Scalable data-driven short-term traffic prediction

K. Friso
DAT.Mobility
Deventer, The Netherlands
kfriso@dat.nl

M.B. Tijink
University of Twente
Enschede, The Netherlands
m.b.tijink@alumnus.utwente.nl

L.J.J. Wismans
DAT.Mobility
Deventer, The Netherlands
University of Twente, Centre for Transport Studies
Enschede, The Netherlands
lwismans@dat.nl
l.j.j.wismans@utwente.nl

Abstract—**Short-term traffic prediction has a lot of potential for traffic management. However, most research has traditionally focused on either traffic models – which do not scale very well to large networks, computationally – or on data-driven methods for freeways, leaving out urban arterials completely. Urban arterials complicate traffic predictions, compared to freeways, because the non-linear effects of traffic are more pronounced on short links and with the presence of more crossings, more modalities and non-necessarily conservation of flow (parking). In this paper we consider several data-driven methods and their prediction performance for various situations, including both freeways and urban arterials, for prediction horizons from five minutes up to one week. The focus lies on predicting the traffic flow and speed given available measured data on a certain location. Thus challenges regarding data fusion, state estimation, and other methods providing complete temporal and spatial data is not addressed.**

The methods evaluated include several naive, parametric and non-parametric methods. For the evaluation of the prediction performance several weeks of data of various locations were used. Performance indicators contained the Root Mean Square Error, Mean Absolute Percentage Error and Mean Absolute Error. Because evaluating average performance might ignore the performance for non-regular traffic conditions, the evaluation also focused on non-regular traffic conditions. Especially these conditions are important in practice, because in these situations the need for accurate information is the highest.

Real-world applicability of traffic prediction requires not only accurate results, but also an indication of the accuracy for each prediction. Earlier research has mostly ignored this, leaving this up to the intuition of users of these predictions. This paper introduces a simple way to calculate confidence intervals, applicable to any traffic prediction method. For comparing these confidence intervals between different prediction horizons, an error measure for the accuracy of a confidence interval is defined. Two methods, SARIMA (Seasonal Auto-Regressive Integrated Moving Average) and NLM (Neighborhood Link Method), proved to be the best. The results also indicate key features necessary for accurate traffic prediction with data-driven methods. The results also show reasonably accurate confidence intervals, with those intervals able to adapt well to different traffic situations.

Keywords—short term traffic prediction, data-driven, SARIMA, NLM

I. INTRODUCTION

Short term traffic prediction (with predictions at most a week forward) for e.g. traffic jam predictions or the effects of a recent accident, is an emerging research field because of its high value for traffic management, logistics and mobility services. The reasons for this are clear: the applications promise less congestion and shorter travel times. Examples where such predictions can be used include route advice in navigation systems and better regulation of traffic flows, by optimizing traffic signal settings.

The inputs used for short-term traffic predictions are current and historic traffic conditions. By analyzing these inputs, e.g. with data mining techniques (data driven approach) or in combination with traffic flow theory (modelling approach), it is possible to create predictions of future traffic situations. Although modelling approaches show potential (see e.g. [6]), data driven approaches are better scalable in general (in terms of size of network and prediction horizon). Two main problems make short-term traffic predictions difficult. Firstly, traffic measurement data is not available for all traffic links, and may have gaps or incorrect data, so the data is not complete in both time and space. Secondly, traffic flows are complex and have non-linear aspects, making prediction non-trivial. On urban roads these two problems are more pronounced: there are far less measurements available compared to freeways, even though there are more roads and routing options. Traffic flow on urban roads is also more complex compared with freeways, because of low speeds, parking, more junctions and interaction with other modalities. Hence, most of earlier research has focused on freeways, but the applicability of their approaches on urban roads is not well researched yet. In this research we mainly focus on the second problem, assuming available sensor data, however including possible missing, noisy or incorrect data, providing short term predictions using data driven approaches.

