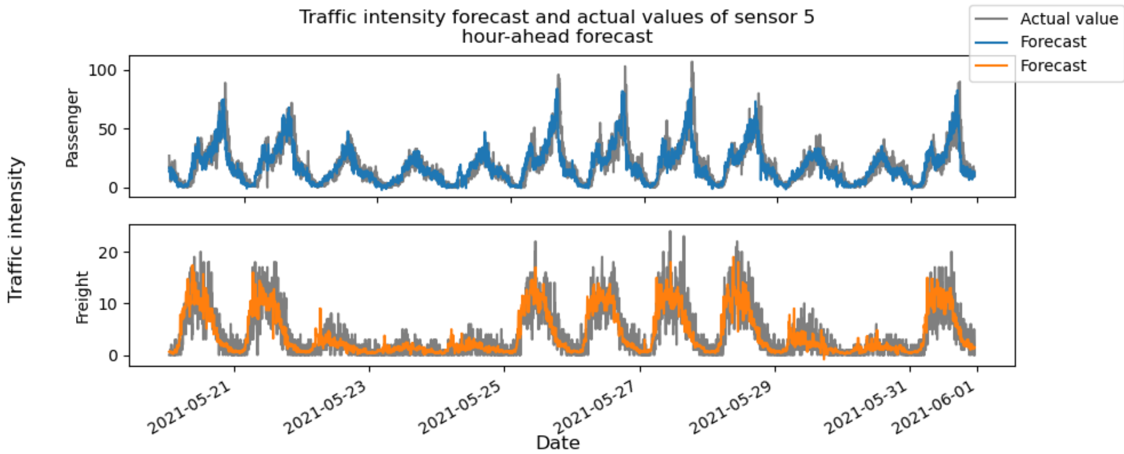


Figure 6.6: Real vs. hour-ahead forecast of traffic intensity for sensor 5.

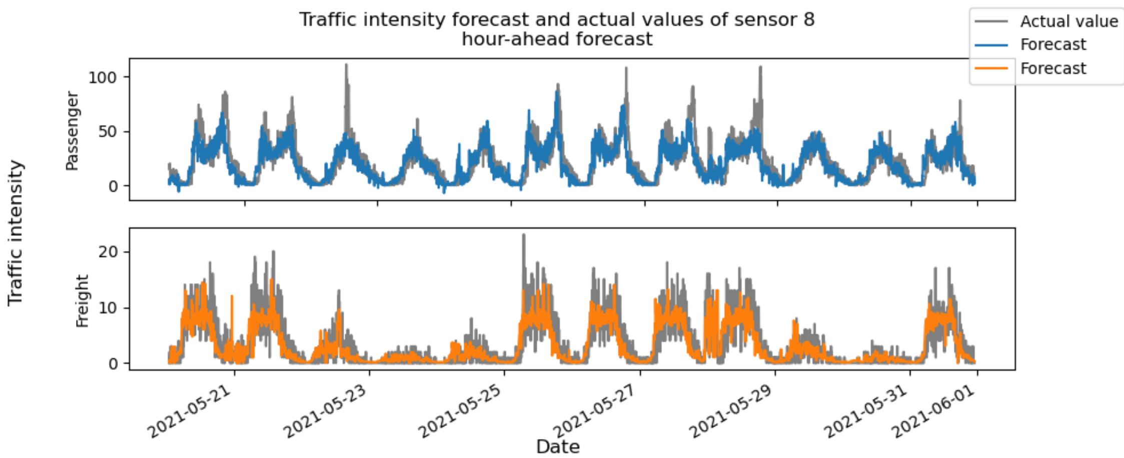


Figure 6.7: Real vs. hour-ahead forecast of traffic intensity for sensor 8.

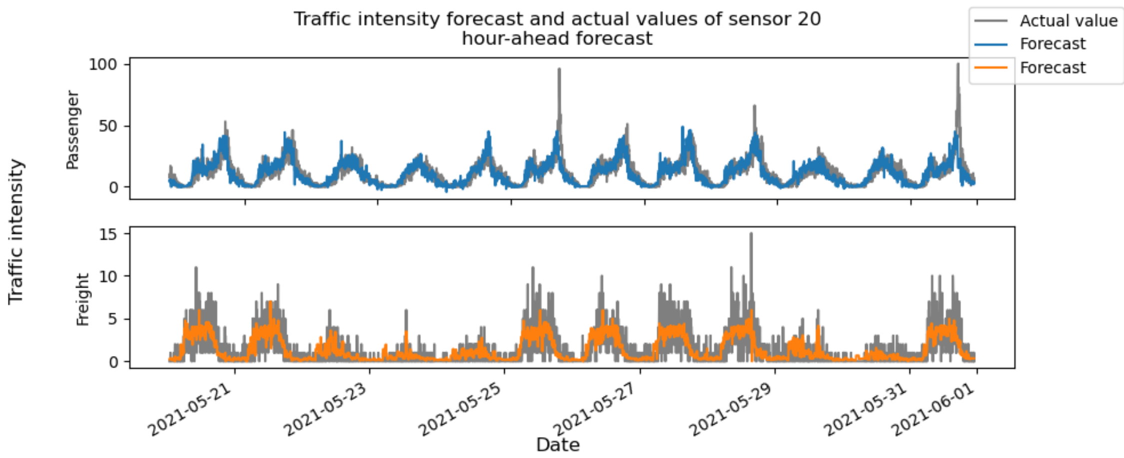


Figure 6.8: Real vs. hour-ahead forecast of traffic intensity for sensor 20.

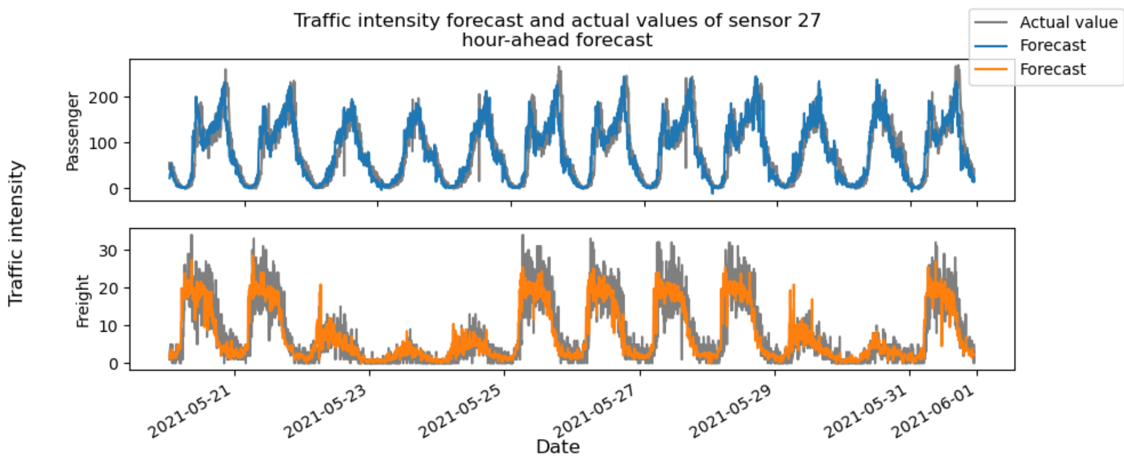


Figure 6.9: Real vs. hour-ahead forecast of traffic intensity for sensor 27.

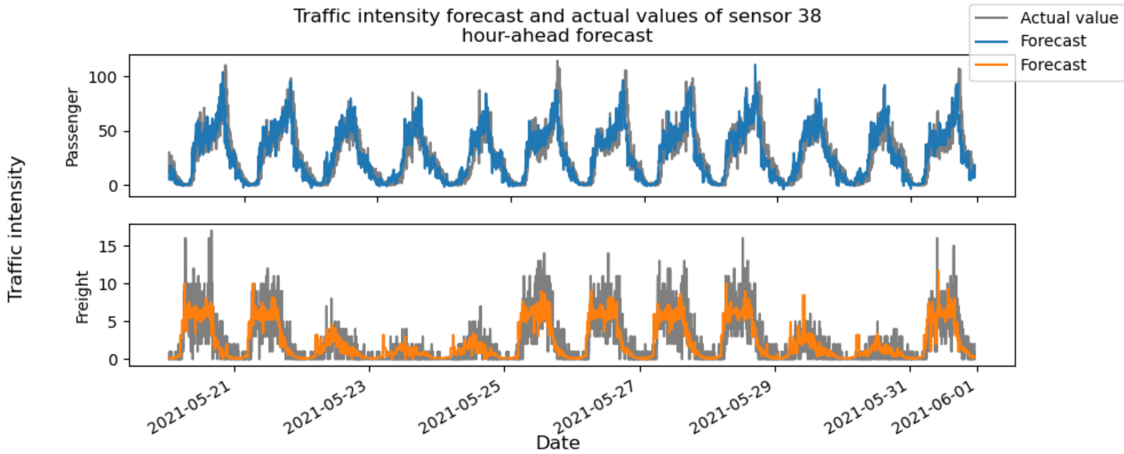
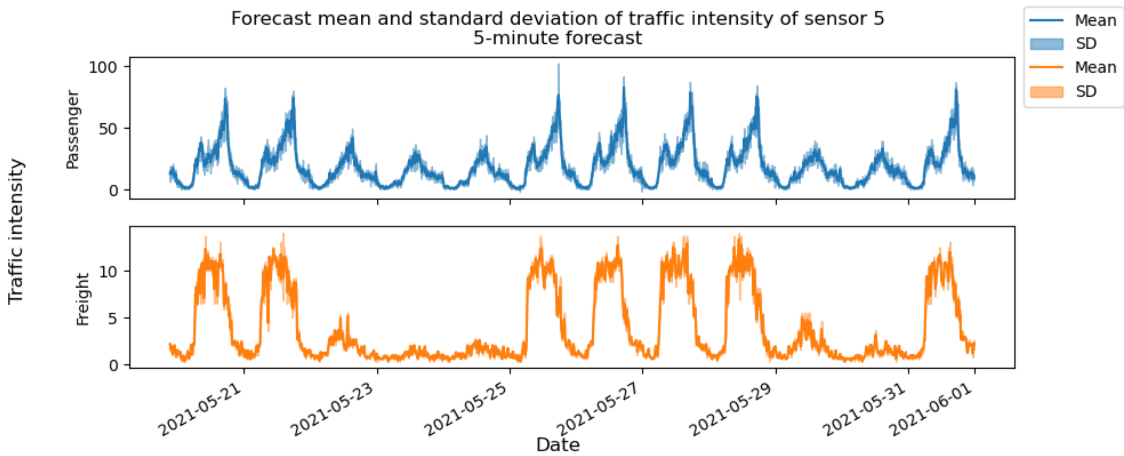


Figure 6.10: Real vs. hour-ahead forecast of traffic intensity for sensor 38.

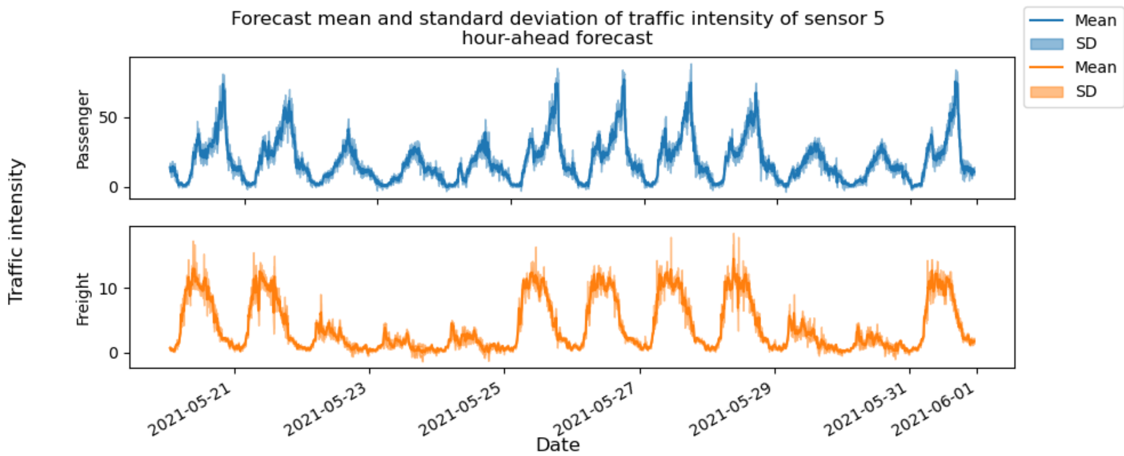
6.2.2 Uncertainty of Forecasts

To determine the uncertainty of the forecasts that the model makes, an XGBoost model was trained 5 times for each of the sensors, with a different random seed, so that each model makes different forecasts for the same point in time. Using the mean and standard deviation of these forecasts, we can gain some insight into the uncertainty of the forecasts.

Figures 6.11, 6.12, 6.13, 6.14, and 6.15 show the forecasting uncertainty for sensors 5, 8, 20, 27, and 38, respectively. In the top figure (a) the mean and standard deviation of the 5-minute forecasts are shown, and in (b) those of the hour-ahead forecast are shown. The standard deviation, the shaded area, around the mean shows how certain or uncertain the model is about forecasts at a time. From these figures, we can see that the 5-minute forecast shows a smaller spread for the standard deviation than that of the hour-ahead forecast for both passenger and freight traffic. Another thing that stands out is that the model is more uncertain of the forecasts when the forecast is fluctuating more or when a peak is forecasted. The uncertainty during fluctuations is especially noticeable for freight traffic, and the uncertainty of peaks is more noticeable for passenger traffic.

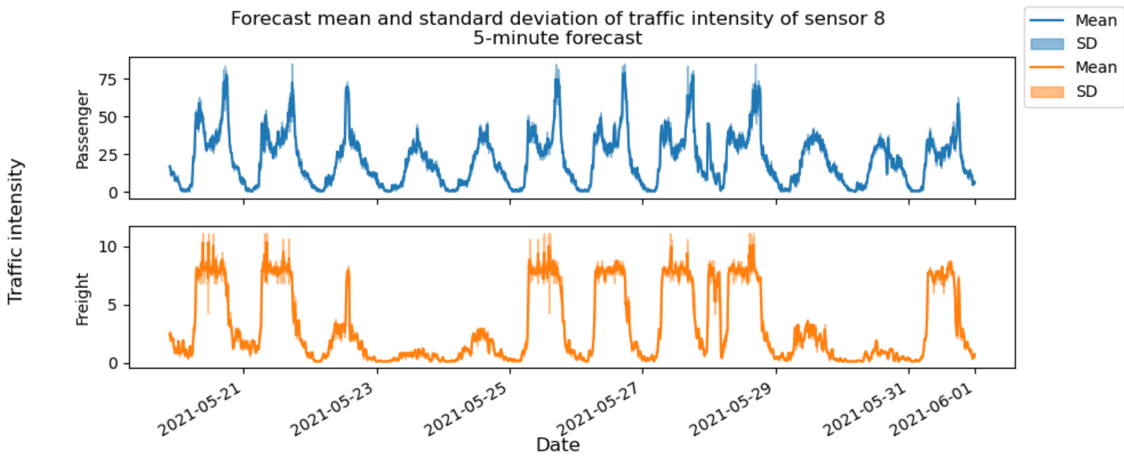


(a) Uncertainty of the 5-minute forecasts of traffic intensity for sensor 5.

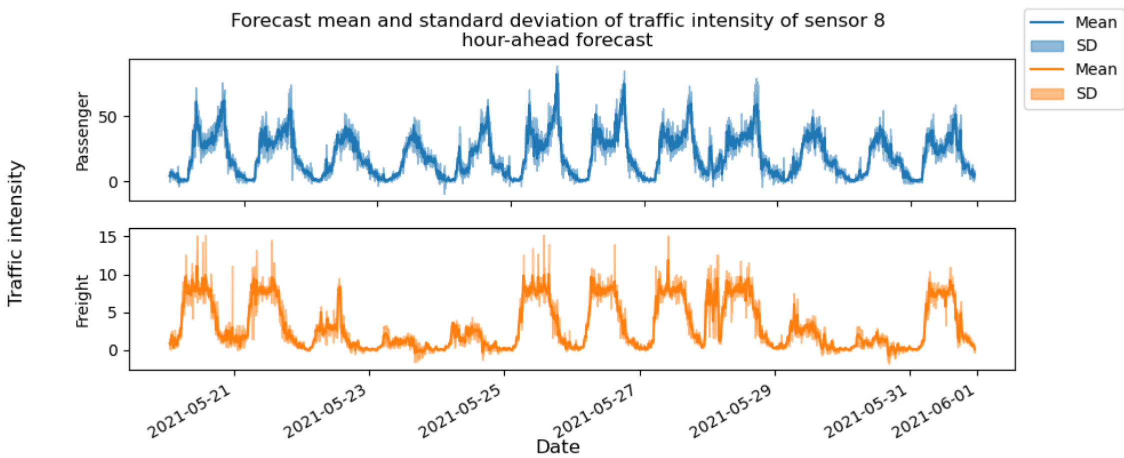


(b) Uncertainty of the hour-ahead forecasts of traffic intensity for sensor 5.

Figure 6.11: Uncertainty of the forecasts of traffic intensity for sensor 5.

