Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 201 (2022) 213–222

Procedia Computer Science

www.elsevier.com/locate/procedia

The 13th International Conference on Ambient Systems, Networks and Technologies (ANT)
March 22 - 25, 2022, Porto, Portugal

A Comparison of Approaches for the Time Series Forecasting of Motorway Traffic Flow Rate at Hourly and Daily Aggregation Levels

Bas van der Bijl[a], Bart Gijsbertsen[a], Stan van Loon[a], Yorran Reurich[a], Tom de Valk[a], Thomas Koch[a,b,*], Elenna Dugundji[b,c]

[a] *Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands*
[b] *CWI National Research Institute for Mathematics & Computer Science, 1098 XG Amsterdam, The Netherlands*
[c] *MIT Center for Transportation & Logistics, 1 Amherst St, Cambridge, MA 02142, United States of America*

Abstract

Congestion forms a large problem in many major metropolitan regions around the world, leading to delays and societal costs. Congestion is generally associated with reduced average speed at a high traffic flow rate. This traffic flow rate is defined as the number of vehicles that pass a certain location at a given time. The modelling and prediction of this traffic flow rate may lead to valuable insights that may be used to reduce congestion and societal costs. This study aims to predict the traffic flow rate for 41 different locations in and around Amsterdam, The Netherlands. Using TBATS, SARIMAX and LSTM models, among others, the traffic flow rate of these locations has successfully been modelled. These models may provide accurate predictions for the future flow rate, which may be useful for the identification of infrastructure bottlenecks and the scheduling of maintenance. Considering the dramatic effects of the COVID-19 pandemic on the traffic flow rate, the inclusion of 2020 data with a number of external factors, could lead to an improvement of the results and the ability to model the future effects of the pandemic.

Keywords: time series; prediction; traffic flow; machine learning; traffic management

Congestion is known to have increased substantially in major cities over the course of the past few decades. Therefore, it is essential that governments stay aware of any trends regarding the traffic flow rate in the major cities. This awareness enables governments to foresee any bottlenecks in the infrastructure in the long term. These bottlenecks lead to a lot of congestion, which results in immense costs. The Dutch institute for mobility policy (KiM) estimated the societal cost of congestion on the primary highway network at 3.3 to 4.3 billion euros in 2018 [7]. Accurately predicting and preventing the peaks in the traffic flow rate could lead to a huge cut in these costs.

Based on mobility data in 2015, the Dutch knowledge platform on traffic and mobility, CROW [15], forecasts that the increase in traffic flow rate will be especially problematic in the city of Amsterdam with its lack of space and large number of narrow streets. Therefore, this study is focused on the analysis of the current traffic flow rate in Amsterdam

* Corresponding author
E-mail address: koch@cwi.nl

10.1016/j.procs.2022.03.030

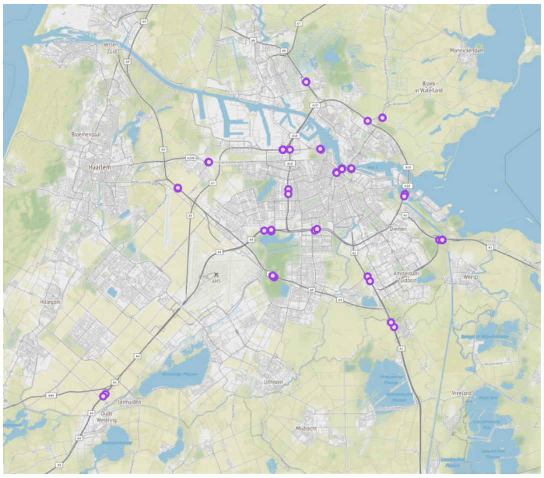

Fig. 1: Visualization of the sensor locations showing a wide variety of different sensor locations in and around Amsterdam.

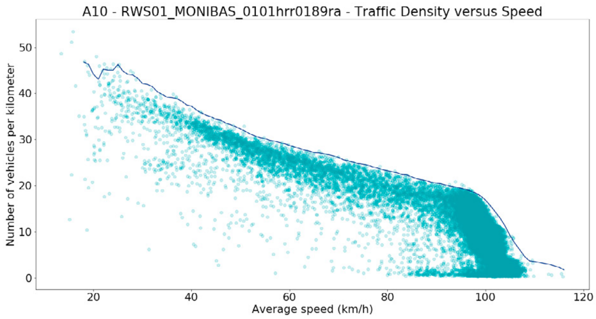

Fig. 2: Scatter plot of the average speed versus traffic density (number of vehicles per kilometer) on the A10. It can be seen that a low speed yields a higher density of vehicles, indicating that vehicles are closer together at lower speeds and suggests that there is congestion.

as well as the prediction of this traffic flow rate for the future. These predictions are done by identifying and using patterns in the historical data.