II. USED DATA AND EVALUATION METHODS

Only Inductive Loop Detector Data are used, since this provides mostly complete measurement data (i.e. flows and speeds) for all used links. The data are processed computing a 10-minute moving average, producing a new measurement at

978-1-5090-6484-7/17/$31.00

time t as average of the measurements from time t−9 till time t. The reason is that the raw minute data fluctuates so much that no patterns are visible or predictable, while short term traffic prediction focuses on estimating average traffic conditions to support decision making (i.e. reliable and to some extent robust decisions). Furthermore, there are some promising method that assume the measurements as noiseless observations, which could otherwise not be used.

The measured flow for location l in time interval t is denoted by $q(l,t)$ and the measured speed by $v(l,t)$. Several predictions horizons τ will be evaluated: τ = [5 min, 10 min, 15 min, 30 min, 1 h, 2 h, 1 day, 1 week]. Furthermore, the estimators for time t at location l as predicted τ time intervals ago will be denoted

$\hat{q}_{-\tau}(l,t)$ and $\hat{v}_{-\tau}(l,t)$.

The used 14 measurement locations in Rotterdam (the Netherlands) include both freeway and urban locations and have relatively complete minute data for the used 7 weeks of data (September 1st 2014 – October 19th 2014) obtained from NDW [4]. The first 5 weeks are used for training of the methods, the last 2 weeks for evaluation.

The performance of the data-driven methods is evaluated by the RMSE (Root Mean Square Error) and for several situations regarding periods and types of roads: Rush hours, Non-rush hours, Weekend, Urban arterial locations, Freeway locations.

Some small modifications were made to the regular formulas of the performance measures to make them more appropriate for our analysis. Firstly, it is known beforehand that some measurements or predictions are incorrect or missing. These are not taken into account. Secondly, sometimes the measurements of the flow q or speed v reach or approach zero. To get useful results, these periods are not taken into account during evaluation (for all measures), namely when $q \leq 60$ veh/h or $v \leq 10$ km/h. Note that this corresponds with traffic conditions with very low traffic flow, which are not all too interesting in any case, or very high congestion for low speed situations. Fortunately, these very high congestion periods do not occur very often, only 9.4 minutes per week on average in the chosen data set.

III. Data-driven Methods

Based on the literature [5,6], short-term traffic prediction can be divided into several categories:

- **Non-parametric methods**. These methods do not use any additional model knowledge (e.g. traffic flow theory) to derive predictions. Counter-intuitively, it is allowed to use parameters in non-parametric methods, as long as these can be estimated purely from the input data. Examples of these methods are neural networks and clustering algorithms. In our research a Neural Network Method is considered.
- **Parametric methods**. These methods use additional parameters or some model to predict traffic. This research will only use models that allow the parameters to be estimated from traffic data. General examples of this category are Extended Kalman Filters, time series analysis and full-fledged traffic models. Our research covers two time series methods (SARIMA and MSTARMA) and a Neighbourhood Link Method.
- **Naive methods.** These methods are actually a subset of the previous category. They are very simplistic and no model structure or parameters are used. Examples for traffic prediction include using the historic average. This paper will cover the historic average and last observed measurement methods, along with a combination of both.

All the chosen methods satisfy the requirements of the research goal: they are scalable in number of links and can be used in a purely data-driven approach. All used methods are described in detail in [8], here the highlights of the methods are presented.

Neural Network (NN): Neural networks [6] are based on the way the human brain works: neurons activate (or trigger) based on their own inputs from other neurons, activating yet other neurons. Together, these neurons can learn complex tasks. Here, these artificial neurons use recent traffic measurements to predict future ones. Since no knowledge about traffic is used to create this model, neural networks are non-parametric methods. An advantage of neural networks is that they are theoretically able to overcome many of the problems associated with the other methods: noisy or incorrect data can be taken into account, small fluctuations can be separated from incidents and the non-linear behavior in traffic flows can be modeled.

Seasonal Auto-Regressive Integrated Moving Average (SARIMA): The SARIMA method looks at traffic measurement data as a time series. The Auto-Regressive and Moving Average models, combined together known as ARMA or ARIMA, can be extended with a periodic part. The result is SARIMA. This method looks at links individually, without any spatial correlation. Any missing or incorrect data is a problem, since the method needs data at any point in time. Here the values discovered by [7] are used: the traffic data will be modeled with a SARIMA(1,0,1) $(0,1,1)_T$ process, for each link separately, and for both predicted variables (v and q), where T is the seasonal period.