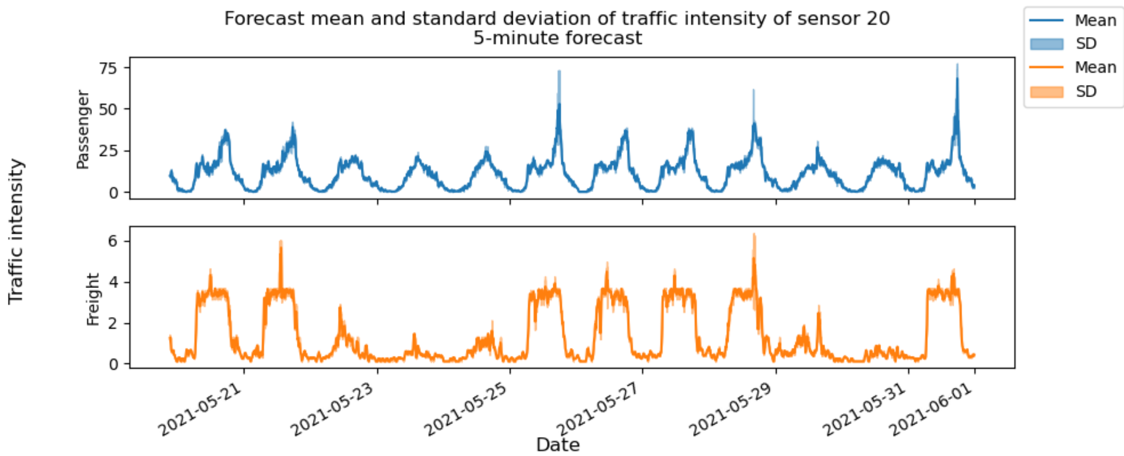


(a) Uncertainty of the 5-minute forecasts of traffic intensity for sensor 8.

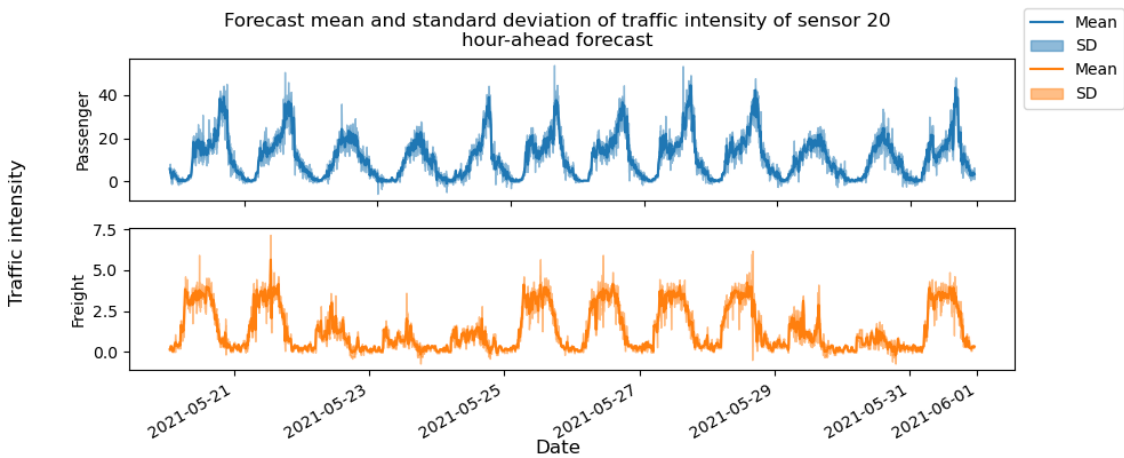


(b) Uncertainty of the hour-ahead forecasts of traffic intensity for sensor 8.

Figure 6.12: Uncertainty of the forecasts of traffic intensity for sensor 8.

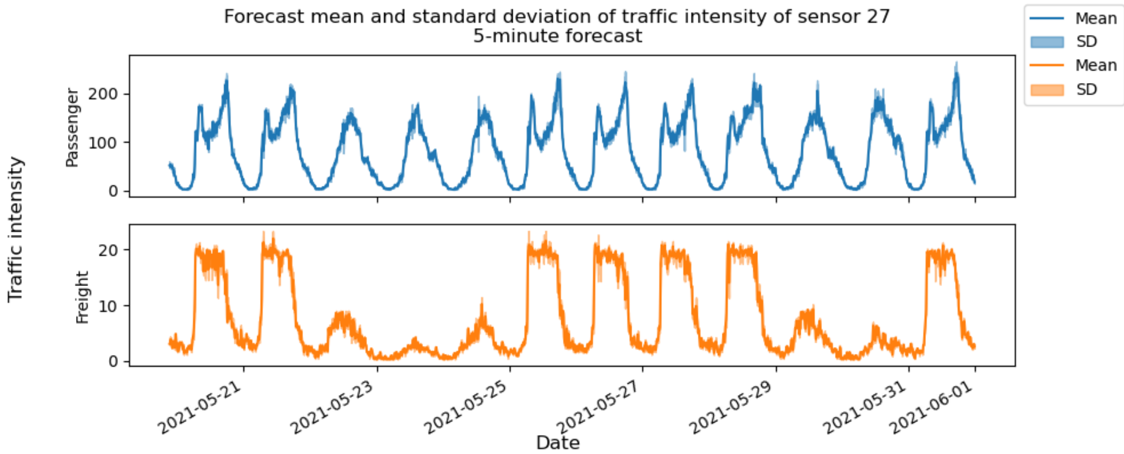


(a) Uncertainty of the 5-minute forecasts of traffic intensity for sensor 20.

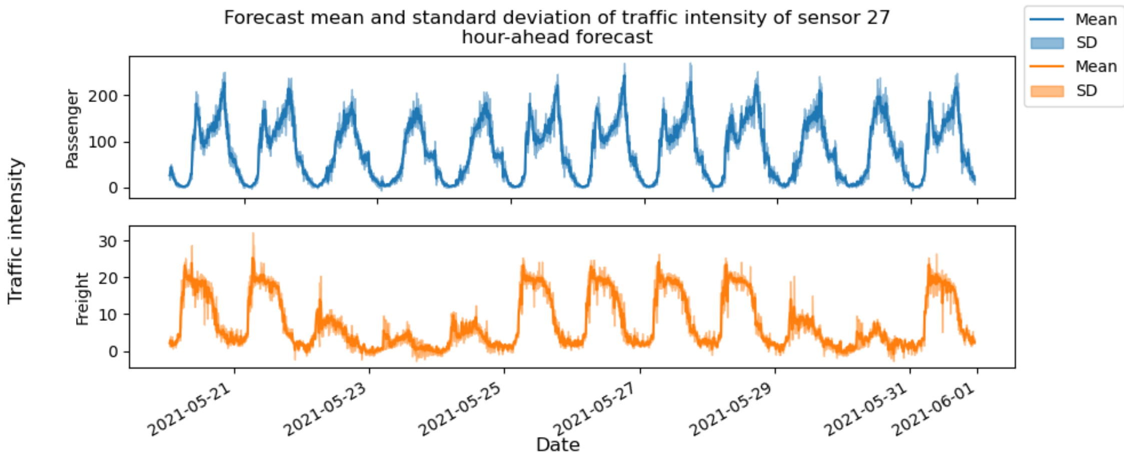


(b) Uncertainty of the hour-ahead forecasts of traffic intensity for sensor 20.

Figure 6.13: Uncertainty of the forecasts of traffic intensity for sensor 20.

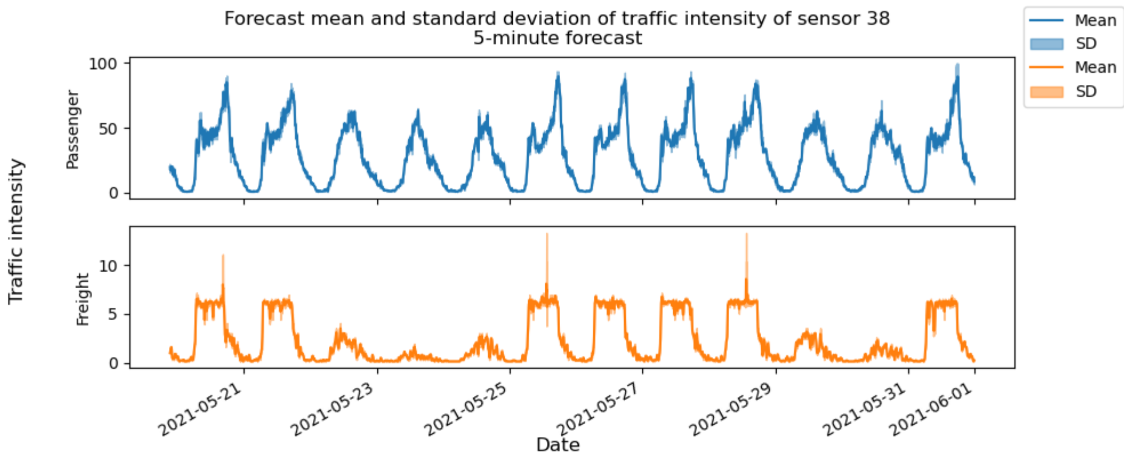


(a) Uncertainty of the 5-minute forecasts of traffic intensity for sensor 27.

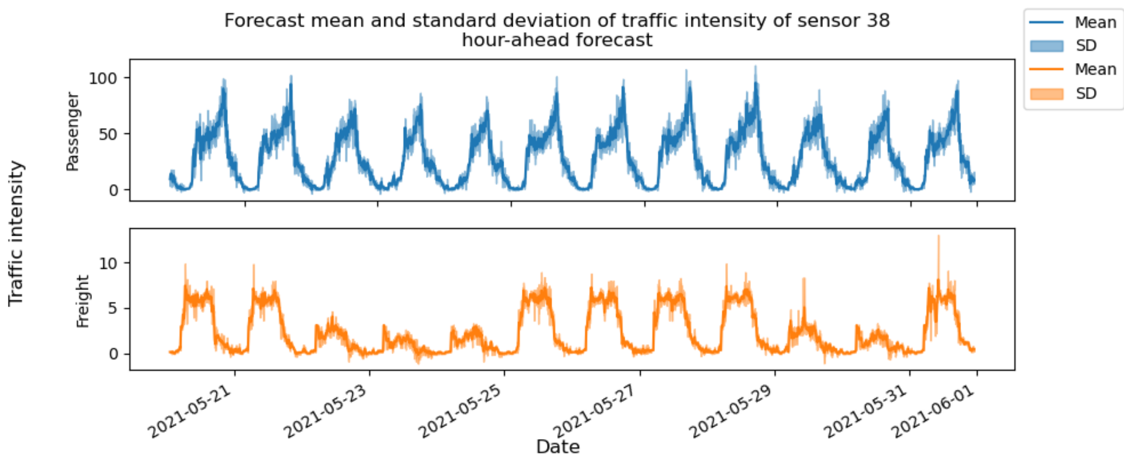


(b) Uncertainty of the hour-ahead forecasts of traffic intensity for sensor 27.

Figure 6.14: Uncertainty of the forecasts of traffic intensity for sensor 27.



(a) Uncertainty of the 5-minute forecasts of traffic intensity for sensor 38.



(b) Uncertainty of the hour-ahead forecasts of traffic intensity for sensor 38.

Figure 6.15: Uncertainty of the forecasts of traffic intensity for sensor 38.

7

Discussion

7.1 Impact of Maintenance

In Section 3.3 the impact of maintenance was studied by inspecting the traffic intensity before and during maintenance by using hypothesis tests. From the traffic intensity inspection, we could see that for sensors where the traffic intensity decreased during Phase 1, this happened more for passenger traffic than for freight traffic. This could be because freight traffic has to continue, even if it will be hindered by road maintenance, while passenger traffic could more easily be rescheduled, especially during weekends. We also saw that for other sensors, those on the A44 and the detours, the traffic intensity increased. So logically, maintenance impacts the quantity of traffic, not only at the maintenance location itself but also on the detours and advisory routes.

The hypothesis tests that were performed also showed some interesting findings. It showed that during Phase 1b, the traffic intensity on the detour was indeed higher than before maintenance. This means that significantly more traffic used the roads on this detour, despite Phase 1b being planned during the night. Rijkswaterstaat should take this into account when planning detours when roads are closed to prevent traffic nuisance to people living next to those detours, especially if detours go through more residential areas,t like the N208 going through a residential area in Sassenheim. Another interesting finding was that during Phase 2, the A44 - part of the advisory route - still had a higher traffic intensity for freight traffic than before maintenance, despite the A4 being open for traffic. This was not the case for passenger traffic. This could indicate that freight traffic does follow advisory routes, and passenger traffic does not. This finding is interesting for Rijkswaterstaat, as this can help them plan these advisory routes better in ways that will facilitate freight traffic.

7.2 Model Performance

From the visualisation of the forecasts in Section 6.2.1, we can see that the models follow the real traffic intensity quite nicely. The models do not seem to have any trouble forecasting the intensity, despite there being changes in the traffic flow due to maintenance. An interesting observation is that for the hour-ahead forecast for freight traffic, the forecast actually is leading to the actual value of the traffic intensity. One possible reason for that is that freight traffic has a very strong weekly pattern, and it seems that the model has learnt this historic pattern very well and takes more recent traffic intensity values less into account. This forecast could be improved by somehow forcing the model to put more emphasis on more recent values of traffic intensity.

7.3 Graph Neural Network

In Section 6 we saw that for the 5-minute forecast for passenger traffic, the GNN had the lowest mean error, and for freight traffic, its errors were very close to that of XGBoost. For the hour-ahead forecast, the GNN had worse errors than that of XGBoost. This means that the GNN, despite having more information available than XGBoost, does not necessarily have better forecasts than XGBoost. But the GNN does have lower errors than the Baseline model, Holt-Winters and the Transformer. The results also show that the GNN might be better at very short-term forecasts than more long-term forecasts.

7.4 Future Research

For future research, it could be interesting to have a more in-depth look into the capabilities of the GNN and the transformer to predict multiple time steps at once, as both of these models are capable of making multi-step ahead predictions. Using these multi-step ahead predictions, it would not be necessary to have two separate models to predict a 5-minute forecast and an hour-ahead forecast. A challenge in this would be to find optimal hyperparameters for the models to accurately forecast the different time steps. Another interesting topic for future research is to study road maintenance that took place on different roads, to investigate whether the forecasting models have similar performance when applied to a new case study.

8

Conclusion

The goal of this research was to study the impact of maintenance on traffic, specifically freight traffic, and whether traffic intensity can still accurately be forecasted, despite the impact maintenance might have. Another goal was to implement a model that incorporates a graph neural network which utilises data from multiple sensors at once to forecast the traffic intensity. The research questions there were developed to study these goals were:

- Does road maintenance have a significant impact on the intensity of freight traffic?
- Can models accurately forecast the traffic intensity during maintenance?
- Can a model that incorporates a graph neural network and utilises data from multiple sensors at once forecast traffic intensity better than other models that use data from only one sensor?

To answer the first question, the traffic intensity was visualised and hypothesis tests were performed. It was found that there were indeed significant changes in the traffic intensity for freight traffic, especially on the detour and advisory route. Rijkswaterstaat can use this information to plan detours and advisory routes in such a way as to facilitate freight traffic and decrease nuisance for residential areas.

The second question can be answered by looking at the forecasts of the XGBoost models and comparing them to the real traffic intensity. The visualisations of the forecasts showed that the models can accurately predict traffic intensity, even during maintenance. So it can be concluded that the XGBoost model is quite robust to changes in traffic flow that were caused by maintenance.

To answer the final question, the errors of the different methods were compared. These showed that the GNN is better than the baseline model, Holt-Winters exponential smoothing, and the Transformer for the 5-minute and the hour-ahead forecasts for both types of traffic. However, the GNN was only better than XGBoost at the 5-minute forecast for passenger traffic. For the 5-minute forecasts for freight traffic, the errors were very close to that of XGBoost, but were slightly higher. This indicates that the GNN is good at making short-term forecasts, like the 5-minute forecast.

In conclusion, this research found that there is a significant change in traffic intensity during maintenance, especially for freight traffic. During maintenance, XGBoost can accurately forecast the traffic intensity, better than the GNN. So the advice to Rijkswaterstaat is to implement XGBoost models if they want to predict traffic intensity.