The prediction of the future traffic flow rate can be used for various purposes. The short-term predictions, for instance, can be used to find a route through Amsterdam that avoids congestion. As previously mentioned, the long-term predictions can be used by the government or the municipality of Amsterdam in the identification of bottlenecks to take appropriate actions. These long-term predictions can also be used to schedule maintenance without drastically affecting the traffic flow rate. The analyses and predictions can also be used to investigate the traffic flow rate with respect to the COVID-19 pandemic.

1. Case study

The data set that has been provided by the Dutch data portal for road traffic (NDW) [9] and consists of 9,619,639 observations with 22 columns. These observations describe the flow rate of traffic, in the number of vehicles that pass over a road sensor per hour, from 41 locations around Amsterdam. These locations are shown in Figure 1. It can be seen that the locations describe major motorways around Amsterdam as well as a few smaller roads in the city center. The data set provides hourly data for the 41 locations for five years from the 1st of January 2016 up until the 1st of January 2021. The data describes an hourly aggregation of the number of vehicles, the average speed and an optional categorization by the length of those vehicles.

To determine which models are applicable to the provided data, an extensive data analysis is carried out. The data analysis will also provide other valuable insights regarding the traffic flow rate in Amsterdam.

Since there is a large number of locations, it is not viable to show the analyses for all locations. Therefore, two locations have been chosen. The two locations that have been chosen are the A10 at `RWS01_MONIBAS_0101hrr0189ra` and the A4 at `RWS01_MONIBAS_0041hrr0197ra`. These locations show relatively large traffic intensities and a large variance in traffic speed. It is assumed that large motorways, such as these, have a larger effect on the traffic flow rate of Amsterdam than smaller roads do. These two locations show a clear relationship of the average speed and traffic flow rate. Most other locations show similar effects, which is why we expect that any conclusions will generalize well. The A10 and A4 locations that have been chosen for the data analysis, will also be used later in the project for the validation of the different models. This drastically simplifies the validation of models.

One metric that gives insight into the traffic flow rate for certain locations, is the density of vehicles. The density of vehicles is defined by the total number of vehicles, divided by the average speed and the total number of lanes, resulting in the number of vehicles per kilometer. Figure 2 shows the average vehicle density versus the average speed for the A10 location. It can be seen that a low speed yields a higher density of vehicles. This means that vehicles are closer together at lower speeds, suggesting that there is congestion. The A4 location showed similar patterns regarding the vehicle density.

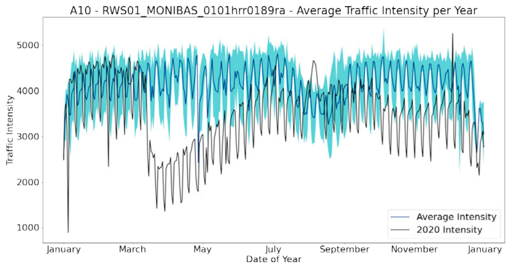

Fig. 3: Plot of the average flow rate per year of the A10. Multiple seasonal patterns can be observed and it can be seen that the traffic flow rate for 2020 deviates from the general patterns. This is a result of the measures to reduce mobility in order to stop the spread of COVID-19.

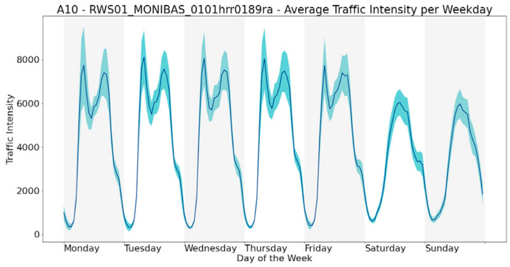

Fig. 4: Plot of the average traffic flow rate per day of the week of the A10. Clear daily and weekly patterns can be observed, with peaks during the morning and evening rush hours.

The average traffic flow rate can also be used to estimate which locations have the potential to be future bottlenecks for traffic in Amsterdam. These locations can be identified by the fraction of time in which they are at their maximum capacity. Unfortunately, the locations that are described by the data set have extremely diverse characteristics and are, therefore, remarkably difficult to compare. This is especially difficult when it comes to speed and flow rate. For that reason, only the roads and motorways with a maximum speed of 100 km/h, have been analysed regarding future bottlenecks. In Figure 2, it can be seen that the largest densities are reached for speeds of less than 30 km/h. With that, potential future bottlenecks can be identified by investigating which motorways spend the largest fraction of time at a speed of less than 30 km/h.

Figure 3 shows the average traffic flow rate per year for the A10, with the standard deviation plotted around the average. Since the different years approximately follow the same patterns, it can be concluded that there is a yearly seasonal component. It can also be seen there are no clear upwards or downwards trends that need to be considered. Furthermore, it can be seen that 2020 significantly deviates from the other years. This is a result of the COVID-19 pandemic. Especially the first national lockdown from March of 2020, shows a remarkable decrease in the traffic flow rate. This deviation will also have a notable impact on the modelling of the data and the forecasting of the future traffic flow rate. For that reason, it has been decided that model validation will be carried out on 2019 rather than 2020. These patterns were also visible for the A4 location.