This allows one-step predictions (for flow):

$$\hat{q}_{-\tau}(l,t+\tau) = \phi_q(l) \cdot (l, t+\tau-1) + q(l, t+\tau-T) - \phi_q(l) \cdot q(l, t+\tau-T-1) + \varepsilon_{q,t+\tau} + \theta_q(l) \cdot \varepsilon_{q,t+\tau-1} + \Theta_q(l) \cdot \varepsilon_{q,t+\tau-T} + \theta_q(l)\,\Theta_q(l) \cdot \varepsilon_{q,t+\tau-T-1}$$

where any measurements $q(l, t')$ with $t' > t$ are replaced with $\hat{q}_{t-t'}(l,t')$ (so actually, one-step predictions are made τ times). For the speed predictor $\hat{v}_{-\tau}(l,t+\tau)$ a similar formula holds.

The result is the prediction, using the week-long difference as trend and using both recent and week-long moving averages. If any of those used measurements fall after time t, they are replaced by their predictions, since these measurements are not yet available for prediction at time t. If any data is missing, it is replaced by either its prediction or, if not available either, its previous value, just to make sure that the prediction process can continue.

To estimate the parameters ((l), (l), Θ(l)) for both predictors of the SARIMA process, generally a maximum likelihood estimation is used to choose the parameters. Since this particular model has only six parameters per location, a simple grid search is used to find initial estimates, which are improved using a local search.

Multivariate Spatial-Temporal Auto-Regressive Moving Average (MSTARMA): The MSTARMA model [3] is similar to SARIMA, as it also uses a time series perspective. Instead of using a seasonal component, it uses spatial information: measurements of other nearby links are used to influence the prediction. This makes it a variant of the Vector Auto-Regressive Moving Average (VARMA) model. Also like SARIMA, it has a problem with incorrect or missing data, since it assumes that all measurements are exact.

The single-step predictor of MSTARMA is defined by:

$$\hat{X}_{t+\tau,-\tau} = \sum_{i=1}^{p} (\phi_q \otimes S_i) X_{t+\tau-i} + \sum_{i=1}^{q} (\theta_q \otimes S_i) \varepsilon_{t+\tau-i}$$

where any values $X_{t'}$ with $t'>t$ are replaced by $\hat{X}_{t',t-t'}$ and the predictors of flow and speed are part of vector $\hat{X}_{t+\tau,-\tau}$ and missing values are replaced with their prediction.

Neighbourhood Link Method (NLM): The neighbourhood link method [1] uses correlations between traffic patterns of separate links to create predictions: if the traffic at a location resembles that of another location in the past, that correspondence can be used to predict future traffic. The NLM extends this simple idea by using more than one of these correlated traffic flows. Missing or incorrect data is no big problem for this method, since a prediction only depends on a few measurements. In this research only the basic variant of NLM prediction as described in [1] is used.

Let (l) be the neighbourhood vector of location l with prediction horizon τ, containing the 4 locations whose flows correlate most with the future flow at location l. The locations in the neighbourhood vector have traffic patterns resembling the traffic pattern at location l with τ time difference. It is possible that locations far away from location l in distance end up in the neighbourhood vector, if these have better correlation in their traffic patterns than nearby locations, even though that might seem unlikely at first sight. The prediction for location l, using the neighbourhood vector (l), is simply the historic average at location l plus the weighted deviation from the historic average at the neighbourhood locations (for flow):

$$\hat{q}_{-\tau}(l, t+\tau) = \overline{\mathbf{q}}(l, t+\tau) + \sum_{i=1}^{4} \mathbf{w}_{l,\tau,i}^{q} \left[q\left(N_{\tau,\mathrm{i}}(l), t\right) - \overline{\mathbf{q}}\left(N_{\tau,\mathrm{i}}(l), t\right) \right]$$

and similar for the speed predictor $\hat{v}_{-\tau}(l, t+\tau)$. To determine the weight vectors $\mathbf{w}$ a least-squares optimization of the above predictors is made using the training data.