References

[1] W. Alajali, W. Zhou, and S. Wen. Traffic flow prediction for road intersection safety. In *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pages 812–820. IEEE, 2018. 5

[2] S. AlKheder, W. Alkhamees, R. Almutairi, and M. Alkhedher. Bayesian combined neural network for traffic volume short-term forecasting at adjacent intersections. *Neural Computing and Applications*, 33:1785–1836, 2021. 8

[3] L. Cai, K. Janowicz, G. Mai, B. Yan, and R. Zhu. Traffic transformer: Capturing the continuity and periodicity of time series for traffic forecasting. *Transactions in GIS*, 24(3):736–755, 2020. 7

[4] C. Chatfield, A. B. Koehler, J. K. Ord, and R. D. Snyder. A new look at models for exponential smoothing. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 50(2):147–159, 2001. 5

[5] J. Chen, S. Liao, J. Hou, K. Wang, and J. Wen. Gst-gcn: a geographic-semantic-temporal graph convolutional network for context-aware traffic flow prediction on graph sequences. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1604–1609. IEEE, 2020. 6

[6] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016. 26

[7] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014. 28

[8] Z. Diao, X. Wang, D. Zhang, Y. Liu, K. Xie, and S. He. Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 890–897, 2019. 6

[9] X. Dong, T. Lei, S. Jin, and Z. Hou. Short-term traffic flow prediction based on xgboost. In *2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS)*, pages 854–859. IEEE, 2018. 5

[10] E. S. Gardner Jr. Exponential smoothing: The state of the art—part ii. *International journal of forecasting*, 22(4):637–666, 2006. 5

[11] K. Guo, Y. Hu, Z. Qian, H. Liu, K. Zhang, Y. Sun, J. Gao, and B. Yin. Optimized graph convolution recurrent neural network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems*, 22(2):1138–1149, 2020. 6

[12] C. He, D. Wang, Y. Yu, Z. Cai, et al. A hybrid deep learning model for link dynamic vehicle count forecasting with bayesian optimization. *Journal of Advanced Transportation*, 2023, 2023. 8

[13] R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2018. 25

[14] W. Jiang and J. Luo. Graph neural network for traffic forecasting: A survey. *Expert Systems with Applications*, page 117921, 2022. 29

[15] K. Jin, J. Wi, E. Lee, S. Kang, S. Kim, and Y. Kim. Trafficbert: Pre-trained model with large-scale data for long-range traffic flow forecasting. *Expert Systems with Applications*, 186:115738, 2021. 7

[16] T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016. 28

[17] X. Kong, W. Xing, X. Wei, P. Bao, J. Zhang, and W. Lu. Stgat: Spatial-temporal graph attention networks for traffic flow forecasting. *IEEE Access*, 8:134363–134372, 2020. 6

[18] B. Lartey, A. Homaifar, A. Girma, A. Karimoddini, and D. Opoku. Xgboost: a tree-based approach for traffic volume prediction. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1280–1286. IEEE, 2021. 5

[19] T. Lumley, P. Diehr, S. Emerson, and L. Chen. The importance of the normality assumption in large public health data sets. *Annual review of public health*, 23(1):151–169, 2002. 21

[20] S. M. Lundberg and S.-I. Lee. A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 4765–4774. Curran Associates, Inc., 2017. 27

[21] T. Ma, C. Antoniou, and T. Toledo. Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. *Transportation Research Part C: Emerging Technologies*, 111:352–372, 2020. 8

[22] M. L. Miguel, M. C. Penna, J. C. Nievola, and M. E. Pellenz. New models for long-term internet traffic forecasting using artificial neural networks and flow based information. In *2012 IEEE Network Operations and Management Symposium*, pages 1082–1088. IEEE, 2012. 5

[23] S. Mohanty, A. Pozdnukhov, and M. Cassidy. Region-wide congestion prediction and control using deep learning. *Transportation Research Part C: Emerging Technologies*, 116:102624, 2020. 6

[24] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu. Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*, 2016. 7

[25] S. Shleifer, C. McCreery, and V. Chitters. Incrementally improving graph wavenet performance on traffic prediction. *arXiv preprint arXiv:1912.07390*, 2019. 7

[26] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. 30, 36

[27] J. Wang, W. Deng, and Y. Guo. New bayesian combination method for short-term traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 43:79–94, 2014. 8

[28] X. Wang, Y. Ma, Y. Wang, W. Jin, X. Wang, J. Tang, C. Jia, and J. Yu. Traffic flow prediction via spatial temporal graph neural network. In *Proceedings of The Web Conference 2020*, pages 1082–1092, 2020. 6, 27

[29] Z.-H. Wang, C.-Y. Lu, B. Pu, G.-W. Li, and Z.-J. Guo. Short-term forecast model of vehicles volume based on arima seasonal model and holt-winters. In *ITM Web of Conferences*, volume 12, page 04028. EDP Sciences, 2017. 5

[30] P. R. Winters. Forecasting sales by exponentially weighted moving averages. *Management science*, 6(3):324–342, 1960. 25

[31] G. Woo, C. Liu, D. Sahoo, A. Kumar, and S. Hoi. Etsformer: Exponential smoothing transformers for time-series forecasting. *arXiv preprint arXiv:2202.01381*, 2022. 8

[32] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang. A hybrid deep learning based traffic flow prediction method and its understanding. *Transportation Research Part C: Emerging Technologies*, 90:166–180, 2018. 8

[33] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang. Connecting the dots: Multivariate time series forecasting with graph neural networks. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 753–763, 2020. 6

[34] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang. Graph wavenet for deep spatial-temporal graph modeling. *arXiv preprint arXiv:1906.00121*, 2019. 7

[35] H. Yang, X. Li, W. Qiang, Y. Zhao, W. Zhang, and C. Tang. A network traffic forecasting method based on sa optimized arima–bp neural network. *Computer Networks*, 193:108102, 2021. 8

[36] B. Yu, H. Yin, and Z. Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875*, 2017. 7, 29

[37] L. Zhang, Y. Yang, Y. Deng, and H. Kang. Forecasting of road traffic flow based on harris hawk optimization and xgboost. *Journal of Advances in Mathematics and Computer Science*, 37(2):21–29, 2022. 6

Appendix A

Case Study

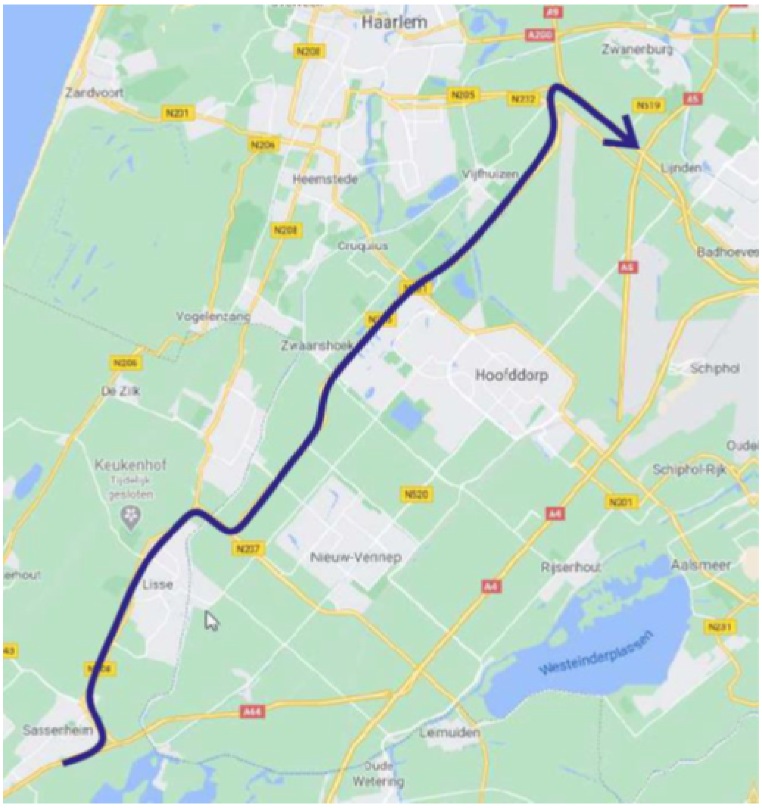


Figure A.1: Map showing the detour used during Phase 1b for traffic coming from the A44 Den Haag/Wassenaar going towards Amsterdam.

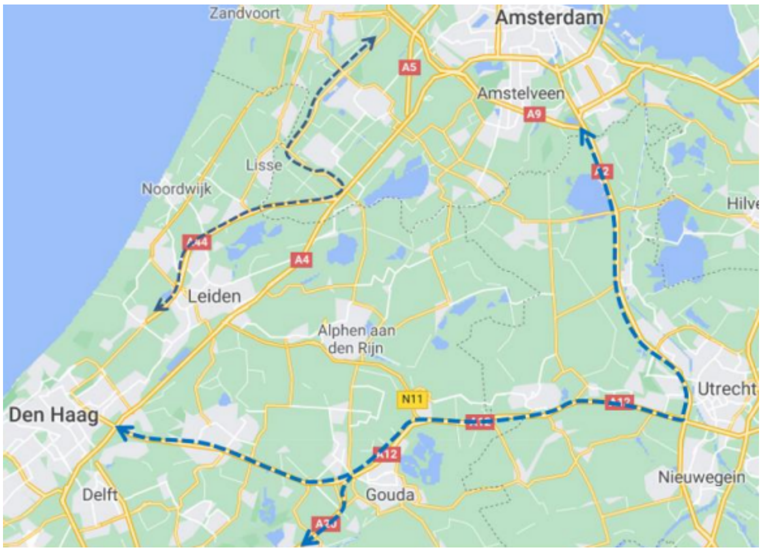


Figure A.2: Map showing the advisory routes during Phase 2.

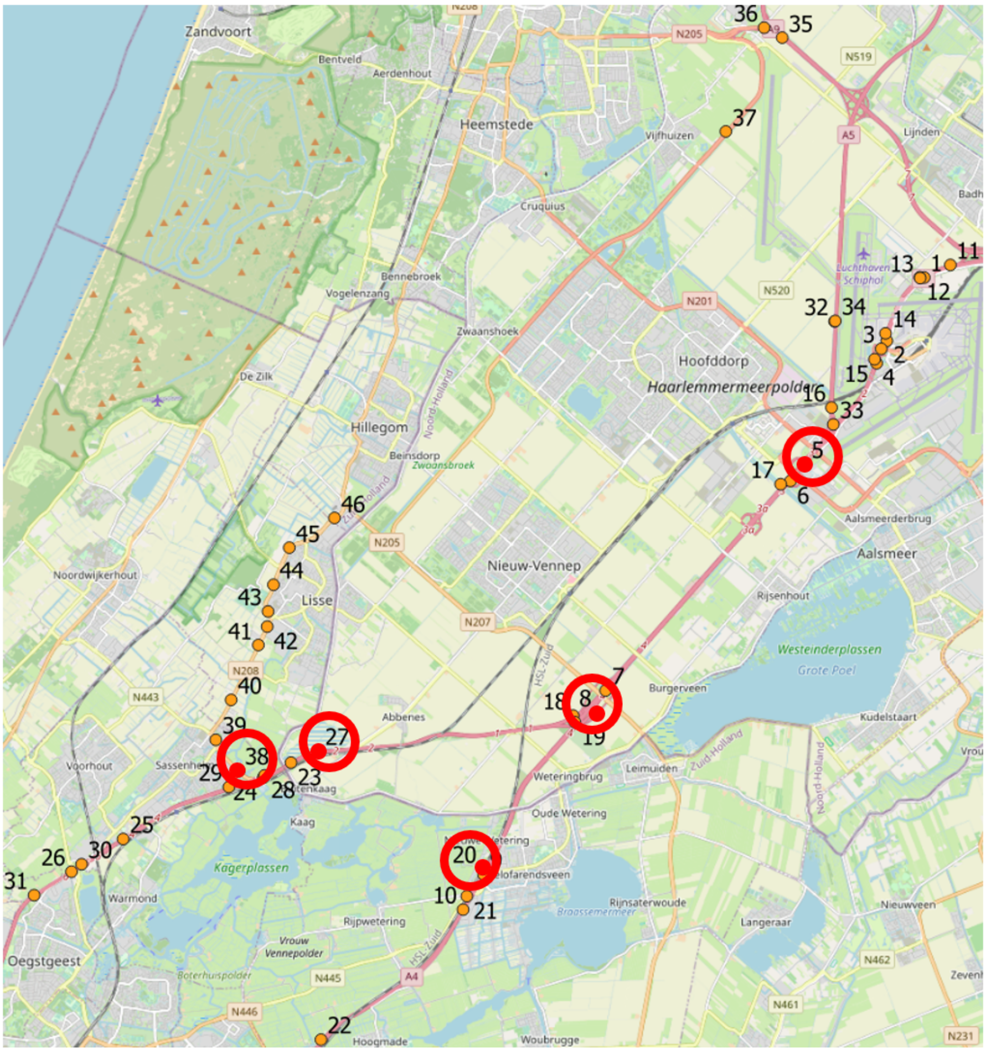
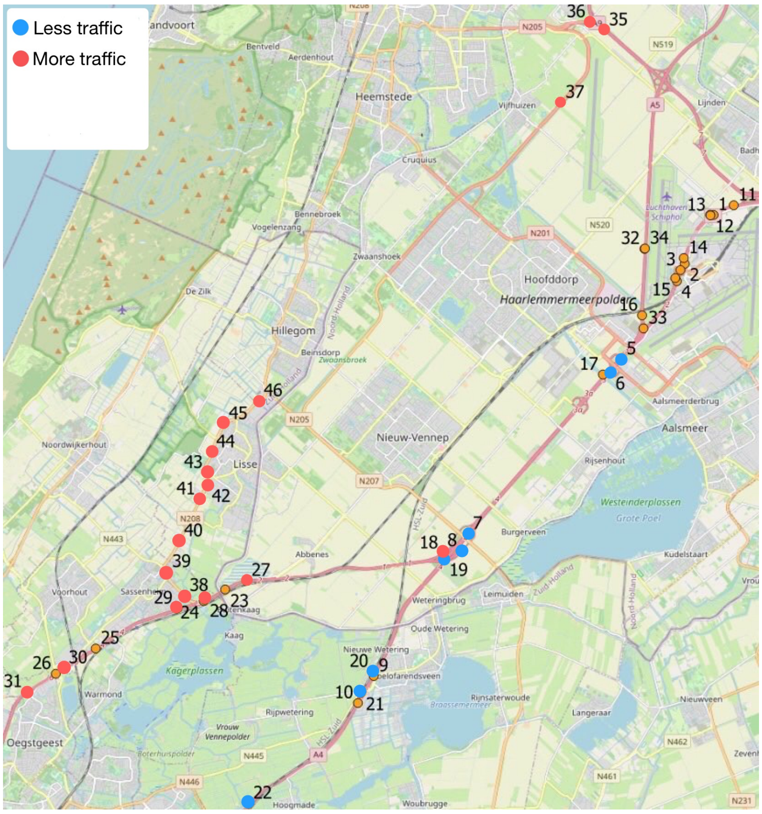
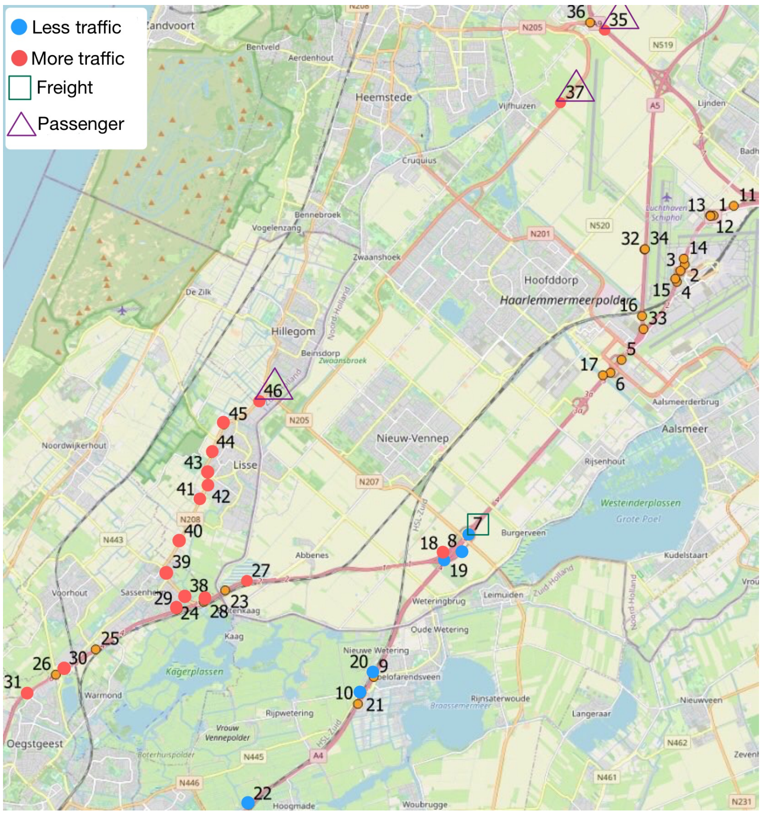


Figure A.3: Map showing the locations of the sensors selected for analysis.

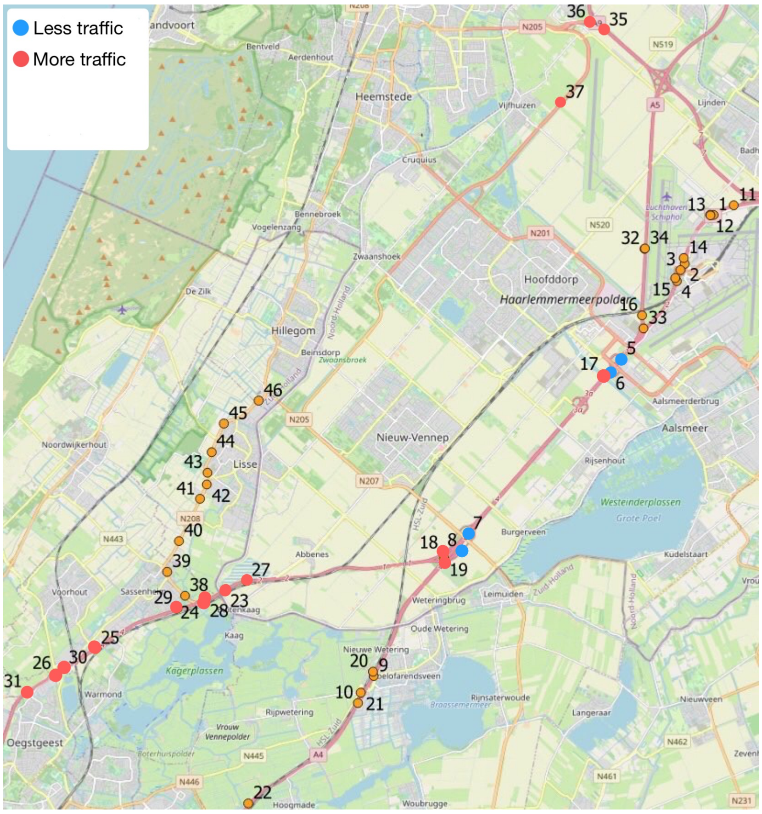


(a) Selected sensors and their hypothesis.