The plot in Figure 4 shows the progression of the average traffic flow rate per hour of the week for the A10. Here, 2020 is disregarded as the COVID-19 pandemic affects the clarity of seasonal patterns. The plot shows a clear daily seasonal component. It can be seen that the traffic flow rate is highest in the morning and in the afternoon. This is shown by the two peaks in the graph of each day. Furthermore, it can be seen that the traffic flow rate is by far the lowest during the nights. The plot also clearly shows that the traffic flow rate is far lower on Saturdays and Sundays than it is the rest of the week. It can also be observed that the nights during the weekend are slightly less quiet than during the rest of the week. Another pattern that can be identified is that the morning rush hour contains a higher peak in traffic flow rate than the evening rush hour does. The traffic flow rate is slightly more spread out during the evening rush hour. These weekly and daily patterns were also visible for the A4.

It is expected that the different seasonal components that have been identified, may interfere with each other. This may result in difficulties when modelling the data.

2. Background

Considering that the traffic flow rate depends on time, it can be interpreted as a time series. There are multiple models that can be used for time series prediction. A number of these models are described in this section. After the elaboration of the different models, the evaluation techniques of time series models are discussed. These evaluation techniques are used to judge the quality of the fit of the model, as well as to compare the performance of the different models. Furthermore, the input data may be enriched with exogenous factors. This section also elaborates on which factors are suitable for this.

2.0.1. Seasonal Naive Model

Seasonal naive models are suitable for data with clear seasonal components [5]. For this model, the predicted value is equal to the last observed value from the same moment of the previous season. Considering that the input data show multiple seasonal components, this approach may be limited. Another limitation of this model is that potential trends are not used in the prediction. Disregarding the limitations, the simplicity of the model and the seasonal patterns of the input data, the seasonal naive model will be used as a baseline throughout this research.

2.0.2. *Autoregressive Integrated Moving Average Models*

Autoregressive integrated moving average, or ARIMA, models are a combination of autoregressive (AR) and moving average (MA) models. The order p in the AR model represents the number of historical values that are used in a linear combination to form the next value. The order q in MA models represents the number of historical values that are used in the moving average. The orders p and q can be determined using the partial autocorrelation function (PACF) and the autocorrelation function (ACF) plots, respectively. The next value X_t can be defined using the following equation:

$$X_t = \sum_{j=0}^{q} \beta_j Z_{t-j} + \sum_{i=1}^{p} \alpha_i X_{t-i} + Z_t \quad (1)$$

In this equation, $\{Z_t\}$ is a white noise time series. $\alpha_1, \alpha_2, ..., \alpha_p$, and $\beta_0, \beta_1, ..., \beta_q$ are the weights that describe the contribution of the previous p and q values, respectively.

ARIMA models are most suitable for the modelling of time series that contain trends. The modelling of time series with seasonal patterns can be overcome by differencing. This differencing is described by parameter d, that ARIMA adds to the AR and MA models, resulting in an order (p, d, q) [4].

ARIMA models require different modelling assumptions to be made. In the first place, the time series is required to be stationary. For stationary time series, $E(X_t)$ and $E(X_t X_{t+h})$ exist and do not depend on t. This means that the time series is consistent over time, indicating that the time series can not contain trends or seasonal patterns. An Augmented Dickey-Fuller test can be used to test whether a given time series is stationary. In second place, the residuals are required to be normally distributed. To determine whether the residuals are normally distributed, a comparison can be made between a kernel density estimation of the distribution of the residuals and the normal distribution. Further, a correlogram of the residuals can be consulted to determine whether there are still patterns present in the residuals. If this is not the case, it can be concluded that the ARIMA model is an appropriate fit.

2.0.3. *Seasonal Autoregressive Integrated Moving Average Models*

ARIMA models are used to make a non-stationary time series stationary by adjusting the trend. Seasonal autoregressive integrated moving average, or SARIMA, models are an extension of this, in which a seasonal component is also considered. This means that a SARIMA model is used to make a non-stationary time series stationary by removing the trend and seasonality [4]. This results in the addition of seasonal parameters P, D, Q and s, where s is the number of observations in a season. With that, a SARIMA model can be described by the order (p, d, q)x$(P, D, Q)[s]$. Similarly to the ARIMA model, the ACF and PACF plots can also be used to determine the values for P and Q.

A SARIMA model uses differencing at a lag equal to the number of seasons (parameter s) to remove additive seasonal effects [3]. The best way to find the most suitable values for the parameters, is by trying a large number of models with different parameter values [1]. These parameter values can be determined using a grid search.

Additionally, SARIMA can also be implemented with exogenous features, resulting in a SARIMAX model [8]. These exogenous features, as the name states, have a different origin than the original time series. This means that the features describe other characteristics of the observations than the time series does. The addition of these features may lead to a more precise forecast of the time series.