As formulated here, the NLM does not scale to large networks, since it needs to calculate the correlation coefficient between every pair of locations. To make this approach scalable, one can simply limit the locations in the neighborhood vector to some constant number of closest locations or highest shown correlation in the past. This way, the number of correlation coefficients that need to be calculated grows linearly in number of locations.

Historic Average (HA): One of the simplest methods available is the historic average method [7]. It simply assumes that any future moment will be identical to the recorded historic average. Missing or incorrect data is not a problem, except in extreme cases, since an average is taken. Note that this means that the prediction for a certain moment in time will be the same, no matter the choice of τ.

Stationary (S): Another simple estimator is the stationary method [7]. It assumes that traffic will remain as it is now. The method is very sensitive to errors, since the last available measurement is used as prediction. These estimators use the latest available measurement (so the measurement at time t) to predict q and v at time $t + \tau$. If this measurement is missing, the measurement at the latest time t' with an available measurement is used, with $t' < t$.

Historic Average with Deviation (HAD): Since both previous methods, historic average and stationary, have clear problems (historic average with very-short term prediction and stationary with longer-term prediction) a simple combination of each will also be used [7], called historic average with deviations in which the historic average for $t + \tau$ is adjusted based on the ratio between current measurement and the historic average for the current time interval t. As a result it combines a smoothly continuing traffic curve with historic patterns to predict the changes in traffic. Like the stationary method, the HAD method is sensitive to errors.

IV. Results

All properties not related to predictive abilities of the methods can be found in Table 1. The number of parameters refers to the parameters needed by each method. These are estimated during the training process, given five weeks of training data (September 1st 2014 – October 5th 2014) of the 14 used locations. The training time needed can be found in Table 1 as well. All methods used the same locations and used the same five weeks training data as warm-up data and two weeks of testing data (October 6th 2014 – October 19th 2014). The running time is the time needed to apply the trained methods, creating a prediction for all 20160 measurements for a single prediction horizon τ. Thus, all methods are more than fast enough to use in practice. Note that training is faster than predicting for SARIMA and MSTARMA. This is because only single-step predictions were used in the training procedure (by nature of those methods) and multi-step predictions are needed for every prediction. Since all these secondary properties are good, the rest of the comparison will only focus on prediction performance.

TABLE I. Comparison of the Properties of the Several Methods

Method	# parameters to estimate	Training time	Running time
NN	$\sim 250\lvert\mathscr{T}\rvert \cdot \lvert\mathscr{L}\rvert$	$\sim \lvert\mathscr{L}\rvert$ min	$\sim \lvert\mathscr{L}\rvert$ s
SARIMA	$3\lvert\mathscr{L}\rvert$	$\sim 10\lvert\mathscr{L}\rvert$ s	$\sim 2\lvert\mathscr{L}\rvert$ min
MSTARMA	$36\lvert\mathscr{L}\rvert$	$\sim 40\lvert\mathscr{L}\rvert$ min	$\sim 10\lvert\mathscr{L}\rvert$ min
NLM	$8\lvert\mathscr{T}\rvert \cdot \lvert\mathscr{L}\rvert$	$\sim \lvert\mathscr{L}\rvert$ s	~ 0 s
HA	0	~ 0 s	~ 0 s
S	0	~ 0 s	~ 0 s
HAD	0	~ 0 s	~ 0 s

Where $\mathcal{L}$ is the set of measurement locations and $\mathcal{T}$ the set of prediction horizons.

The resulting RMSE comparison of the methods can be seen in Figure 1 for the flow predictions and Figure 2 for the speed predictions. For clarity, the lines connecting performances for the different prediction horizons $\tau \in \mathcal{T}$ is for visualization purposes and not an estimate of performance in between. Furthermore the x-as does not contain equal step sizes, so relative slopes of these lines should not be interpreted. No matter the situation (freeway or urban links, rush hours or weekend, long or short prediction intervals τ) the results remain roughly the same in shape and ordering for the researched methods.