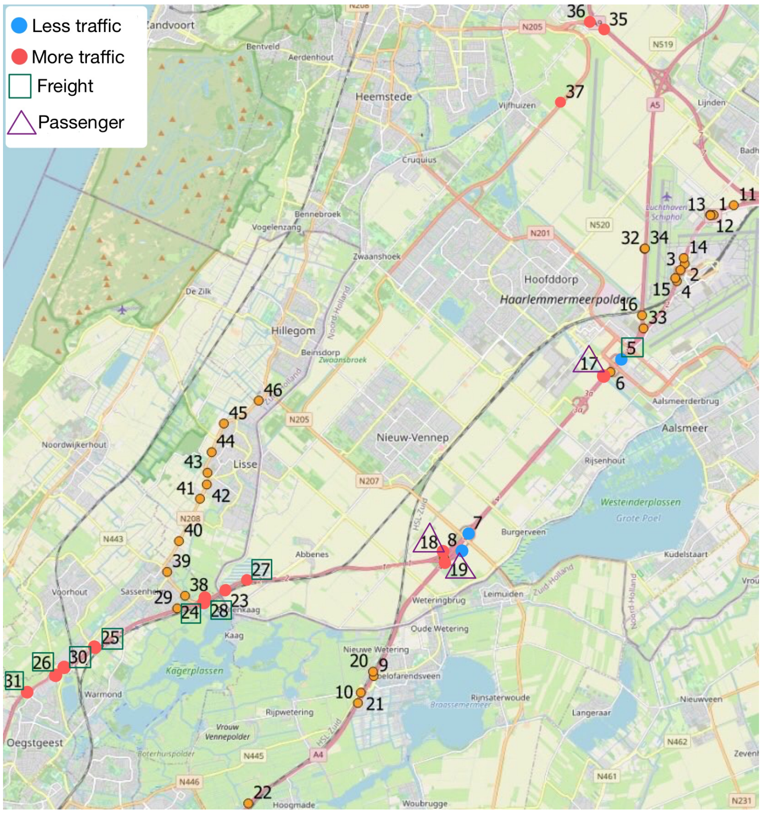


(b) Results of hypothesis testing.

Figure A.4: Sensors selected for hypothesis testing for Phase 1.



(a) Selected sensors and their hypothesis.



(b) Results of hypothesis testing.

Figure A.5: Sensors selected for hypothesis testing for Phase 2.

Table A.1: p-value sensor for the Mann–Whitney U test and the t-test to compare traffic intensity before and during Phase 1 of the maintenance.

	Freight traffic		Passenger traffic	
sensor	p-value MWU test	p-value t-test	p-value MWU test	p-value t-test
5	0.4771	0.2324	0.3006	0.28
6	0.379	0.1907	0.6149	0.4972
7	<**0.01**	<**0.01**	0.2158	<**0.01**
8	<**0.01**	<**0.01**	<**0.01**	<**0.01**
18	<**0.01**	<**0.01**	<**0.01**	<**0.01**
19	<**0.01**	<**0.01**	<**0.01**	<**0.01**
20	<**0.01**	<**0.01**	<**0.01**	<**0.01**
21	<**0.01**	<**0.01**	<**0.01**	<**0.01**
22	<**0.01**	<**0.01**	<**0.01**	<**0.01**
27	<**0.01**	<**0.01**	<**0.01**	<**0.01**
28	<**0.01**	<**0.01**	<**0.01**	<**0.01**
29	<**0.01**	<**0.01**	<**0.01**	<**0.01**
30	<**0.01**	<**0.01**	<**0.01**	<**0.01**
31	<**0.01**	<**0.01**	<**0.01**	<**0.01**
35	0.576	0.6317	<**0.01**	<**0.01**
36	<**0.01**	<**0.01**	<**0.01**	<**0.01**
37	0.856	0.5959	<**0.01**	<**0.01**
38	<**0.01**	<**0.01**	<**0.01**	<**0.01**
39	<**0.01**	<**0.01**	<**0.01**	<**0.01**
40	<**0.01**	<**0.01**	<**0.01**	<**0.01**
41	<**0.01**	<**0.01**	<**0.01**	<**0.01**
42	<**0.01**	<**0.01**	<**0.01**	<**0.01**
43	<**0.01**	<**0.01**	<**0.01**	<**0.01**
44	<**0.01**	<**0.01**	<**0.01**	<**0.01**
45	<**0.01**	<**0.01**	<**0.01**	<**0.01**
46	0.1556	0.1667	<**0.01**	<**0.01**

Table A.2: p-value sensor for the Mann–Whitney U test and the t-test to compare traffic intensity before and during Phase 2 of the maintenance.

	Freight traffic		Passenger traffic	
sensor	p-value MWU test	p-value t-test	p-value MWU test	p-value t-test
5	<**0.01**	<**0.01**	0.0929	0.5608
6	0.6911	0.5089	<**0.01**	0.1076

Continued on next page

Table A.2: p-value sensor for the Mann–Whitney U test and the t-test to compare traffic intensity before and during Phase 2 of the maintenance.

sensor	Freight traffic p-value MWU test	p-value t-test	Passenger traffic p-value MWU test	p-value t-test
7	<**0.01**	<**0.01**	<**0.01**	<**0.01**
8	<**0.01**	<**0.01**	<**0.01**	<**0.01**
17	0.9519	0.7612	<**0.01**	<**0.01**
18	0.0625	<**0.01**	<**0.01**	**0.0116**
19	0.4332	0.6973	<**0.01**	<**0.01**
23	<**0.01**	<**0.01**	**0.0251**	<**0.01**
24	<**0.01**	<**0.01**	0.1509	<**0.01**
25	<**0.01**	<**0.01**	0.423	<**0.01**
26	<**0.01**	<**0.01**	0.0532	<**0.01**
27	<**0.01**	<**0.01**	0.0664	0.6829
28	<**0.01**	<**0.01**	0.2936	0.335
29	0.0627	<**0.01**	**0.0289**	0.9408
30	<**0.01**	<**0.01**	0.2914	0.3768
31	**0.012**	**0.0231**	0.6147	0.1484

Appendix B

Implementation Details

B.1 Holt-Winters

Table B.1: Parameters used by the Holt-Winters method for freight traffic.

sensor	initial α	initial β	α	β	γ
1	1.1972	0.0000	0.0588	0.0000	0.0001
2	0.9032	0.0000	0.0635	0.0000	0.0000
3	10.2227	0.0205	0.2849	0.0020	0.0538
4	1.3483	0.0001	0.0719	0.0000	0.0001
5	6.8738	0.0004	0.1222	0.0000	0.0000
6	2.3069	0.0000	0.1005	0.0000	0.0016
7	1.0835	0.0000	0.0969	0.0000	0.0008
8	2.7323	0.0001	0.1168	0.0001	0.0022
9	2.1390	0.0001	0.0989	0.0000	0.0050
10	1.3397	0.0000	0.0970	0.0000	0.0003
11	26.3092	0.1443	0.3140	0.0106	0.0491
12	13.0883	0.0030	0.1968	0.0026	0.0553
13	57.0143	0.3599	0.3950	0.0148	0.0927
14	12.7726	0.0031	0.2463	0.0004	0.0509
15	1.2771	0.0001	0.0601	0.0000	0.0002
16	0.3963	0.0000	0.0980	0.0000	0.0000
17	2.0702	0.0000	0.0884	0.0000	0.0000
18	1.4799	0.0000	0.0837	0.0000	0.0023
19	4.1871	-0.0000	0.1373	0.0002	0.0089
20	1.3224	0.0000	0.0840	0.0000	0.0001
21	1.4128	0.0000	0.0874	0.0000	0.0006

Continued on next page

Table B.1: Parameters used by the Holt-Winters method for freight traffic.

sensor	initial α	initial β	α	β	γ
22	2.2892	0.0000	0.1022	0.0000	0.0018
23	10.6214	0.0006	0.1491	0.0002	0.0034
24	6.8618	0.0000	0.1446	0.0004	0.0178
25	7.9209	-0.0001	0.1503	0.0000	0.0001
26	7.2917	0.0000	0.1338	0.0000	0.0000
27	8.3651	-0.0012	0.1630	0.0003	0.0016
28	7.8401	0.0041	0.1480	0.0005	0.0085
29	8.3269	-0.0008	0.1529	0.0004	0.0433
30	3.9537	0.0011	0.1565	0.0007	0.0210
31	0.8587	0.0000	0.0797	0.0000	0.0000
32	6.1680	0.0002	0.1328	0.0003	0.0065
33	1.0407	0.0000	0.1237	0.0000	0.0000
34	5.1494	-0.0002	0.1133	0.0004	0.0104
35	1.4435	0.0000	0.0863	0.0000	0.0002
36	9.5605	-0.0052	0.1646	0.0019	0.0002
37	4.1692	-0.0000	0.1566	0.0001	0.0044
38	2.1162	0.0000	0.1049	0.0000	0.0010
39	3.0755	0.0000	0.1112	0.0000	0.0022
40	2.4781	0.0001	0.0959	0.0000	0.0023
41	1.7285	0.0000	0.0946	0.0000	0.0002
42	1.4641	0.0000	0.1043	0.0000	0.0001
43	1.5895	0.0001	0.0914	0.0000	0.0008
44	1.4719	0.0000	0.0877	0.0000	0.0010
45	2.5649	0.0001	0.1050	0.0000	0.0006
46	3.3268	0.0000	0.1147	0.0001	0.0026

Table B.2: Parameters used by the Holt-Winters method for passenger traffic.

sensor	initial α	initial β	α	β	γ
1	1.1972	0.0000	0.0588	0.0000	0.0001
2	0.9032	0.0000	0.0635	0.0000	0.0000
3	10.2227	0.0205	0.2849	0.0020	0.0538
4	1.3483	0.0001	0.0719	0.0000	0.0001
5	6.8738	0.0004	0.1222	0.0000	0.0000
6	2.3069	0.0000	0.1005	0.0000	0.0016
7	1.0835	0.0000	0.0969	0.0000	0.0008
8	2.7323	0.0001	0.1168	0.0001	0.0022

Continued on next page

Table B.2: Parameters used by the Holt-Winters method for passenger traffic.

sensor	initial α	initial β	α	β	γ
9	2.1390	0.0001	0.0989	0.0000	0.0050
10	1.3397	0.0000	0.0970	0.0000	0.0003
11	26.3092	0.1443	0.3140	0.0106	0.0491
12	13.0883	0.0030	0.1968	0.0026	0.0553
13	57.0143	0.3599	0.3950	0.0148	0.0927
14	12.7726	0.0031	0.2463	0.0004	0.0509
15	1.2771	0.0001	0.0601	0.0000	0.0002
16	0.3963	0.0000	0.0980	0.0000	0.0000
17	2.0702	0.0000	0.0884	0.0000	0.0000
18	1.4799	0.0000	0.0837	0.0000	0.0023
19	4.1871	-0.0000	0.1373	0.0002	0.0089
20	1.3224	0.0000	0.0840	0.0000	0.0001
21	1.4128	0.0000	0.0874	0.0000	0.0006
22	2.2892	0.0000	0.1022	0.0000	0.0018
23	10.6214	0.0006	0.1491	0.0002	0.0034
24	6.8618	0.0000	0.1446	0.0004	0.0178
25	7.9209	-0.0001	0.1503	0.0000	0.0001
26	7.2917	0.0000	0.1338	0.0000	0.0000
27	8.3651	-0.0012	0.1630	0.0003	0.0016
28	7.8401	0.0041	0.1480	0.0005	0.0085
29	8.3269	-0.0008	0.1529	0.0004	0.0433
30	3.9537	0.0011	0.1565	0.0007	0.0210
31	0.8587	0.0000	0.0797	0.0000	0.0000
32	6.1680	0.0002	0.1328	0.0003	0.0065
33	1.0407	0.0000	0.1237	0.0000	0.0000
34	5.1494	-0.0002	0.1133	0.0004	0.0104
35	1.4435	0.0000	0.0863	0.0000	0.0002
36	9.5605	-0.0052	0.1646	0.0019	0.0002
37	4.1692	-0.0000	0.1566	0.0001	0.0044
38	2.1162	0.0000	0.1049	0.0000	0.0010
39	3.0755	0.0000	0.1112	0.0000	0.0022
40	2.4781	0.0001	0.0959	0.0000	0.0023
41	1.7285	0.0000	0.0946	0.0000	0.0002
42	1.4641	0.0000	0.1043	0.0000	0.0001
43	1.5895	0.0001	0.0914	0.0000	0.0008
44	1.4719	0.0000	0.0877	0.0000	0.0010
45	2.5649	0.0001	0.1050	0.0000	0.0006
46	3.3268	0.0000	0.1147	0.0001	0.0026

B.2 XGBoost

Table B.3: Features used by the XGBoost models for the 5-minute forecast for freight traffic.

sensor	Model features
1	[SMA_12]
2	[SMA_11]
3	[SMA_4]
4	[hour, SMA_12]
5	[SMA_5]
6	[SMA_9]
7	[SMA_11]
8	[SMA_7]
9	[SMA_10]
10	[SMA_9]
11	[SMA_2]
12	[SMA_5]
13	[SMA_2]
14	[SMA_4]
15	[SMA_12]
16	[SMA_12]
17	[SMA_10]
18	[SMA_6]
19	[SMA_5]
20	[SMA_9]
21	[SMA_10]
22	[SMA_6]
23	[SMA_4]
24	[SMA_4]
25	[SMA_4]
26	[SMA_9]
27	[SMA_4]
28	[SMA_4]
29	[SMA_4]
30	[SMA_5]
31	[SMA_9, hour]
32	[SMA_5]
33	[SMA_6]
34	[SMA_8, hour, SMA_7, SMA_2, SMA_6, SMA_12]
35	[SMA_11]
36	[SMA_6]
37	[SMA_6]
38	[SMA_6]
39	[SMA_6]

Continued on next page

Table B.3: Features used by the XGBoost models for the 5-minute forecast for freight traffic.

sensor	Model features
40	[SMA_10]
41	[SMA_9]
42	[SMA_6]
43	[SMA_8]
44	[SMA_8]
45	[SMA_7]
46	[SMA_6]

Table B.4: Features used by the XGBoost models for the 5-minute forecast for passenger traffic.

sensor	Model features
1	[SMA_3]
2	[SMA_3]
3	[SMA_2]
4	[SMA_3]
5	[SMA_4, SMA_3, SMA_5, SMA_2, SMA_6, hour]
6	[SMA_5]
7	[SMA_5]
8	[SMA_3]
9	[SMA_6]
10	[SMA_6]
11	[SMA_2]
12	[SMA_4]
13	[SMA_2]
14	[SMA_2, SMA_3, SMA_4, gem_intensiteit_1, hour, gem_intensiteit_12]
15	[SMA_4]
16	[SMA_8, SMA_11]
17	[SMA_4]
18	[SMA_5]
19	[SMA_4]
20	[SMA_4]
21	[SMA_6]
22	[SMA_4]
23	[SMA_4, SMA_5, SMA_2, SMA_6, SMA_3, hour]
24	[SMA_3, SMA_4, SMA_2, hour, SMA_6, gem_intensiteit_1]
25	[SMA_2]

Continued on next page

Table B.4: Features used by the XGBoost models for the 5-minute forecast for passenger traffic.

sensor	Model features
26	[SMA_2]
27	[SMA_2]
28	[SMA_4]
29	[SMA_4, SMA_2, SMA_3, SMA_5, hour, gem_intensiteit_1]
30	[SMA_3, SMA_5, SMA_4, SMA_2, SMA_6, hour]
31	[SMA_4]
32	[SMA_2, SMA_3, hour, SMA_5, gem_intensiteit_1, SMA_6]
33	[SMA_2]
34	[SMA_2, SMA_3, hour, gem_intensiteit_1, gem_intensiteit_10, SMA_5]
35	[SMA_3, SMA_4, SMA_2, gem_intensiteit_1, hour, SMA_5]
36	[SMA_3, SMA_2, SMA_5, hour, SMA_6, SMA_4]
37	[SMA_2]
38	[SMA_3]
39	[SMA_4]
40	[SMA_3]
41	[SMA_4]
42	[SMA_4]
43	[SMA_4]
44	[SMA_4]
45	[SMA_4]
46	[SMA_6]

Table B.5: Features used by the XGBoost models for the hour-ahead forecast for freight traffic.