2.0.4. *(T)BATS Models*

(T)BATS is an acronym to describe the key features of the model: (**T**rigonometric seasonality), **B**ox-Cox transformations, **A**RIMA errors, **T**rends, **S**easonal components. The way in which these model components are used is fully autonomous [10].

TBATS models use exponential smoothing to model highly seasonal data. Whereas most popular models can incorporate at most one seasonal pattern, a TBATS model is known to incorporate multiple, complex seasonal patterns. TBATS models can incorporate seasonal patterns of which the frequency is not constant [11]. According to [12], each seasonality in a TBATS model is modelled by a trigonometric representation based on a Fourier series. A Fourier series is an expansion of a periodic function in terms of an infinite sum of sines and cosines. It is predominantly used to break up periodic functions to calculate them individually [16]. The Box-Cox transformations that are incorporated in the model, enable the model to handle non-linearity in the data. This stabilizes the variance in the predictions [2].

One major drawback of TBATS models is that forecasting may be slow, especially for long time series [5]. Another drawback of TBATS models is that other explanatory variables may not be added, potentially limiting the overall performance. The long-term predictions of TBATS models are thereby not always adequate. Furthermore, TBATS models are used under the assumption that the residuals are normally and independently distributed [2]. According to [5], the autonomous nature of the model components may thus lead to predictions that are not useful.

2.0.5. *Recurrent Neural Networks*

A recurrent neural network, or RNN, is a neural network that is specifically designed for the analysis of sequences. A specific subsection of RNN's that are proven to be useful in time series modelling are Long Short-Term Memory models, or LSTM models. These models have a similar structure to neural networks, where there are input and output layers, with at least one hidden layer to connect them both. The difference between neural networks and LSTM models, is the option to make connections between neighbouring nodes within the same (hidden) layer. The connections make it possible to attain and retain 'memory'.

When training LSTM models, parameter-tuning is crucial. This means that optimal values need to be found for the parameters of the LSTM. In the first place, the number of input nodes determines how many historical values are used. In second place, the number of output nodes determine the number of future values that are predicted. When more output nodes are present, the accuracy and reliability of the prediction will most likely decrease. The number of hidden layers determines the number of node layers between the input and output layers. Each hidden layer contains a given number of (hidden) nodes, each with a non-linear activation function. Increasing the number of layers and nodes leads to an increase in complexity. This means that a trade-off has to be made between the complexity and the training time. The training of a LSTM model is done for a number of epochs, which is equal to the number of times that the model will iterate through the entire data set. The most suitable values for these parameters can be found using a grid search, where the different models are evaluated using the evaluation metrics that are discussed later.

LSTM models generally have one output node, giving one predicted value. In time series prediction, however, it is essential to predict further into the future. For that reason, the rolling update method is used. This method is also known as walk-forward validation [13]. The essence of the method is to use observed data as well as predictions to predict future values. This prevents leakage of the validation data and it enables the LSTM model to predict farther into the future.

2.0.6. *Alternative Models*

Other models that have been considered for time series prediction are exponential smoothing, the Holt-Winters method, autoregressive moving average models and variations thereof, vector autoregressive moving average models and variations thereof, k-nearest neighbors, convolutional neural networks and graph neural networks. These models have not been discussed as they are too complex, lack flexibility or the ability to accommodate complex seasonal patterns.

2.1. *Evaluation of Models*

The evaluation of a model is crucial for the understanding of the performance of the model and the comparison of different models. There is a large array of metrics that can be used for this task. These metrics can be categorized as follows:

- Metrics to indicate the distance between the predicted and actual values
- Metrics that show the fraction of the variance in the data that is explained by model
- Metrics that assist in the trade-off between complexity and likelihood

Some of these metrics are more suitable than others for the evaluation of time series forecasting. In this case, the MAE, or mean absolute error, is the best metric to use for the comparison of different models [6]. The MAE is a measure of the distance between the predicted and actual values, calculated using the following equation: $\text{MAE} = \frac{\sum_{i=1}^{n} |\hat{y}_i - y_i|}{n}$ Moreover, the AIC is used to select the most appropriate orders for the ARIMA and SARIMA models. This measure penalizes the use of a large number of parameters and reduces the probability of overfitting.

2.2. *Extension of Data*

Next to the data that have been provided, there may be more characteristics of the different locations that also affect the traffic flow rate.

One of the external features that may affect traffic flow is the weather [14]. Weather characteristics such as precipitation, fog and wind affect the traffic flow negatively. For that reason, obtaining weather data for the different locations in Amsterdam, may prove to be extremely valuable.

Next to the weather, large events in Amsterdam may also contribute to the traffic flow rate. Therefore, information regarding events in venues such as the RAI Convention Center, the Johan Cruijf Arena or the Ziggo Dome, may provide additional value to the traffic flow rate data.