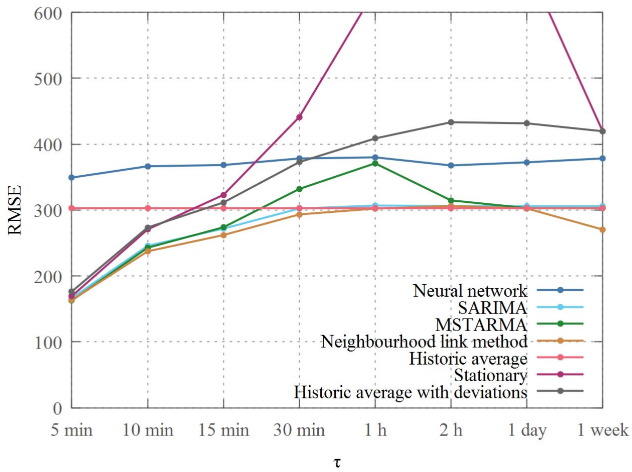

Fig. 1. RMSE of flow predictions

Figures 1 and 2 show that the historic average method has a flat line, i.e. identical performance for all prediction horizons τ. This is because the historic average prediction does not depend on prediction horizon τ or the measurements leading up to the predicted time point: so the errors at all time points only depend on the actual measurement and the historic average, which remains equal for different τ. The other methods do use those measured values, so their performance changes based on what values are available before the prediction horizon τ.

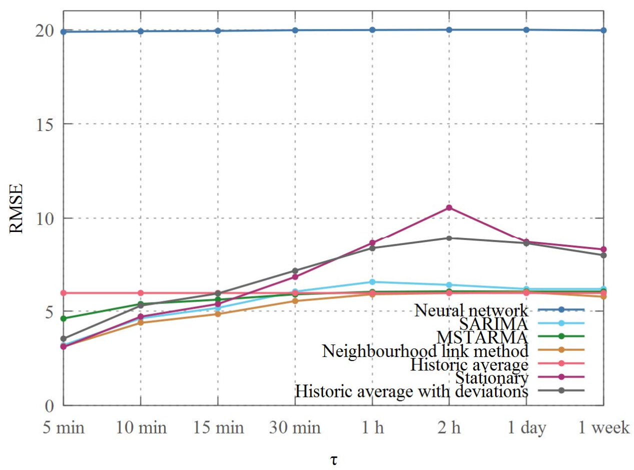

Fig. 2. RMSE of speed predictions

The historic average with deviations and stationary methods perform the worst. On time frames up to ten minutes, the results are acceptable, since the traffic does not change very much in that time frame, but longer horizons do not give acceptable results. The neural network method also does not work very well, it always predicts worse than the historic average method. Its almost-constant performance is due to bad performance at two locations, where it gave extremely high flows and speeds for some periods in the evaluation data. Leaving out those locations gives a slightly better score, but not enough to change its position compared to the other methods. All other methods, including the simple historic average method, outperform those three methods for larger prediction intervals.

Another interesting thing is that the historic average method is performing relatively well, especially for larger prediction horizons. That means that when the prediction horizon increases that taking into account current traffic conditions including current existent non-regular situations (i.e. incidents) is becoming less relevant or can even lead to worse predictions.

SARIMA and NLM do best on all prediction horizons, even though these methods are not very complicated. No conclusion can be made on which method is better, because the results are very close. The NLM prediction drops down a bit for $\tau = 1$ week, which is unexpected, since the found correlation for the neighbourhood vector is not higher. But since traffic a week ahead is more like the current traffic than a day ahead, this results is not particularly strange either. MSTARMA comes very close to SARIMA and NLM for the first prediction intervals, but for $\tau > 15$ min, it does not predict well.

From all locations and time periods in the evaluation data, a period was chosen arbitrarily to contain both a traffic incident and a normal traffic situation. When looking at this incident, from 7:30 AM to 8:45 AM in Figures 3 and 4, there is only a small difference between NLM and SARIMA. The shown incident is illustrative of the results for incidents in general, although the exact behavior is not the same for every incident. Both methods do not capture this incident very well, but SARIMA is closer to the measurements than the NLM, although they have a similar shape. The other methods do not perform better than these two, except the stationary and HAD methods, but these have other shortcomings. The NLM follows the historic average more, even on short prediction horizons, indicating that its low values on the measures are likely because of good prediction during normal situations. SARIMA performs best with regard to incidents subjectively, but even then its prediction lies somewhere between the historic average and measurements. The neural network method predicts very similar to the historic average, but still scores worse in the measures. The stationary and historic average with deviations methods look good for $\tau \leq 15$ minutes but do not predict changes in the traffic very well. Finally, the historic average method is a good baseline for prediction, but not very useful to gain insight into the current and near-future traffic situations.