sensor	Model features
1	[hour, SMA_12]
2	[hour, SMA_12]
3	[gem_intensiteit, hour]
4	[SMA_11]
5	[SMA_7, hour]
6	[SMA_5, hour]
7	[SMA_8]
8	[SMA_5, hour]
9	[SMA_4, hour]
10	[SMA_7, hour]

Continued on next page

Table B.5: Features used by the XGBoost models for the hour-ahead forecast for freight traffic.

sensor	Model features
11	[hour, SMA_2, SMA_3, gem_intensiteit, dayofweek, SMA_12]
12	[hour, SMA_3]
13	[SMA_2, gem_intensiteit, hour, SMA_3, SMA_12, gem_intensiteit_7]
14	[SMA_2, hour, SMA_3, SMA_5, dayofweek, SMA_12]
15	[hour, SMA_9]
16	[hour, SMA_12]
17	[SMA_7, hour]
18	[SMA_8, hour]
19	[SMA_5, hour]
20	[SMA_5, hour]
21	[hour, SMA_6]
22	[SMA_5, hour]
23	[hour, SMA_5]
24	[SMA_2, hour]
25	[SMA_2, hour]
26	[SMA_4, hour]
27	[SMA_3, hour]
28	[SMA_4, hour]
29	[SMA_3, hour]
30	[SMA_3, hour]
31	[hour, SMA_6]
32	[SMA_4, hour, SMA_5, SMA_11, SMA_9, SMA_10]
33	[hour, SMA_7]
34	[SMA_3]
35	[SMA_8]
36	[hour, SMA_5]
37	[SMA_4, hour, SMA_5, dayofweek, SMA_3, SMA_12]
38	[SMA_6, hour]
39	[SMA_6, hour]
40	[SMA_6, hour]
41	[SMA_5, hour]
42	[SMA_4, hour]
43	[SMA_5, hour]
44	[hour, SMA_6]
45	[hour, SMA_7]
46	[SMA_5, hour]

Table B.6: Features used by the XGBoost models for the hour-ahead forecast for passenger traffic.

sensor	Model features
1	[hour]
2	[hour, SMA_4]
3	[gem_intensiteit, hour, SMA_2, SMA_3, dayofweek, SMA_12]
4	[hour, SMA_2]
5	[hour, SMA_5, SMA_6, SMA_12, SMA_2, dayofweek]
6	[hour, SMA_3, SMA_2, SMA_4, SMA_5, dayofweek]
7	[SMA_3, hour]
8	[SMA_2, hour, SMA_3, SMA_4, dayofweek]
9	[hour, SMA_3]
10	[SMA_6, hour]
11	[SMA_2, hour, SMA_3, gem_intensiteit, dayofweek, SMA_8]
12	[hour, SMA_3]
13	[gem_intensiteit, hour, SMA_2, SMA_3, gem_intensiteit_2, dayofweek]
14	[SMA_2, hour, gem_intensiteit, SMA_3, gem_intensiteit_1, dayofweek]
15	[SMA_4, hour, SMA_12, SMA_10, SMA_2, SMA_3]
16	[SMA_9]
17	[hour, SMA_3]
18	[SMA_5, hour]
19	[SMA_2, hour, SMA_4, gem_intensiteit, SMA_5, dayofweek]
20	[SMA_8, hour, SMA_3, SMA_2, SMA_5, dayofweek]
21	[hour, SMA_3]
22	[hour, SMA_3]
23	[hour, SMA_3, SMA_4, SMA_2, gem_intensiteit, SMA_12]
24	[SMA_2, hour, SMA_3, gem_intensiteit, SMA_6, SMA_4]
25	[gem_intensiteit, hour, SMA_2, SMA_3, dayofweek, gem_intensiteit_2]
26	[gem_intensiteit, SMA_2, hour, SMA_3, dayofweek, gem_intensiteit_9]
27	[SMA_2, hour, SMA_3, gem_intensiteit, SMA_5, gem_intensiteit_12]
28	[SMA_3, hour, SMA_2, SMA_4, SMA_5, gem_intensiteit]
29	[SMA_3, hour, SMA_2, SMA_5, gem_intensiteit, SMA_10]
30	[SMA_2, SMA_3, hour, SMA_4, gem_intensiteit]
31	[hour, SMA_2, SMA_4, SMA_3, SMA_5, gem_intensiteit]
32	[gem_intensiteit, hour, SMA_2, dayofweek, SMA_3, SMA_8]
33	[SMA_2]
34	[SMA_2, hour, gem_intensiteit, dayofweek]
35	[hour, SMA_2, SMA_3, gem_intensiteit, SMA_12, SMA_4]
36	[SMA_2, hour, gem_intensiteit, SMA_3, gem_intensiteit_1, dayofweek]
37	[SMA_2, hour, gem_intensiteit, SMA_3, dayofweek]
38	[SMA_2, hour, SMA_4, gem_intensiteit, SMA_6]

Continued on next page

Table B.6: Features used by the XGBoost models for the hour-ahead forecast for passenger traffic.

sensor	Model features
39	[SMA_2, hour, SMA_4, SMA_3, gem_intensiteit, gem_intensiteit_12]
40	[SMA_2, hour, SMA_3, gem_intensiteit, SMA_12, SMA_5]
41	[SMA_2, hour]
42	[SMA_2, hour, SMA_3, gem_intensiteit, SMA_4, dayofweek]
43	[SMA_2, hour, SMA_5, SMA_3, gem_intensiteit_6, SMA_12]
44	[SMA_2, hour]
45	[SMA_2, hour, SMA_6, SMA_3, SMA_9, SMA_5]
46	[SMA_3, hour, SMA_2, SMA_4, gem_intensiteit, SMA_5]

B.3 GNN

B.4 Transformer

Table B.7: Parameters used for the 5-minute forecast by the transformer model for freight traffic.

sensor	lr	encoder sequence length
1	0.0001	6
2	0.0001	1
3	0.0001	6
4	0.0001	6
5	0.0001	3
6	0.0001	6
7	0.0010	6
8	0.0001	6
9	0.0001	6
10	0.0100	1
11	0.0100	6
12	0.0001	6
13	0.0010	3
14	0.0010	6
15	0.0001	6
16	0.0001	1
17	0.0001	1
18	0.0001	6

Continued on next page

Table B.7: Parameters used for the 5-minute forecast by the transformer model for freight traffic.

sensor	lr	encoder sequence length
19	0.0001	6
20	0.0100	6
21	0.0001	6
22	0.0001	6
23	0.0001	6
24	0.0001	6
25	0.0001	6
26	0.0001	6
27	0.0001	6
28	0.0001	6
29	0.0001	6
30	0.0001	6
31	0.0100	3
32	0.0001	3
33	0.0001	6
34	0.0010	6
35	0.0001	6
36	0.0001	3
37	0.0001	6
38	0.0001	6
39	0.0001	6
40	0.0001	6
41	0.0100	3
42	0.0001	6
43	0.0001	6
44	0.0100	1
45	0.0001	6
46	0.0001	6

Table B.8: Parameters used for the 5-minute forecast by the transformer model for passenger traffic.

sensor	lr	encoder sequence length
1	0.0001	6
2	0.0001	6
3	0.0100	6

Continued on next page

Table B.8: Parameters used for the 5-minute forecast by the transformer model for passenger traffic.

sensor	lr	encoder sequence length
4	0.0100	6
5	0.0100	3
6	0.0100	6
7	0.0001	6
8	0.0001	6
9	0.0100	6
10	0.0001	6
11	0.0100	3
12	0.0001	6
13	0.0100	6
14	0.0100	6
15	0.0001	6
16	0.0100	1
17	0.0001	6
18	0.0100	1
19	0.0001	12
20	0.0001	6
21	0.0001	12
22	0.0001	12
23	0.0100	12
24	0.0100	24
25	0.0100	24
26	0.0100	6
27	0.0100	6
28	0.0100	1
29	0.0100	1
30	0.0100	1
31	0.0001	6
32	0.0100	3
33	0.0100	1
34	0.0100	6
35	0.0100	24
36	0.0100	12
37	0.0001	6
38	0.0001	6
39	0.0001	1
40	0.0001	12
41	0.0001	12
42	0.0001	12

Continued on next page

Table B.8: Parameters used for the 5-minute forecast by the transformer model for passenger traffic.

sensor	lr	encoder sequence length
43	0.0001	12
44	0.0001	6
45	0.0001	12
46	0.0001	1

Table B.9: Parameters used for the hour-ahead forecast by the transformer model for freight traffic.

sensor	lr	encoder sequence length
1	0.0010	3
2	0.0001	24
3	0.0001	1
4	0.0010	3
5	0.0001	6
6	0.0001	6
7	0.0001	3
8	0.0010	12
9	0.0001	12
10	0.0100	1
11	0.0001	12
12	0.0001	24
13	0.0001	3
14	0.0010	24
15	0.0010	24
16	0.0100	1
17	0.0100	1
18	0.0001	12
19	0.0100	1
20	0.0010	1
21	0.0001	12
22	0.0010	24
23	0.0001	12
24	0.0001	6
25	0.0001	12
26	0.0001	3
27	0.0001	12

Continued on next page

Table B.9: Parameters used for the hour-ahead forecast by the transformer model for freight traffic.

sensor	lr	encoder sequence length
28	0.0001	12
29	0.0001	6
30	0.0001	12
31	0.0001	6
32	0.0001	6
33	0.0001	6
34	0.0100	12
35	0.0100	1
36	0.0100	24
37	0.0010	12
38	0.0001	6
39	0.0001	6
40	0.0001	3
41	0.0001	6
42	0.0100	1
43	0.0100	24
44	0.0100	12
45	0.0001	6
46	0.0001	6

Table B.10: Parameters used for the hour-ahead forecast by the transformer model for passenger traffic.

sensor	lr	encoder sequence length
1	0.0001	3
2	0.0001	12
3	0.0100	6
4	0.0001	3
5	0.0001	1
6	0.0001	3
7	0.0001	6
8	0.0001	3
9	0.0001	3
10	0.0001	6
11	0.0100	1
12	0.0001	288

Continued on next page

Table B.10: Parameters used for the hour-ahead forecast by the transformer model for passenger traffic.

sensor	lr	encoder sequence length
13	0.0100	24
14	0.0100	24
15	0.0001	6
16	0.0001	6
17	0.0001	6
18	0.0001	1
19	0.0001	6
20	0.0001	3
21	0.0001	3
22	0.0001	3
23	0.0100	1
24	0.0100	12
25	0.0100	24
26	0.0100	288
27	0.0100	12
28	0.0100	12
29	0.0100	24
30	0.0100	288
31	0.0001	1
32	0.0100	24
33	0.0001	12
34	0.0100	24
35	0.0100	24
36	0.0100	6
37	0.0100	24
38	0.0001	1
39	0.0001	1
40	0.0001	1
41	0.0001	3
42	0.0001	6
43	0.0001	1
44	0.0001	1
45	0.0001	1
46	0.0001	1

Appendix C

Results

Table C.1: RMSE per sensor for each model for 5-minute forecasts for freight traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
1	1.9107	1.6705	1.2923	1.3443	**1.2563**
2	1.4779	1.2419	1.0184	1.207	**0.996**
3	13.9972	17.9189	5.3058	**5.0729**	14.5434
4	2.1312	1.8268	1.575	**1.4551**	1.7374
5	4.3507	4.4637	**2.292**	2.4153	3.9085
6	2.8152	2.8225	**1.6764**	1.7244	2.3704
7	1.8253	1.7974	1.1626	1.3212	**1.0748**
8	4.1068	4.0057	2.128	2.1191	**1.9719**
9	2.5770	2.6462	**1.5512**	1.7175	2.2433
10	2.2311	2.0751	1.4678	1.4208	**1.2843**
11	18.1279	35.1563	8.6874	**7.6776**	26.3645
12	7.8423	15.2511	4.9897	**4.5944**	5.3653
13	30.9534	60.9886	11.3023	**10.0308**	46.6572
14	13.2040	16.0733	4.9166	**4.7696**	13.5588
15	1.9923	1.7982	**1.4144**	1.4528	1.4436
16	0.9393	0.7840	**0.6598**	1.1206	0.6679
17	2.5436	2.6435	1.5202	1.9117	**1.468**
18	2.2353	2.2711	**1.3951**	1.6149	1.8757
19	4.4487	4.9650	2.3311	2.4185	**2.2084**
20	2.0636	1.9044	1.2864	1.3954	**1.2235**
21	2.0601	1.9383	1.2707	1.4518	**1.1352**
22	2.6209	2.4981	1.5581	1.6235	**1.4457**
23	7.3741	8.5882	**3.6168**	3.8096	7.1222
24	6.0660	6.7180	**2.944**	3.0075	5.7804

Continued on next page

Table C.1: RMSE per sensor for each model for 5-minute forecasts for freight traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
25	6.0722	6.8206	**2.8597**	2.8823	5.9495
26	5.3261	5.7806	**2.5761**	2.7245	5.161
27	7.1045	8.1983	**3.2712**	3.2796	7.2868
28	6.9379	7.5179	**3.1528**	3.321	6.6446
29	7.4037	8.6631	**3.412**	3.6104	7.6089
30	6.1631	6.8542	**2.9856**	3.045	6.2329
31	1.4835	1.3918	1.0028	1.2331	**0.8861**
32	5.4696	5.7264	2.9939	**2.9725**	5.0848
33	1.4898	1.8287	**0.9926**	1.2767	1.1013
34	4.3005	5.1144	**2.7973**	2.89	3.9494
35	2.1014	2.0887	**1.4125**	1.4261	1.6832
36	7.4395	9.0729	3.6177	**3.4043**	8.1397
37	4.4395	5.4438	2.4038	**2.3012**	2.319
38	2.8975	3.0502	1.6766	1.7625	**1.5101**
39	3.2168	3.5252	**1.8029**	1.9191	1.9819
40	2.7237	2.8243	1.5775	1.656	**1.4432**
41	2.4494	2.3986	1.4419	1.459	**1.2776**
42	2.3032	2.1437	1.295	1.3567	**1.113**
43	2.2725	2.2385	1.3427	1.3901	**1.2019**
44	2.1832	2.0930	1.3421	1.3392	**1.1481**
45	3.0806	3.0565	1.7411	1.7437	**1.5595**
46	3.6946	4.0616	**2.011**	2.3785	2.1624

Table C.2: MAE per sensor for each model for 5-minute forecasts for freight traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
1	1.3422	1.2601	**0.9616**	0.9806	0.983
2	0.9991	0.9307	**0.7835**	0.8985	0.8081
3	8.9972	12.7927	3.5153	**3.4004**	12.8017
4	1.5262	1.3458	1.137	**1.0617**	1.3618
5	2.8803	3.1419	**1.6058**	1.6557	3.2919
6	1.8209	1.9675	**1.13**	1.1359	1.9297
7	1.1468	1.2738	**0.7928**	0.9186	0.838
8	2.6955	2.8107	1.4135	**1.3981**	1.4561
9	1.6836	1.8916	**1.0878**	1.1669	1.8747
10	1.3187	1.3976	**0.9257**	0.9563	0.9411
11	11.6165	27.7871	6.247	**5.4086**	23.5355

Continued on next page

Table C.2: MAE per sensor for each model for 5-minute forecasts for freight traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
12	5.5716	12.4262	3.6601	**3.3118**	3.9617
13	18.9051	47.9373	7.8495	**6.78**	41.8591
14	8.3357	11.3777	3.2925	**3.218**	11.8982
15	1.4058	1.3224	**1.0396**	1.0658	1.1005
16	0.4683	0.4334	**0.3773**	0.7979	0.4798
17	1.6434	1.9176	**1.0744**	1.3418	1.1355
18	1.3909	1.5901	**0.9279**	1.0051	1.5259
19	2.9109	3.4378	**1.5441**	1.6266	1.6024
20	1.2962	1.3381	0.883	0.9486	**0.851**
21	1.2610	1.3400	0.8338	0.9921	**0.8332**
22	1.7038	1.7693	1.0775	1.1068	**1.0323**
23	4.8200	6.2330	**2.3983**	2.5713	6.1735
24	4.1520	4.9309	**2.0976**	2.1129	4.9935
25	4.0889	5.0201	**2.0417**	2.0599	5.1823
26	3.6559	4.2679	**1.8436**	1.9469	4.4778
27	4.7770	5.7412	**2.2726**	2.293	6.5228
28	4.7055	5.2411	**2.1189**	2.2607	5.966
29	4.9797	6.1317	**2.3696**	2.4599	6.7507
30	4.2022	4.8759	**2.0634**	2.127	5.5455
31	0.8330	0.8937	0.6017	0.8273	**0.5965**
32	3.6200	4.1763	2.0381	**2.0291**	4.315
33	0.8527	1.2144	**0.6407**	0.8546	0.8749
34	2.9530	3.6730	1.9957	**1.9931**	3.3796
35	1.3064	1.4634	0.9289	**0.9264**	1.3029
36	4.7998	6.6866	2.4115	**2.3113**	7.0655
37	2.7701	3.7158	1.5652	**1.491**	1.7119
38	1.8357	2.1468	1.119	1.1594	**1.1019**
39	2.0582	2.4982	**1.2432**	1.2801	1.4846
40	1.7143	1.9826	**1.0703**	1.11	1.0741
41	1.5139	1.6163	0.9459	0.9676	**0.9231**
42	1.3508	1.4229	0.8111	0.9005	**0.7655**
43	1.3808	1.5156	0.8789	0.9206	**0.8622**
44	1.3062	1.4174	0.8526	0.887	**0.8315**
45	1.8977	2.0956	1.1409	**1.1282**	1.1647
46	2.2977	2.8222	**1.3342**	1.4958	1.5788