Considering that the COVID-19 pandemic has had a remarkable effect on the traffic flow rate for 2020, using COVID-19 statistics to enrich statistics may provide more precise predictions of the traffic flow rate. The following statistics, among others, may provide valuable information in the prediction of the traffic flow rate: the number of positive COVID-19 tests, the number of COVID-19 patients admitted to the hospital, the number of COVID-19 patients admitted to intensive care units, the number of deaths per day as a result of COVID-19 infections. Since the Dutch government bases its COVID-19 measures on these statistics, it is assumed that these statistics will affect the traffic flow rate.

3. Results

In this section, the results of the modelling stage are thoroughly discussed. As previously mentioned, the models were initially fitted to hourly data. This worked adequately for the simple models, but the computational limits formed a problem when training larger, more complex, models. Therefore, the complex models were only applied to the daily-aggregated data.

The seasonal naive method formed a baseline for the performance of the models. Therefore, the models that underperformed with respect to the seasonal naive method, have been disregarded.

3.1. Hourly Results

3.1.1. Seasonal Naive Model

The seasonal naive model achieves a MAE of 2999.63 for the A10 and 1898.77 for the A4. Since the seasonal naive model is used as a baseline model, these MAEs represent the baseline statistics for the validations for the traffic flow rate on 2019.

3.1.2. TBATS Model

The TBATS model achieves a MAE of 777.68 for the A10 and 578.10 for the A4. It is expected that the performance of TBATS models is better for short-term predictions, than for long-term predictions. To investigate this hypothesis, the MAE for the first six months of the validation was compared to the last six months of the validation. These MAEs were extremely similar, which is why the expectation was incorrect.

The TBATS models were trained with seasonalities of 24, 168 and 8766 hours. These seasonalities are equal to a day, 7 days and 365.25 days (a year with the consideration of leap years). During the training of the model, Box-Cox transformations and ARIMA errors were used. The data were also transformed using a log-transformation before being entered into the model. These measures ensure that no negative predictions can be made.

To determine whether the TBATS fit is adequate, the modelling assumptions need to be verified. In the case of TBATS models, the residuals are required to be normally distributed. According to kernel density estimation plots, the residuals are approximately normally distributed for the A10 and A4 locations. Therefore, it can be stated that the TBATS models fit the data adequately.

3.2. Daily-Aggregated Results

3.2.1. Seasonal Naive Model

The seasonal naive model achieves a MAE of 11110.53 for the A10 and 6880.92 for the A4 on daily-aggregated data. These values form the baseline for the daily-aggregated models.

3.2.2. (Seasonal) ARIMA Model

ARIMA models were not appropriate for the hourly traffic flow rate due to the interference of different seasonal patterns. Since the daily-aggregated data have less seasonalities, the ARIMA model may be suitable for the daily-aggregated data. In our data analysis was concluded that the traffic flow rate contains both a weekly pattern and a yearly pattern. Therefore, a difference of order 365 and a difference of order 7 were used on the log-transformed time series of the A10 location. These differences remove any seasonal patterns and result in a stationary time series. This stationarity is tested using the Augmented Dickey-Fuller test, resulting in a p-value of $1.21 \cdot 10^{-14}$. This means that with an α-value of 0.05, the null-hypothesis of the time series being stationary, is rejected. Therefore, ARIMA models may not be suitable for the modelling of the data.

Similarly to the A10, the differencing transformations were also carried out for the A4. For the resulting time series, the Augmented Dickey-Fuller test led to a p-value of $1.23 \cdot 10^{-17}$, indicating that the null-hypothesis is rejected once again. This implies that the data are not stationary. Just as for the A10 that was described previously, a standard ARIMA model may not suffice for the modelling of the traffic flow rate at the A4 location.

The PACF of the transformed time series of the A10 shows that there is a periodic deviation around lag seven. This indicates that there may still be some weekly seasonality in the transformed data. This seasonality explains why the transformed time series is not stationary. A seasonal ARIMA (SARIMA) model could potentially solve this issue. This leads to a SARIMA model of the order $(2, 0, 1)$x$(1, 0, 1)[7]$. The same conclusions hold for the A4, leading to a SARIMA model of the order $(1, 0, 1)$x$(1, 0, 1)[7]$.

The MAE for the SARIMA model for the A10 is equal to 7976.32. The MAE for the A4 is equal to 5118.73, which is a notable improvement with respect to the baseline model.

SARIMA models require certain modelling assumptions, with that, the residuals need to be normally distributed and uncorrelated. These assumptions are verified using a plot of the fitted values against the observed values, a kernel density plot of the residuals and a correlogram. These plots suggest that the residuals are indeed normally distributed and uncorrelated. This holds for the A10 and A4 locations. Therefore, it can be concluded that the SARIMA models fit the data adequately.