Because these are only examples we identified incidents based on the difference between measurements and the historic average for a certain time period (i.e. if the measurement deviates more than 50 % from the historic average for more than ten consecutive time steps). The performance of the methods on the remaining data set shows again that SARIMA and NLM perform best for these situations.

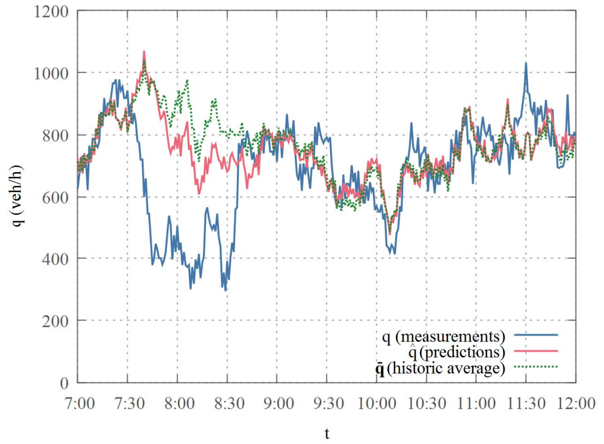

Fig. 3. Example flow prediction for NLM with τ = 15 min

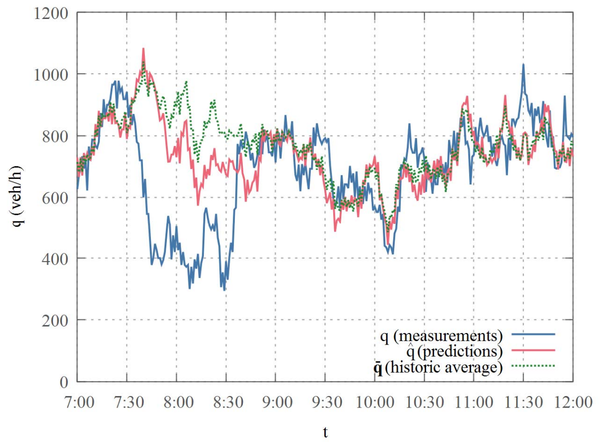

Fig. 4. Example flow prediction for SARIMA with τ = 15 min

V. PREDICTING CONFIDENCE INTERVALS

Since the predictions of the discussed methods can show large and varying errors, it is useful to see if the methods can be modified to also provide a confidence interval for their predictions. The goal of this section is to get accurate confidence intervals [8] providing a means of additional information for decision makers, not to make them as small as possible. Small confidence intervals imply accurate predictions. In this section 80%-confidence intervals will be used, to allow reasonably accurate predictions without having very large confidence intervals.

Since prediction horizons $\tau > 1$ hour were not better than the historic average, only horizons up to one hour will be evaluated. To fully evaluate a confidence interval, it is not sufficient to look at the percentage of values falling inside the interval. Certainly, to get a feeling for how good an interval is, it is useful to look at what percentage of measurements fall inside. But for comparing the accuracy of the confidence intervals, this is not enough: for example, what is a better (more accurate) result for an 80%-confidence interval? 78.1 % or 82 % of the measurements falling inside? There is no way to know just by looking at those numbers, especially if different methods do not have exactly the same number of measurements to compare.

When assuming the confidence interval is correct, the number of measurements that fall inside the confidence interval is distributed binomially: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, where n is the number of measurements, X the number of measurements inside the confidence interval (which is modeled as the random variable) and p is the chance of falling in the confidence interval.

It turns out that the logarithm of the likelihood $L = P(X = k)$ is approximately linear in n for fixed k/n, i.e. $\ln L \approx nc(k/n)$ with c a constant depending on k/n. This is a result of Stirling's approximation for factorials [2]. Since the logarithm of a likelihood preserves strict monotonicity, this allows comparing different results with this score (where the negative log-likelihood is used so a lower-valued score is a better result): $\mathrm{CI}_{\mathrm{score}}(n, k) = \frac{-\ln\left(\binom{n}{k} p^k (1-p)^{n-k}\right)}{n}$.