Table C.3: RMSE per sensor for each model for 5-minute forecasts for passenger traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
1	11.0442	24.4552	7.9398	**7.078**	9.1408
2	5.6642	10.5863	**3.9162**	4.4294	3.9197
3	59.0431	164.6680	25.3162	**21.0954**	120.0764
4	9.0765	18.4620	5.9819	**5.7453**	13.9778
5	10.9340	20.2694	6.4064	**6.2438**	14.8382
6	14.0140	29.4487	8.2078	**7.2687**	22.3746
7	6.1899	10.4566	3.8132	4.5776	**3.6703**
8	13.0234	23.3870	6.9342	**6.1797**	8.1125
9	7.6171	13.8282	4.6985	**4.6598**	9.7875
10	6.9871	10.0659	4.699	**4.0726**	4.1304
11	18.0962	40.1318	10.0416	**8.9537**	29.3644
12	6.8352	12.5582	4.8289	**4.5538**	5.0094
13	44.4368	127.4040	18.9797	**16.2896**	92.2563
14	59.9679	171.3426	24.5361	**21.7453**	124.4363
15	8.0501	16.7470	5.7243	**5.5184**	6.3327
16	3.1019	2.9896	2.4237	**2.2247**	2.6115
17	8.8263	17.9221	**5.863**	6.0067	7.1156
18	9.1920	15.4468	6.0158	**5.9721**	12.1212
19	12.2167	22.3857	6.7029	**5.9937**	7.43
20	7.2841	13.8739	4.779	**4.4177**	5.0976
21	5.2443	9.5664	**3.5179**	4.386	3.632
22	7.8746	16.6447	5.2425	5.1726	**5.14**
23	34.1078	81.4901	17.3726	**15.6591**	56.9224
24	26.3696	65.6758	13.1124	**11.443**	46.1288
25	28.1064	75.2202	12.4186	**11.393**	53.473
26	24.6657	63.5910	10.9592	**10.2872**	44.2965
27	32.9025	84.8826	15.3182	**14.667**	64.1597
28	35.8264	66.6888	16.2362	**14.643**	51.4973
29	32.0327	79.0474	16.6226	**15.2371**	60.3862
30	27.1488	64.4771	13.7988	**12.606**	48.8847
31	16.7140	32.6119	9.2283	**8.9036**	9.7434
32	42.7697	94.8555	18.5508	**16.7173**	68.4237
33	23.4574	49.7520	**8.4488**	9.5724	23.7275
34	38.3650	92.4856	**16.6644**	17.2625	56.0056
35	51.5072	109.3221	23.3126	**19.3507**	81.3718
36	33.5904	74.5241	16.1209	**13.7177**	54.772
37	22.5798	55.8717	12.1311	**10.0985**	21.7614
38	13.0119	31.2034	**7.4868**	7.6614	8.3101
39	13.9192	33.7765	7.9014	**7.6383**	10.0487

Continued on next page

Table C.3: RMSE per sensor for each model for 5-minute forecasts for passenger traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
40	11.7661	27.4683	**6.6685**	6.7004	7.0592
41	9.9686	21.4792	**5.8253**	5.8475	5.9297
42	14.5324	22.2506	5.9278	5.8864	**5.6644**
43	9.0826	19.4221	**5.5413**	5.6182	5.7351
44	9.0388	19.5073	**5.3515**	5.5506	5.7386
45	11.7415	26.9415	**6.8489**	7.1698	8.2468
46	15.3147	38.1624	9.6893	**9.4495**	11.9711

Table C.4: MAE per sensor for each model for 5-minute forecasts for passenger traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
1	7.7988	20.0938	5.7908	**5.1682**	6.7749
2	3.6993	7.9832	**2.718**	3.0732	2.7750
3	34.0074	133.8372	16.9942	**14.0305**	109.1041
4	6.5480	15.4193	4.3251	**4.102**	12.0257
5	6.6657	15.8403	4.4549	**4.403**	12.0780
6	8.6124	23.7716	5.5591	**5.1234**	19.3270
7	4.1537	8.6803	**2.7083**	3.1193	2.7314
8	8.0154	18.2801	4.67	**4.2178**	5.5103
9	4.6924	10.7235	**3.1746**	3.2869	8.2898
10	4.1479	8.0077	2.8791	**2.8126**	2.8800
11	11.7022	32.7849	7.3725	**6.5019**	25.1000
12	4.5078	10.0722	3.4685	**3.3043**	3.6879
13	26.2464	106.0320	13.0876	**11.171**	79.8679
14	34.7753	141.8373	16.2945	**14.3026**	108.5122
15	5.8675	13.4598	4.2059	**4.0532**	4.7789
16	1.8573	2.0444	**1.4667**	1.5646	2.0997
17	5.9760	14.9934	**4.1637**	4.3118	5.1293
18	5.9159	12.1563	**4.018**	4.0286	10.2995
19	7.2737	17.5881	4.4359	**4.0932**	5.2690
20	4.4183	10.7241	3.0161	**2.9986**	3.7065
21	3.5557	7.8347	**2.5487**	3.0002	2.7809
22	5.1510	13.6707	**3.5075**	3.563	3.7889
23	19.0550	66.6392	10.9223	**9.5053**	51.1067
24	15.1693	53.9843	8.8303	**7.7638**	41.4096
25	16.1466	62.1337	8.764	**8.0902**	48.1971
26	14.1225	52.0775	7.8127	**7.2943**	39.3000

Continued on next page

Table C.4: MAE per sensor for each model for 5-minute forecasts for passenger traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
27	19.2355	69.6475	10.2544	**9.5007**	56.5620
28	20.2853	54.5417	9.9549	**8.8669**	45.5989
29	18.6939	65.1191	10.7027	**9.5993**	53.6770
30	15.2251	53.0687	8.8904	**8.2496**	43.6677
31	9.3472	24.7899	5.8041	**5.5757**	6.7574
32	23.5172	77.9571	12.7413	**11.3068**	61.6508
33	13.3256	36.8701	**5.5963**	5.9954	20.6060
34	23.9552	76.1985	11.7609	**11.7461**	49.7125
35	27.7045	87.9573	14.4804	**12.0987**	73.2669
36	18.5890	60.1718	10.9155	**9.4801**	47.8725
37	13.1733	44.6238	8.0037	**6.9837**	15.2112
38	8.0872	25.6015	**5.0726**	5.2089	5.9068
39	8.5783	27.7894	5.404	**5.3468**	7.3344
40	7.2292	22.4585	**4.5867**	4.6646	5.0981
41	6.1988	17.3466	**4.0051**	4.0775	4.3736
42	8.9003	17.5579	**3.888**	3.9877	4.1114
43	5.6400	15.5523	**3.7857**	3.853	4.2125
44	5.6343	15.8963	**3.7435**	3.855	4.0749
45	7.4247	22.2921	**4.8007**	4.9322	6.2049
46	9.7940	31.4005	6.4735	**6.2099**	8.7054

Table C.5: RMSE per sensor for each model for hour-ahead forecasts for freight traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
1	1.9488	1.6616	**1.3279**	1.8497	1.5054
2	1.5209	1.2432	1.0327	1.8401	**0.9634**
3	15.1570	17.7744	6.8217	7.8141	**5.0854**
4	2.2173	1.8163	**1.5419**	2.1186	1.7486
5	4.5654	4.4112	2.3546	3.0335	**2.0809**
6	2.9397	2.8082	1.7238	2.2683	**1.5302**
7	1.8927	1.7908	1.2421	1.8937	**1.0713**
8	4.2687	3.9721	**2.4206**	2.7356	3.5598
9	2.6734	2.6235	1.6113	2.3198	**1.4234**
10	2.4170	2.0572	**1.5081**	1.9614	1.8662
11	20.8401	34.8489	10.5835	**9.6049**	10.3876
12	9.2496	15.1571	6.6947	6.7431	**5.8445**
13	35.6100	60.4034	**13.179**	14.0132	23.2854

Continued on next page

Table C.5: RMSE per sensor for each model for hour-ahead forecasts for freight traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
14	14.1411	15.9123	**5.4554**	6.2119	13.6091
15	2.0800	1.7814	**1.5004**	2.0304	1.7163
16	0.9621	0.7811	0.7184	1.7745	**0.6672**
17	2.6279	2.6237	**1.5462**	2.2205	2.2089
18	2.3197	2.2529	1.4529	2.0455	**1.2626**
19	4.6908	4.9269	**2.4814**	2.883	4.1893
20	2.1097	1.8857	**1.3136**	1.9301	1.6496
21	2.1573	1.9246	1.3017	1.9168	**1.1414**
22	2.6791	2.4780	**1.626**	2.2492	2.1469
23	7.6953	8.5011	3.8535	4.1282	**3.1857**
24	6.2995	6.6419	3.0959	3.4727	**2.7343**
25	6.3776	6.7448	2.992	3.3225	**2.6513**
26	5.5755	5.7160	2.7303	3.2028	**2.539**
27	7.5847	8.1320	3.4726	3.9022	**2.9675**
28	7.2898	7.4417	3.4203	3.988	**2.864**
29	7.8381	8.5806	3.6404	4.1832	**3.1049**
30	6.4786	6.7850	3.3886	3.8539	**2.743**
31	1.5773	1.3945	1.0449	1.8445	**0.8875**
32	5.7652	5.6706	3.1434	3.5364	**2.5791**
33	1.5056	1.8258	1.1278	1.8609	**0.8702**
34	4.5036	5.0632	**3.098**	3.4338	3.9707
35	2.1266	2.0809	**1.4934**	1.9348	1.6914
36	8.1462	8.9861	**4.0957**	4.2437	8.1042
37	4.8910	5.4109	**2.8655**	2.8814	4.4971
38	3.0738	3.0232	1.7499	2.2717	**1.5467**
39	3.4254	3.4971	1.8645	2.4307	**1.6854**
40	2.8558	2.8041	1.5864	2.139	**1.4468**
41	2.5541	2.3814	1.4718	1.9792	**1.27**
42	2.3632	2.1243	**1.3609**	1.9503	1.8512
43	2.3457	2.2247	**1.4088**	1.9631	1.9238
44	2.2449	2.0751	**1.3261**	1.9177	1.8315
45	3.1989	3.0301	1.8047	2.2862	**1.5393**
46	3.8971	4.0235	2.044	2.5341	**1.8109**

Table C.6: MAE per sensor for each model for hour-ahead forecasts for freight traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
1	1.3770	1.2550	**0.9813**	1.3406	1.214
2	1.0381	0.9341	0.7838	1.3175	**0.7785**
3	9.9945	12.6459	4.3812	4.8396	**3.7233**
4	1.6060	1.3317	**1.1293**	1.5199	1.363
5	3.0701	3.0950	1.6193	2.1065	**1.5825**
6	1.9147	1.9537	1.1528	1.5395	**1.1419**
7	1.1786	1.2651	0.8662	1.3147	**0.8126**
8	2.8432	2.7870	**1.5538**	1.8122	2.9056
9	1.7810	1.8750	1.1213	1.6004	**1.0533**
10	1.4459	1.3843	**0.9384**	1.3383	1.4477
11	14.1058	27.4288	7.4599	**6.7318**	7.4995
12	6.7072	12.3216	4.7348	4.8087	**4.3152**
13	23.7186	47.2011	**9.1339**	9.5097	17.3337
14	9.2871	11.2089	**3.6325**	4.0091	11.9999
15	1.4897	1.3087	**1.0835**	1.4606	1.3086
16	0.4923	0.4327	**0.4032**	1.2189	0.4895
17	1.7389	1.8936	**1.0663**	1.531	1.7986
18	1.4592	1.5687	0.9503	1.3935	**0.9323**
19	3.1016	3.4049	**1.6362**	1.973	3.4771
20	1.3367	1.3204	**0.8996**	1.3244	1.2869
21	1.3358	1.3286	0.8575	1.3106	**0.8343**
22	1.7451	1.7530	**1.1008**	1.5261	1.7287
23	5.1841	6.1644	2.5659	2.8145	**2.3721**
24	4.3909	4.8694	2.2079	2.4806	**2.0989**
25	4.4129	4.9588	2.1403	2.3956	**2.0387**
26	3.9007	4.2119	**1.9458**	2.3179	1.9565
27	5.1594	5.6855	2.3956	2.7157	**2.2387**
28	4.9925	5.1842	2.285	2.725	**2.0737**
29	5.3287	6.0536	2.5138	2.8686	**2.3283**
30	4.4608	4.8098	2.2628	2.6965	**2.0659**
31	0.8891	0.8939	0.637	1.2311	**0.6043**
32	3.9195	4.1319	2.1473	2.4305	**1.9686**
33	0.8765	1.2103	0.7093	1.2616	**0.6482**
34	3.1182	3.6298	**2.2226**	2.4061	3.3896
35	1.3445	1.4526	**1.0095**	1.3156	1.276
36	5.3542	6.5959	**2.6749**	2.8482	7.063
37	3.0836	3.6829	**1.7957**	1.9162	3.7417
38	1.9666	2.1215	1.1534	1.5295	**1.1159**
39	2.2300	2.4742	1.2582	1.6507	**1.2496**

Continued on next page

Table C.6: MAE per sensor for each model for hour-ahead forecasts for freight traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
40	1.8263	1.9686	**1.0647**	1.461	1.0758
41	1.5877	1.6025	0.9645	1.34	**0.9198**
42	1.4064	1.4063	**0.8575**	1.3072	1.4957
43	1.4522	1.5050	**0.9119**	1.3275	1.516
44	1.3746	1.4021	**0.8487**	1.2927	1.4434
45	2.0133	2.0774	1.146	1.5386	**1.1288**
46	2.4844	2.7881	1.3451	1.6927	**1.3274**

Table C.7: RMSE per sensor for each model for hour-ahead forecasts for passenger traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
1	15.3329	24.1964	11.4396	9.7982	**8.5229**
2	7.8425	10.5462	4.5909	7.2333	**3.8845**
3	76.1459	163.2609	41.1001	**35.1775**	119.8272
4	11.1812	18.3133	7.8055	11.3859	**6.3414**
5	13.7297	20.1545	6.5312	9.1783	**6.4203**
6	16.7635	29.1989	**8.6702**	9.301	10.7435
7	7.0356	10.3574	5.8987	8.4475	**3.6891**
8	16.0049	23.2095	10.0688	9.3226	**7.8548**
9	8.8954	13.7538	5.1085	7.2902	**4.6678**
10	8.1112	9.9655	5.2213	6.7405	**4.0924**
11	22.3102	39.7757	**13.2329**	16.3978	29.4505
12	8.8444	12.5132	6.299	11.8849	**5.2598**
13	58.5615	126.1086	**25.6116**	28.2865	92.8055
14	79.0482	169.6340	**34.0845**	35.9921	125.1627
15	9.8699	16.5949	7.195	8.6994	**5.8739**
16	2.8550	2.9754	2.2859	4.3663	**1.9334**
17	10.5620	17.7492	6.8422	8.3274	**5.8849**
18	10.6199	15.3223	6.8606	7.7132	**6.0095**
19	15.9899	22.2159	9.294	8.7484	**7.1775**
20	8.6231	13.7530	5.5786	7.2739	**4.4908**
21	6.0952	9.4845	3.8073	6.6287	**3.6772**
22	9.3383	16.4720	6.1027	7.2917	**5.0268**
23	39.6765	80.7429	**21.4325**	22.0291	56.6955
24	31.1774	65.0800	**16.2824**	16.8818	45.8621
25	34.1409	74.5161	**16.4238**	16.9014	53.4882
26	30.9426	62.9888	**13.5663**	15.3077	44.4207

Continued on next page

Table C.7: RMSE per sensor for each model for hour-ahead forecasts for passenger traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
27	41.4196	84.2454	**18.8866**	23.1376	63.8996
28	43.6828	66.0628	**20.207**	21.6912	51.2665
29	40.3233	78.2863	**19.812**	22.7418	60.4375
30	35.5048	63.8352	**19.0323**	19.399	48.9246
31	20.8525	32.3938	10.9463	12.1882	**9.2476**
32	51.3978	94.1291	**22.592**	23.9096	68.7791
33	22.5204	49.5574	16.0364	16.9062	**13.2104**
34	42.1245	91.7987	**24.5235**	27.263	56.4713
35	62.7322	108.4368	29.1804	**28.7773**	81.3517
36	41.4778	73.9862	**18.5883**	20.1659	54.6054
37	31.4542	55.4255	**14.8856**	15.6147	42.6068
38	15.6886	30.8983	9.3262	10.2938	**7.7588**
39	16.9427	33.4109	9.4089	10.4645	**8.5752**
40	14.7183	27.1820	7.7494	8.6839	**6.7888**
41	12.1295	21.2505	6.5867	7.5795	**5.7996**
42	15.2701	22.0339	7.8387	8.3378	**7.8329**
43	11.0745	19.2246	6.4953	7.4528	**5.3448**
44	11.1397	19.2918	6.3789	7.316	**5.3487**
45	14.1275	26.5988	8.6326	9.6903	**6.8612**
46	18.3438	37.7136	**11.5878**	12.1672	13.3756