3.2.3. SARIMAX Model

Next to the SARIMA model that has been elaborated in the previous section, a SARIMAX model has been developed. This model is equal to the SARIMA in the way that it was developed, with one crucial difference: namely, the SARIMAX model incorporates weather data from Amsterdam. In our previous analysis it was shown that the weather seems to affect the traffic flow rate.

To determine the influence of the weather on the traffic flow rate, the correlation between the traffic flow rate and several weather characteristics has been investigated. This analysis showed little correlation between the traffic flow rate and the following daily weather characteristics: wind gusts, average temperature in degrees Celsius, number of sun hours, daily fraction of fog, daily fraction of rain, amount of snow and ice formation.

The SARIMAX model that has been fitted to the A10 has an order of $(2, 0, 1)$x$(2, 0, 1)[7]$. This model results in a MAE of 11110.43 for the 2019 validation of the model. For the A4 location, the SARIMAX has an order of $(1, 0, 0)$x$(1, 0, 1)[7]$. This model results in a MAE of 6880.82 for the 2019 validation of the model. Considering the MAE of the SARIMAX models, it can be stated that the SARIMA model outperforms the SARIMAX model that incorporates the weather data. This indicates that the chosen weather variables do not have the desired effect on the traffic flow rate.

Furthermore, the use of weather data is fairly limited as future forecasts of Amsterdam weather are not provided. This means that only past weather data can be used to predict the future. Past weather data are expected to be invaluable with regard to future predictions.

3.2.4. TBATS Model

The TBATS model achieves a MAE of 5365.72 for the A10 and 16380.73 for the A4. Considering the results of the baseline model, it can be stated that the TBATS model is adequate for the A10 location. However, it can also be seen that the TBATS model is far from adequate for the A4 location.

The TBATS models were trained with seasonalities of 7 and 365.25 days. These seasonalities are equal to a week and a year with the consideration of leap years. During the training of the model, Box-Cox transformations and ARIMA errors were used. Log-transformations of the data did not further improve the performance.

TBATS models require the residuals to be normally distributed. Using kernel density estimation plots, this assumption has been verified for both the A10 and A4 locations. Therefore, the TBATS models can be considered to be an adequate fit.

Since the TBATS model has led to inconsistent results based on the MAE for the A10 and A4 locations. The MAEs show that the TBATS model is not always better than the baseline seasonal naive model. Therefore, a model is trained for a number of other locations. The MAEs for these locations are used in the final decision regarding the use of TBATS models for the modelling of the traffic flow rate. This decision led to the use of TBATS models to model the traffic flow rate.

3.2.5. LSTM Model

The final type of model that is applied to the daily-aggregated data is the LSTM model. A grid search could be used to find the most suitable parameters for the LSTM models.

The parameters that were investigated using the grid search are described in Table 1. This grid search was conducted for both the A10 and A4 locations. The shape of the input layer of the LSTM was equal to (365,1), meaning that the LSTM will receive 365 input values. This indicates that the preceding 365 days of data are used to predict the next day.

Table 1: LSTM Grid Search Parameters

Parameter	Possible Values
Number of Epochs	10, 50, 100, 150
Batch Size	1, 5, 10, 20, 50
Number of Nodes Layer 1	1, 3, 5, 10, 20, 50, 100.
Number of Nodes Layer 2	0, 1, 3, 5, 10, 20, 50, 100
Activation Functions	hyperbolic tangent, sigmoid, ReLu

The grid search led to different models for the A10 and A4 locations. In the trade-off between complexity and performance, however, a simpler model was selected. This simpler model performed nearly as well as the model that was selected using the grid search. Moreover, the simpler model was identical for both the A10 and A4 locations. This indicates that a model with the suggested architecture generalizes well and is likely suitable for other locations. The characteristics of the final model and training procedure can be described as follows:

- Number of Epochs: 100
- Batch Size: 1
- Number of Nodes Layer 1: 20
- Number of Nodes Layer 2: 0 (indicating that the model will only have one hidden layer)
- Activation Functions: hyperbolic tangent

Using the rolling update method, or with walk forward validation the two models were validated. This validation led to a MAE of 4503.00 for the A10 and 3423.48 for the A4.

4. Discussion

Considering the results mentioned in Section 3, the MAE on the hourly data of 2019 is shown for a variety of models in Table 2. In this table, a distinction has been made between the A10 and A4 locations.

The best model that has been fitted to hourly data is the TBATS model. On average, this model performed 72% better than the seasonal naive baseline model, regarding the MAE.

In Figure 5 and Figure 6, the 2019 validation of the TBATS models has been plotted. It can be seen that the forecasts follow the observed data fairly well. An exception to this is the fact that the forecast nearly always underestimates the peaks of the observed data. One possible cause for this is that the traffic flow rate of 2018 was slightly lower than that of 2019. This trend was not as present in preceding years, which is why the TBATS model may not incorporate it sufficiently. As a result, the MAE is still notably high.