Note that the $\mathrm{CI}_{\mathrm{score}}$ (short for Confidence Interval score) has a lower bound of 0, since probabilities are never larger than 1. An interesting property of this score is that two different percentages of measurements falling inside the confidence interval can result in the same $\mathrm{CI}_{\mathrm{score}}$, because it is not a linear score. For example: for n = 100,000 measurements, with 85,500 and 74,200 measurements (85.5 % and 74.2 % of all measurements respectively) falling inside the interval, the result is a $\mathrm{CI}_{\mathrm{score}}$ of approximately 0.1 for both of those (with $p = 0.8$).

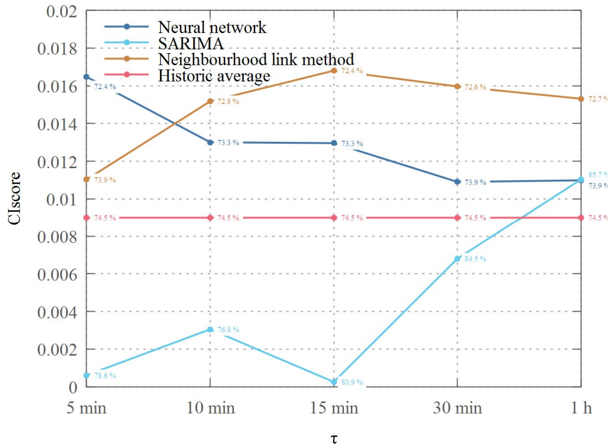

Fig. 5. $\mathrm{CI}_{\mathrm{score}}$ of flow predictions

Figures 5 and 6 show the resulting $\mathrm{CI}_{\mathrm{score}}$'s for the four methods and all evaluated prediction horizons τ. Each evaluation point has the percentage of measurements falling inside the intervals labeled next to it. As in the results for the predictions in section 5, the lines in the graphs do not indicate real values, but are only there to aid visual comparison. The graphs show that the calculated confidence intervals are not too far from the goal of 80 %: all lie between 70 % and 90 %. Also, although the $\mathrm{CI}_{\mathrm{score}}$'s and associated percentages do not remain constant for any method (except the historic average) while changing τ, they do not change very fast either. Thus, the resulting $\mathrm{CI}_{\mathrm{score}}$'s are mostly independent of the prediction horizon τ, which is useful for prediction. Some methods have more accurate confidence intervals than others but, as discussed earlier, this is not very important. The results for HA are average, for the same reason as the results in section 5: it does not depend on recent measurements or the prediction horizon τ.

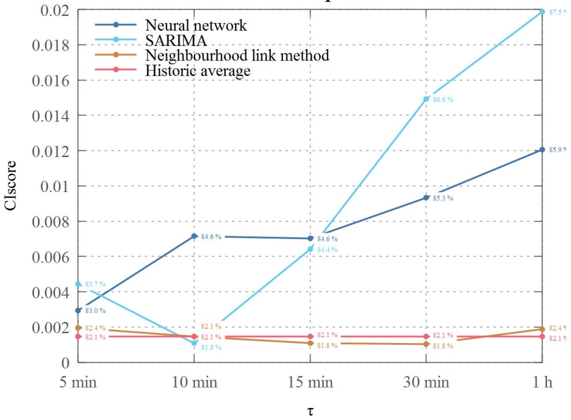

Fig. 6. $\mathrm{CI}_{\mathrm{score}}$ of speed predictions

The NN results fluctuate, but within a reasonable margin. SARIMA has a strange effect in its results: the percentage at ten minutes is lower than at five minutes, but after that it keeps increasing with larger τ. There are several possible explanations, but all such explanations relate to the fact that SARIMA is not a perfect model for traffic prediction. Whereas the NN and HA methods have their confidence intervals based on their own predictions, SARIMA has a confidence interval derived from its model. So, any inaccuracies in this model, as compared to real traffic, are magnified with larger τ. The NLM confidence intervals for flow are not the most accurate, but still reasonable. For speed the NLM intervals are quite accurate.