Table C.8: MAE per sensor for each model for hour-ahead forecasts for passenger traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
1	11.4248	19.7903	8.3556	7.0021	**6.2802**
2	5.2800	7.9355	3.0013	4.894	**2.8017**
3	51.4023	132.1711	26.1821	**22.9198**	109.032
4	8.1613	15.2486	5.6068	7.482	**4.7222**
5	9.1136	15.6791	**4.5584**	6.5477	4.6451
6	11.3788	23.4958	**6.032**	6.5593	8.0656
7	4.8577	8.5797	3.7424	5.3253	**2.7612**
8	10.5757	18.1051	6.3938	6.1698	**5.3596**
9	5.7561	10.6278	3.3771	4.9636	**3.264**
10	4.9897	7.8992	3.1641	4.5752	**2.8311**
11	15.8791	32.4039	**9.4151**	11.4029	25.1804
12	6.1958	10.0166	4.1932	7.1864	**4.0629**
13	39.9200	104.6565	**17.5121**	19.2837	81.0443

Continued on next page

Table C.8: MAE per sensor for each model for hour-ahead forecasts for passenger traffic.

sensor	Baseline	Holt-Winters	XGBoost	GNN	Transformer
14	53.5864	139.9524	**23.2283**	24.5336	110.0472
15	7.3995	13.2815	5.3015	6.3415	**4.4279**
16	1.8946	2.0211	1.4892	2.9171	**1.4435**
17	7.5591	14.7930	4.7974	6.0224	**4.3384**
18	7.0560	12.0201	4.5397	5.2963	**4.1753**
19	10.0121	17.4022	5.5984	5.8386	**4.9942**
20	5.3130	10.5865	3.5394	4.9205	**3.2239**
21	4.2556	7.7354	**2.7361**	4.4775	2.7877
22	6.4273	13.4418	4.017	5.0115	**3.6699**
23	25.8824	65.8669	**14.114**	14.311	50.8678
24	20.6108	53.4510	**11.2208**	11.6842	41.1602
25	22.6576	61.4103	**11.5449**	11.9372	48.196
26	20.5590	51.4802	**9.3762**	10.8279	39.3657
27	27.5645	69.0129	**12.7151**	15.1458	56.4626
28	27.5835	53.8345	**12.9864**	13.7758	45.4849
29	26.6114	64.2396	**13.1773**	14.795	53.9381
30	22.8362	52.3942	**12.1373**	12.695	43.6814
31	13.0167	24.5275	7.0202	7.7173	**6.1991**
32	33.0327	77.0920	**15.6111**	16.8512	61.994
33	13.5520	36.5428	10.723	**9.7788**	9.9579
34	27.8523	75.3923	**17.4701**	18.732	50.2431
35	40.2539	86.8209	**18.9478**	19.0686	73.2901
36	26.0979	59.4526	**12.8178**	13.8748	47.701
37	20.5805	44.0882	**9.8397**	10.4193	36.2593
38	10.4190	25.2645	6.4371	6.9006	**5.5525**
39	11.3618	27.3762	6.503	7.0601	**6.2902**
40	9.9256	22.1187	5.3934	5.9541	**4.9272**
41	8.2489	17.1188	4.5136	5.2357	**4.244**
42	10.1062	17.3150	**5.1201**	5.4991	5.8087
43	7.4495	15.3313	4.4092	5.1579	**3.8352**
44	7.5203	15.6593	4.3885	5.0461	**3.9443**
45	9.7017	21.9541	5.9084	6.5663	**5.0661**
46	12.6314	30.8872	**7.851**	8.1549	10.1269

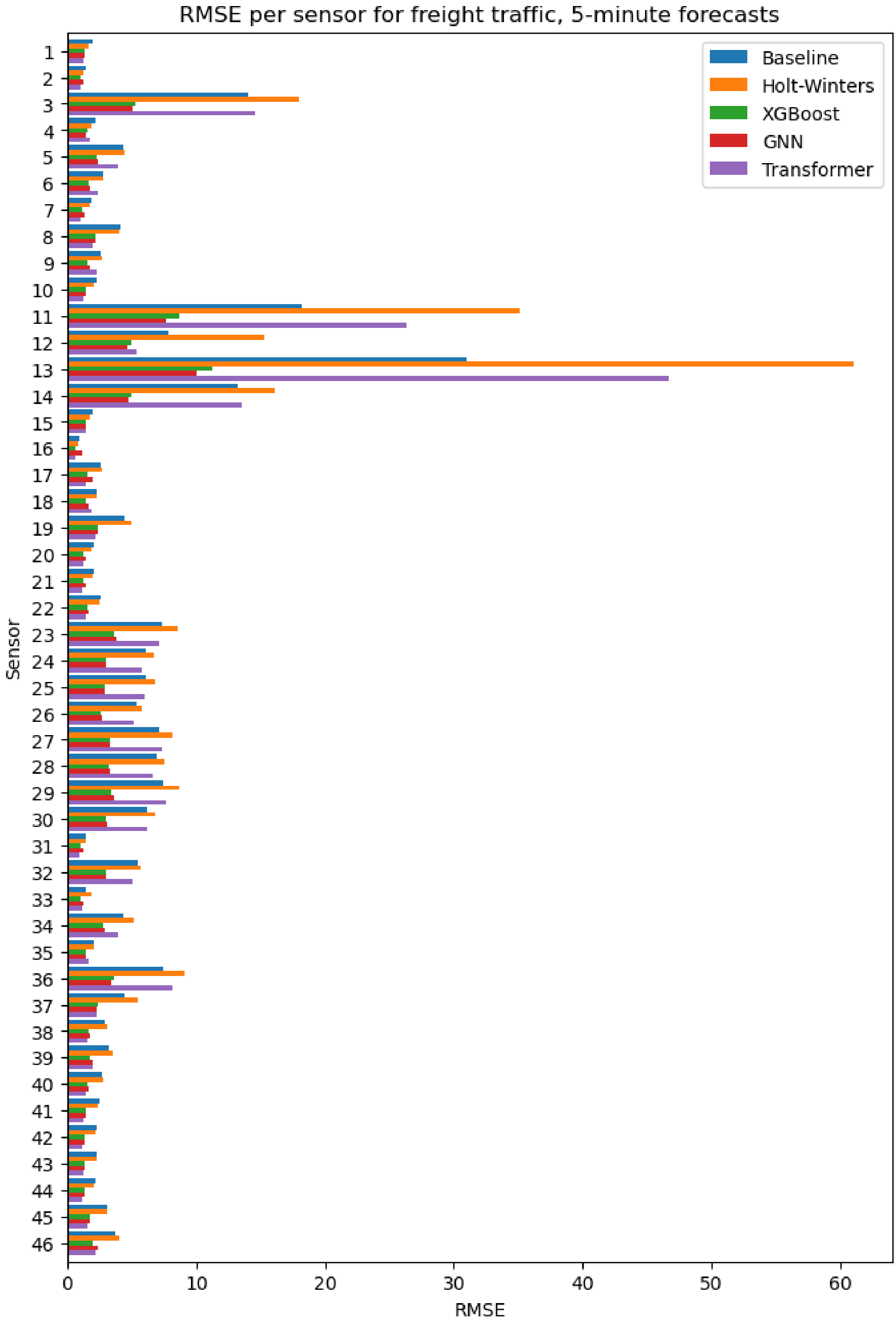


Figure C.1: RMSE of the 5-minute forecasts for freight traffic per sensor for each of the methods.

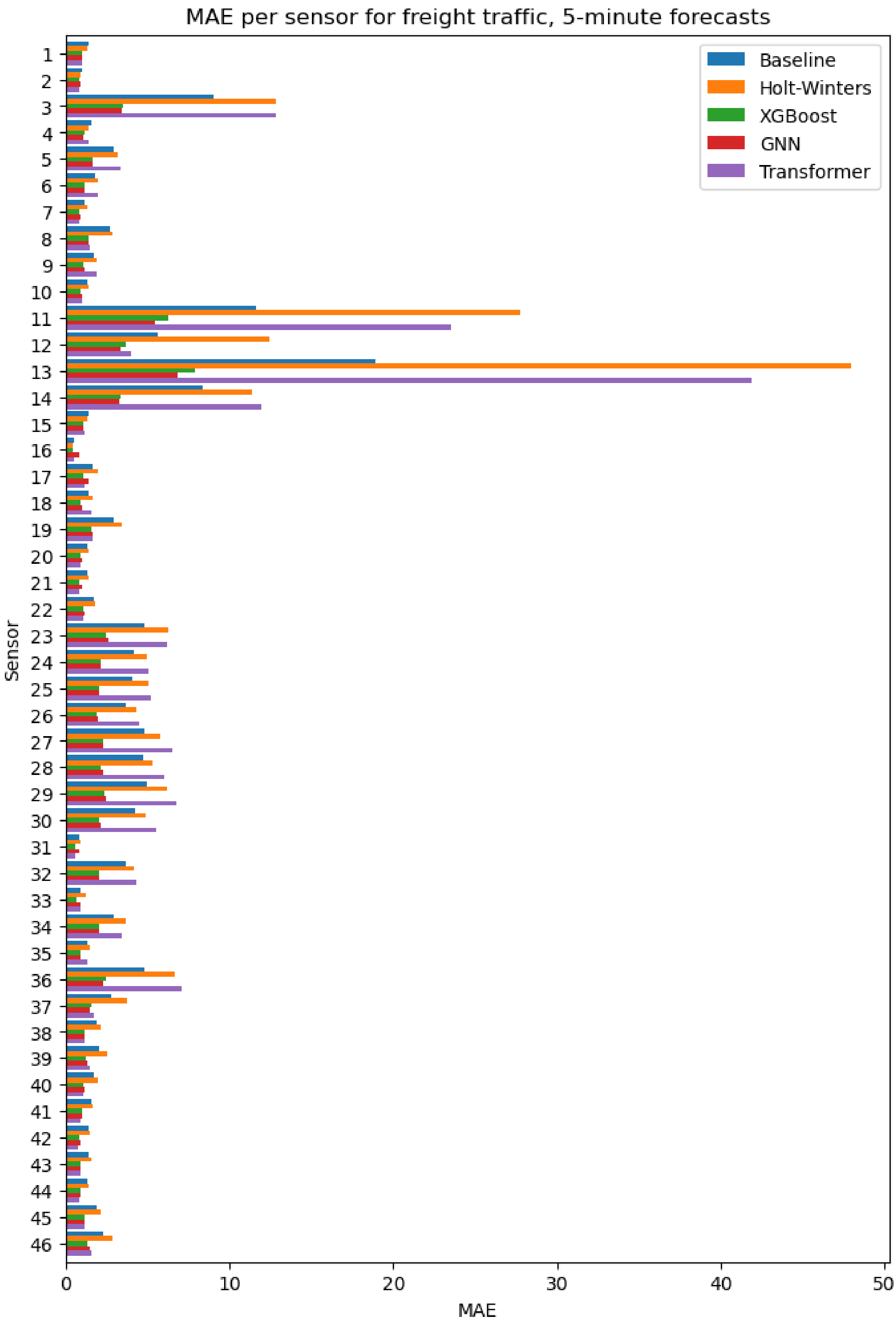


Figure C.2: MAE of the 5-minute forecasts for freight traffic per sensor for each of the methods.

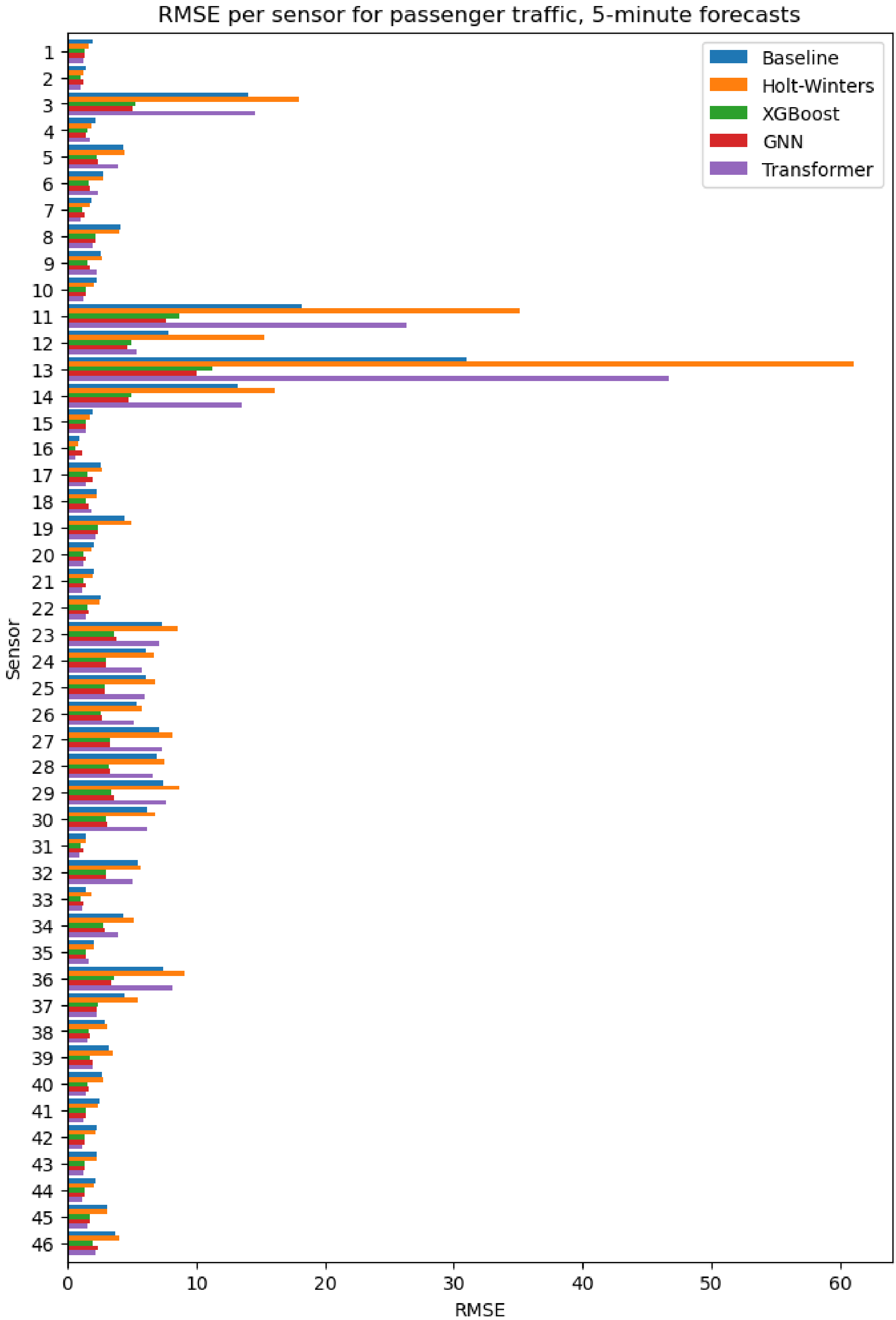


Figure C.3: RMSE of the 5-minute forecasts for passenger traffic per sensor for each of the methods.

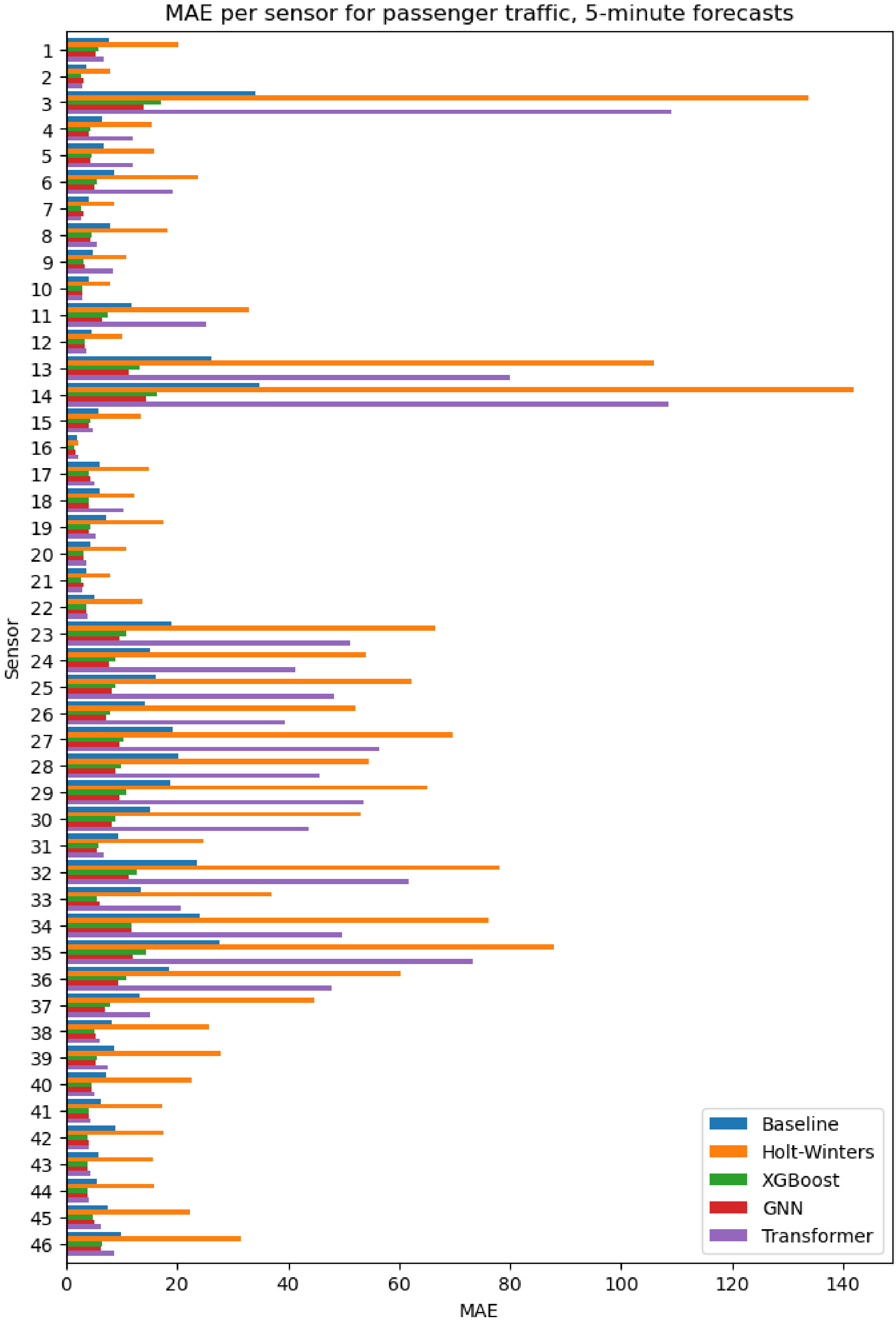


Figure C.4: MAE of the 5-minute forecasts for passenger traffic per sensor for each of the methods.