In Figure 5 and Figure 6, it can also be seen that the forecasts contain an unusual pattern, where the traffic flow rate is predicted to be low at the start of a week and higher towards the end. When this pattern is compared to the observed

Model	Location A10 MAE 2019	Location A4 MAE 2019
Seasonal Naive	2999.63	1898.77
TBATS	**777.68**	**578.10**

Table 2: Hourly Results

Model	Location A10 MAE 2019	Location A4 MAE 2019
Seasonal Naive	11110.53	6880.92
SARIMA	7976.32	5118.73
SARIMAX	11110.43	6880.82
TBATS	5365.72	16380.73
LSTM	**4503.00**	**3423.48**

Table 3: Daily-Aggregated Results

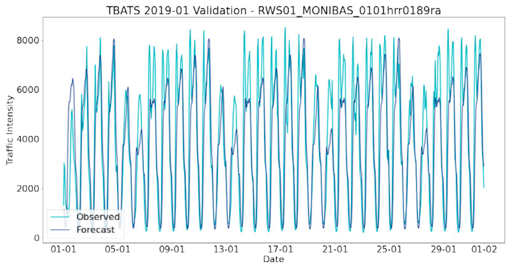

Fig. 5: The validation on 2019 of the A10 motorway using the TBATS model with hourly data. The forecast seems to follow the observed data fairly well, with the exception of the underestimation of the peaks. The ascending nature of the peaks is not found in the observed data and is likely a result of the sinusoidal nature of the TBATS model. This model performed 74.1% better than the seasonal naive baseline model, regarding the MAE.

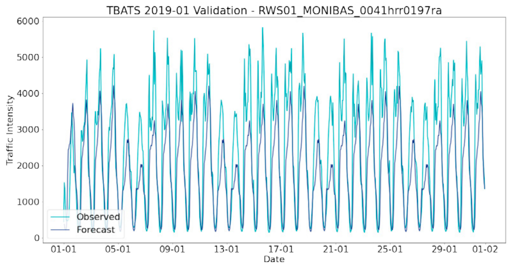

Fig. 6: The validation on 2019 of the A4 motorway using the TBATS model with hourly data. The predicted values seem to follow the observed data fairly well, with the exception of the underestimation of the peaks. The ascending nature of the peaks from Monday until Friday is not found in the observed data and is likely a result of the sinusoidal nature of the TBATS model. This model performed 69.6% better than the seasonal naive baseline model, regarding the MAE.

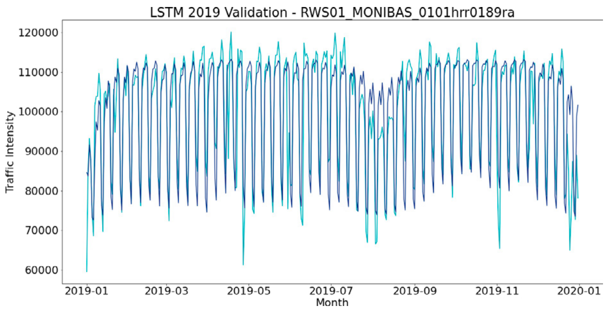

Fig. 7: The validation on 2019 of the A10 location using the LSTM model with daily-aggregated data. The forecast seems to follow the weekly and yearly seasonal patterns fairly well. This model performed 59.5% better than the seasonal naive baseline model, regarding the MAE.

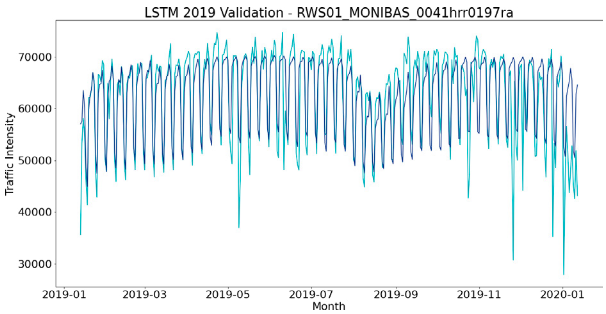

Fig. 8: The validation on 2019 of the A4 location using the LSTM model with daily-aggregated data. The forecast seems to follow the weekly and yearly seasonal patterns fairly well. This model performed 50.2% better than the seasonal naive baseline model, regarding the MAE.

traffic flow rate, it can be seen that this pattern is not equal to the observed weekly patterns. This is likely due to the sinusoidal nature of the TBATS model.

Due to the limit of computational resources, the implementation of other models to the hourly data was not possible. Therefore, a number of models have been implemented on daily-aggregated data. The MAE on the daily-aggregated data of 2019 is shown for a variety of models in Table 3. The Holt-Winters method is omitted from this table as it underperformed with respect to the seasonal naive model.

The best model that has been trained on the daily-aggregated data is the LSTM model. On average, this model performed approximately 55% better than the seasonal naive baseline model, regarding the MAE. Figure 7 shows the daily validation of 2019 for the A10 using the LSTM model and Figure 8 shows this validation for the A4. It can be seen that the two LSTM models describe the data in both figures fairly well.