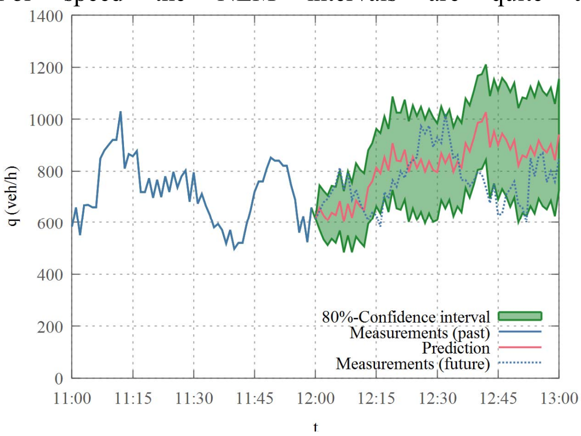

Fig. 7. SARIMA forecast at 12:00 PM

To get a feeling for the how prediction works in practice, Figure 7 shows a SARIMA flow. In this graph the prediction starting moment t remains the same and τ changes. So given all the information up to 12:00 PM, a prediction of the coming hour is made. In general, the size of the confidence interval grows with larger prediction horizon τ, converging to some limit. This is logical, because a longer time without measurements means more time for model errors to build up. Since the flow and speed are bounded, the confidence interval has a maximum size too. This growing of the confidence interval is also visible in the Figure 7.

VI. Conclusions

Short-term traffic prediction has the potential for less congestion and lower travel times, by improving navigation system information and allowing better traffic management. In this paper we looked at several data-driven methods for short-term traffic predictions, namely neural networks (NN), seasonal auto-regressive moving-average (SARIMA), multivariate spatial-temporal auto-regressive moving-average (MSTARMA), neighbourhood link method (NLM), historic average (HA), stationary (S) and historic average with deviations (HAD).

The resulting best methods were SARIMA and NLM: despite the simplicity of these methods, they performed best on all different situations, both when looking at all data and when looking at incident data. After that, the MSTARMA method performed best, but only for short prediction horizons, after which it did not perform better than the historic average method. The NN method performed worst, except for the S and HAD methods for longer prediction horizons. The secondary objectives of all methods, like computation time and scalability to larger networks, were appropriate for usage in practice.

Additionally, predictions for more than an hour ahead, using any of the researched methods, were indistinguishable from just using the historic average. In addition, it does not seem likely that any data-driven method will be able to do so. Even on shorter time frames, some methods have difficulties beating the historic average.

Some methods were expanded with confidence interval prediction, to give an accuracy estimate corresponding with a prediction. These methods, the NN, SARIMA, NLM and HA methods, gave good results, with observed intervals all between 70 % and 90 %, for a goal of 80%-confidence intervals. Thus, these confidence intervals seem good enough to use in practice.

References

[1] Luuk De Vries. “Urban Traffic State Estimation & Prediction”. MA thesis. 2016

[2] Jacques Dutka. “The early history of the factorial function”. In: Archive for History of Exact Sciences 43.3 (1991), pp. 225–249. DOI: 10.1007/BF00389433

[3] Wanli Min and Laura Wynter. “Real-time road traffic prediction with spatio-temporal correlations”.In: Transportation Research Part C: Emerging Technologies 19.4 (2011), pp. 606–616

[4] Nationale Databank Weggegevens. Oct. 3, 2016. URL: http://www.ndw.nu/

[5] C.P. Van Hinsbergen, J.W. Van Lint, and F.M. Sanders. “Short term traffic prediction models”. In: Proceedings of the 14th World Congress on Intelligent Transport Sytems (ITS), held Beijing, October 2007. 2007

[6] E.I. Vlahogianni, M.G. Karlaftis, and J.C. Golias. “Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach”. In: Transportation Research Part C: Emerging Technologies 13.3 (2005), pp. 211–234. DOI: 10.1016/j.trc.2005.04.007

[7] Billy M. Williams and Lester A. Hoel. “Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results”. In: Journal of transportation engineering 129.6 (2003), pp. 664–672

[8] M.B. Tijink, ‘Scalable data-driven short-term traffic prediction’, Internship, University of Twente, January 2017