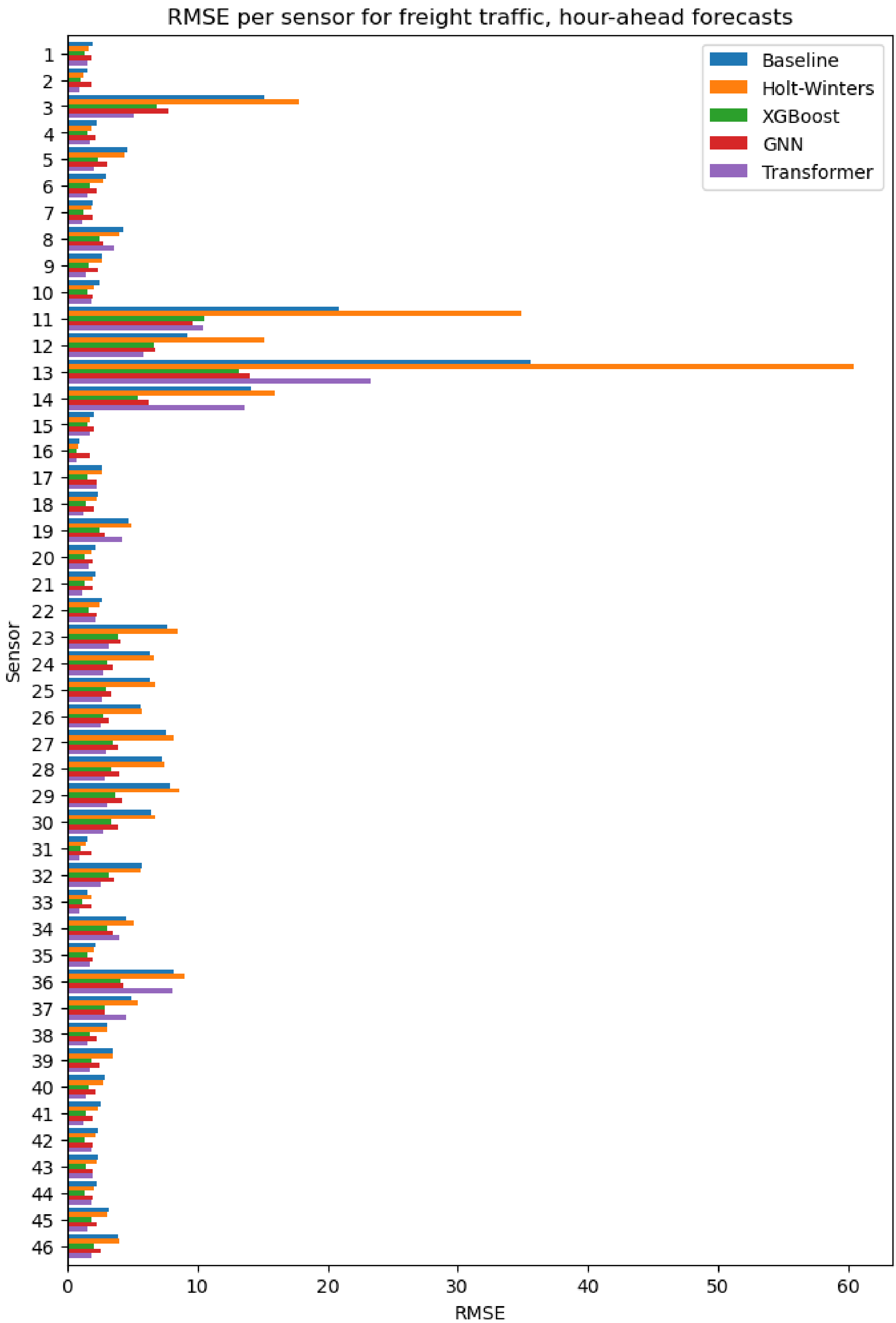


Figure C.5: RMSE of the hour-ahead forecasts for freight traffic per sensor for each of the methods.

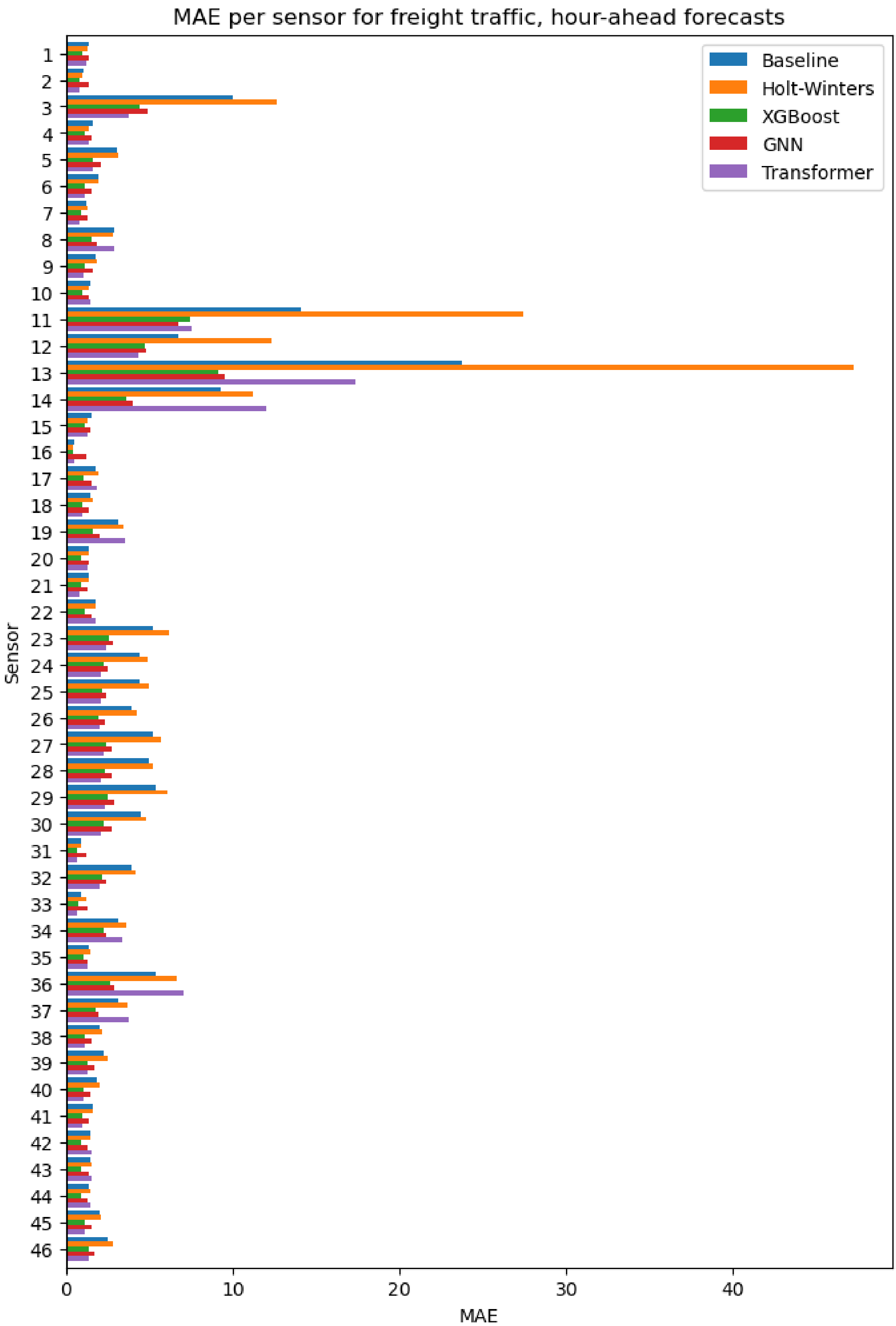


Figure C.6: MAE of the hour-ahead forecasts for freight traffic per sensor for each of the methods.

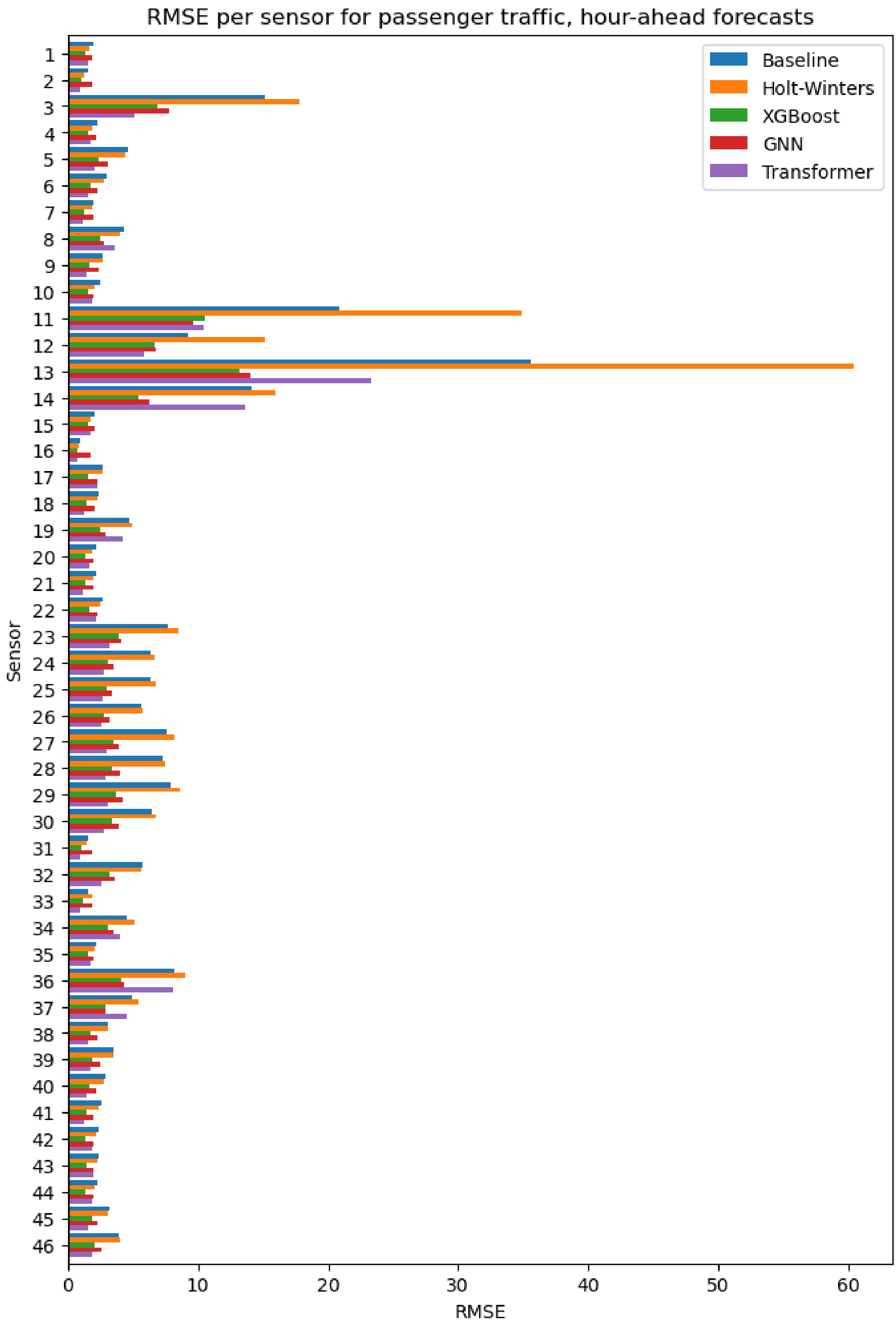


Figure C.7: RMSE of the hour-ahead forecasts for passenger traffic per sensor for each of the methods.

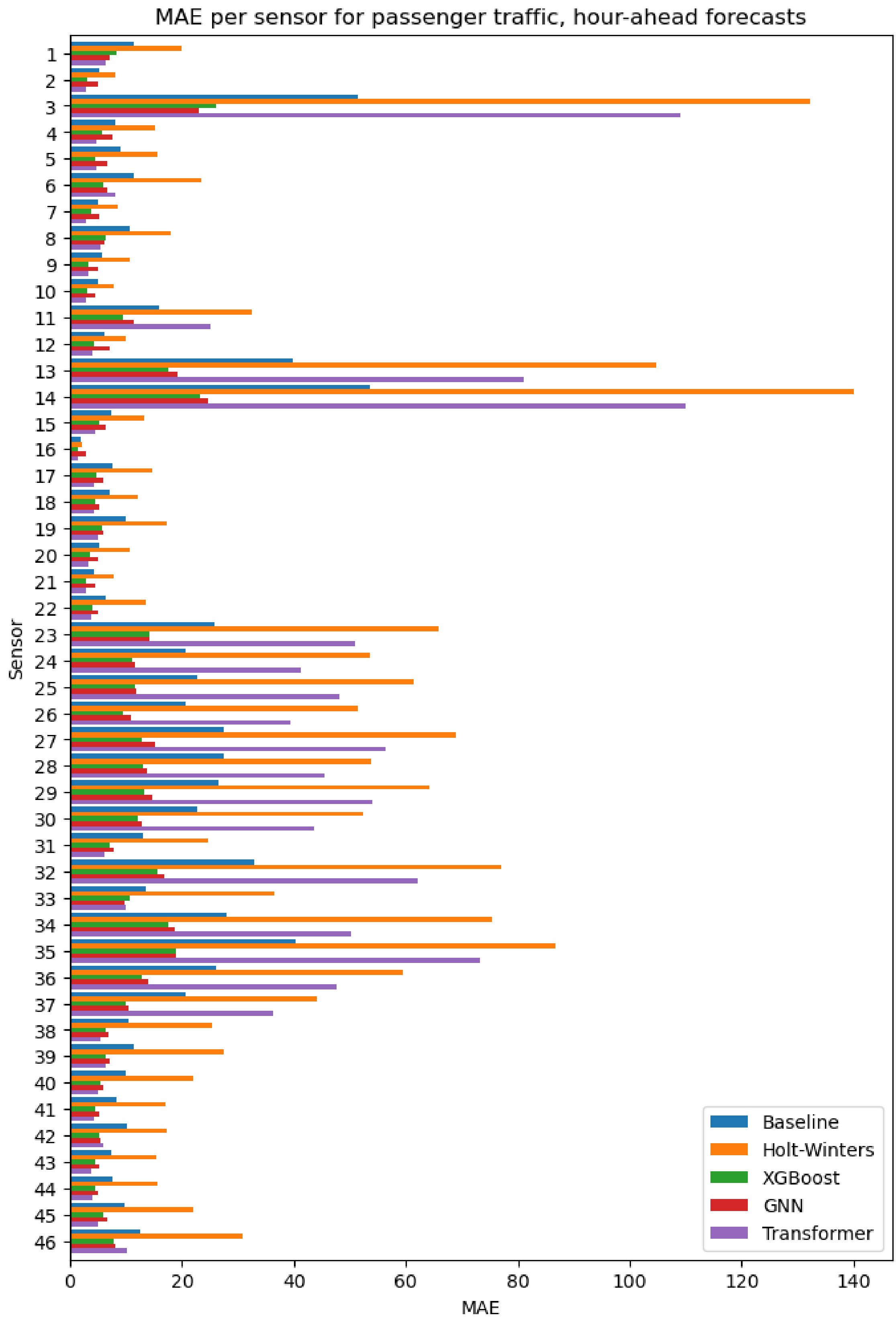


Figure C.8: MAE of the hour-ahead forecasts for passenger traffic per sensor for each of the methods.