5. Conclusion

Considering the results that were discussed in Section 4, it can be stated the modelling of the traffic flow rate has been carried out successfully. That said, the TBATS model is considered to be the most suitable model for the hourly traffic flow rate data. This model performs 72% better than the baseline model, the seasonal naive model. With respect to the daily-aggregated data, the LSTM model is the most suitable. This model performs 55% better than the seasonal naive model on the daily-aggregated data.

6. Future study

There are numerous, interesting additions that could enrich the current research.

Firstly, the manner in which models are implemented could be improved. The validation of the models was done on 2019 for two locations. With this, it was assumed that the model performance would be similar for other locations. In practice, this turned out not to be the case in all instances. Therefore, it is suggested that the models should be validated on all locations in the future. This ensures that the chosen model performs well on all locations. Another way in which the implementation of models can be improved is through the use of multiple evaluation metrics. In this study, the mean absolute error was used. To complement this, the use of metrics such as root mean squared error (RMSE), mean absolute percentage error (MAPE) or adjusted R-square are recommended.

Lastly, more extensions of the data could be considered. Some possible extensions were mentioned in Section 2.2. In this research, weather data were used to enrich the daily-aggregated data for a SARIMAX model. Since the weather is subject to change during a day, it is expected that the use of hourly weather data may be beneficial for the predictions. Next to the weather data, data regarding large events as well as COVID-19 statistics could be used to further enrich the data. This enables the predictions to include the effects of COVID-19. The addition of live data streams in combination with dynamic systems, would also allow for more accurate short-term predictions, accounting for unforeseen circumstances.

These additions may lead to better model performance or improved insights regarding the traffic flow rate in and around Amsterdam.

Acknowledgements

This research has been conducted in the framework of the Impact Study North/Southline research theme ”Mobility and Accessiblity” hosted at CWI and funded in part by the Municipality of Amsterdam and the regional transportation authority of Amsterdam. We would like to thank the Faculty of Science at the Vrije Universiteit Amsterdam for giving us the opportunity to initially work on the research during the Project Optimization of Business Processes (POBP).

References

[1] Brownlee, J., 2019. A gentle introduction to SARIMA for time series forecasting in Python. URL: https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/.
[2] Chou, Y.T., 2017. BATS and TBATS model. URL: https://yintingchou.com/posts/bats-and-tbats-model/.
[3] Cowpertwait, P.S.P., Metcalfe, A.V., 2009. Introductory time series with R. Springer.
[4] Esprabens, J., Arango, A., 2020. Time series for beginners. URL: https://bookdown.org/JakeEsprabens/431-Time-Series/modelling-time-series.html.
[5] Hyndman, R.J., Athanasopoulos, G., 2018. Forecasting: principles and practice. OTexts.
[6] JJ, 2016. MAE and RMSE - Which metric is better? URL: https://medium.com/human-in-a-machine-world/mae-and-rmse-which-metric-is-better-e60ac3bde13d.
[7] Kennisinstituut voor Mobiliteit, 2019. Mobiliteitsbeleid 2019. Ministerie van Infrastructuur en Waterstaat. URL: https://www.kimnet.nl/publicaties/rapporten/2019/11/12/mobiliteitsbeeld-2019-vooral-het-gebruik-van-de-trein-neemt-toe.
[8] Khandelwal, R., 2020. Time series prediction using SARIMAX. URL: https://medium.com/datadriveninvestor/time-series-prediction-using-sarimax-a6604f258c56.
[9] Nationaal Dataportaal Wegverkeer, 2021. Nationaal dataportaal wegverkeer. URL: https://ndw.nu.
[10] NHS England and NHS Improvement, 2020. Advanced forecast techniques. NHS URL: https://www.england.nhs.uk/wp-content/uploads/2020/01/advanced-forecasting-techniques.pdf.
[11] Pentaho, 2020. Time-series forecasting using TBATS model. URL: https://blog.tenthplanet.in/time-series-forecasting-tbats/.
[12] Skorupa, G., 2019. Forecasting time series with multiple seasonalities using TBATS in Python. URL: https://medium.com/intive-developers/forecasting-time-series-with-multiple-seasonalities-using-tbats-in-python-398a00ac0e8a.
[13] Tutorialspoint, 2021. Time series - Walk forward validation. URL: https://www.tutorialspoint.com/time_series/time_series_walk_forward_validation.htm.
[14] US Department of Transportation, 2018. How do weather events impact roads? URL: https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm.
[15] Voerknecht, H., 2016. De wegen slibben dicht. hoe komt CROW tot die conclusie? URL: https://www.crow.nl/kennis/bibliotheek-verkeer-en-vervoer/kennisdocumenten/de-wegen-slibben-dicht,-hoe-komt-crow-aan-die-info.
[16] Weisstein, E.W., 2020. Fourier series. URL: https://mathworld.wolfram.com/FourierSeries.html.