Overcoming Traffic Sensors Malfunctions with Deep Learning

Catalán Pastor, V.; Congeduti, E.; Oliehoek, F.A.

Publication date
2022

Document Version
Accepted author manuscript

Published in
34th Benelux Conference on Artificial Intelligence (BNAIC) and the 30th Belgian Dutch Conference on Machine Learning (Benelearn)

Overcoming Traffic Sensors Malfunctions with Deep Learning

Victoria Catalán Pastor[1][0000−0003−0196−7961], Elena Congeduti[1][0000−0002−9568−6409], Aleksander Czechowski[1][0000−0002−6054−9842], and Frans A. Oliehoek[1][0000−0003−4372−5055]

Delft University of Technology, Delft 2628XE, Netherlands
{V.CatalanPastor,E.Congeduti,A.T.Czechowski,F.A.Oliehoek}@tudelft.nl

Abstract. Constant growth of cities and their rapid urbanization contribute significantly to an increase in traffic congestion, leading to high costs both in terms of time and fuel consumption. Intelligent Transportation Systems (ITSs) play an important role in managing traffic in urban areas by reducing accidents and increasing road capacity. To accomplish these tasks, these systems require live traffic data provided by different sources. The most common one are induction loop sensors, located on roads. When induction loop malfunctions occur, the data streams become unavailable, making crucial ITS services inoperative until maintenance can take place. This research explores a proxy solution to such problem: using predictions of deep learning models, such as Long Short-Term Memory networks (LSTMs) and Temporal-Convolutional Networks (TCNs) as temporary replacement data streams for faulty loops, based on data from neighboring, functioning sensors. This method is presented in a real-world scenario using data from a road segment in The Netherlands. The results show that the deep learning models can effectively predict the data from malfunctioning sensors, thus allowing to overcome the issues due to the missing information. The two models are compared on this task to conclude that even though the TCN running times are shorter, LSTM reaches similar levels of accuracy and provides more robust predictions toward short-term sensor failures.

Keywords: Traffic Prediction · Time-Series Forecasting · LSTM · TCN · Sensor Malfunction

1 Introduction

Smart city services are widely spread, continuous development and expansion as cities are constantly growing and becoming overcrowded due to urbanization and a rise in the global population [8]. In these highly populated metropolitan areas, traffic congestion is progressively becoming a critical issue. In fact, it entails considerable costs related to hurdles in the productivity of economic activities in a city and vehicle fuel consumption [2]. In addition, individuals are affected in terms of time loss, increasing stress levels and high levels of air pollution, especially during or close to traffic peak hours.

With the aim of optimizing transport networks in order to reduce congestion and provide users with real-time traffic data, Intelligent Transportation Systems (ITSs) have become an essential part of urban life. These systems consist of broad, accurate and real-time networks for transportation management [4] which strongly rely on forecasting algorithms for short-term traffic flow to identify potential congested areas. Nowadays, technological advances such as the development of 5G communication systems, or increases in the number of vehicle-detection sensors on the roads help to improve the efficiency and accuracy of ITSs. This turns into a reduction of traffic accidents and an increase in the capacity of the lanes [3].

Due to the influence of the accuracy of the data used as input for these systems, the methods applied for collecting this data are of extreme importance. The most common way of measuring the traffic flow is by employing induction-loop sensors, located on the road. These act as counters, detecting single vehicles traveling across them by measuring changes in the magnetic field [11]. The reliability of the information provided, as well as its quality, turns out to depend substantially on the failure rate of the sensors [30].

Once a failure of one or more of the on-road sensors occurs, maintenance is required to restore their functioning and continue streaming the data to the transportation systems. However, the repair costs can be unsustainably high [9]. With the high level goal of mitigating these costs, or potentially vanishing them to a certain extent, this work focuses on leveraging deep learning models for sequence predictions [12] to compensate for the lack of data due to faulty sensors. Given the extended use of these sensors, the idea is to leverage the traffic flow information collected from neighboring sensors as 'predictors' for other sensor locations. In fact, since such sensors are widely located in roads, neighboring networks of sensors are able to preserve enough information to retrieve data from few faulty sensors.

In this study, we use two state-of-the-art methods for time series forecasting, the first based on Long Short-Term Memory networks, previously used for traffic prediction problems in [1], and the second one based on convolutional networks [6]. The approach to determine the suited configuration of sensors to use as predictors, is based on the previous work [1]. In this paper, the authors distinguish between three types of sensor locations on a lane: central, entry, and exit positions. The idea is to limit the amount of inputs to the networks and reduce their complexity.

Our main contributions in this paper are the following:

- A framework which allows to replace malfunctioning traffic sensors by deep learning models.
- Evaluation and comparison of two deep learning models for this task in a real-world application domain.

2 Background

In this section, sequence modeling with deep learning models is briefly presented, together with previous research on traffic forecasting.

2.1 Deep Learning in Short-Term Traffic Forecasting

Traffic predictions can be seen as sequence forecasting problems, for which time series prediction methods can be applied. In such problems, a temporal sequence of data $(x^1, x^2, ..., x^t)$ is used as an input for a model in order to predict the value at the next time step x^{t+1}. A generalization of such problems allows different input $(x^1, x^2, \ldots, x^t)$ and output sequences $(y^1, y^2, \ldots, y^t)$ such that $(x^1, x^2, \ldots, x^t)$ serves as predictor for the target value y^{t+1}. Most of the research done for traffic forecasting in the past focused on approaches based on statistical models such as Auto Regressive Integrated Moving Average (ARIMA) and its variants [22][27].

Nowadays deep learning models have gained popularity and have been widely used for traffic prediction tasks [18][20][24], slowly overcoming statistical methods. At first, simple models, as Feed-Forward Networks, were mostly considered for time series forecasting [23], until Recurrent Neural Networks (RNNs) proved to reach better performances for sequence predictions [13]. Among them the Long Short-Term Memory network (LSTM), [14] has been extensively used in traffic forecasting [1][25][29], showing to outperform the previous ARIMA models [5].

Recently, [6] introduced the Temporal Convolutional Networks (TCNs) as convolution-based network architectures for modeling sequential data, showing comparable or superior performances to recurrent-based models on a variety of time series tasks. This specific type of network was then employed for traffic prediction problems [15][21]. Later on, it was demonstrated that it indeed achieved highly accurate predictions, outperforming LSTMs under certain conditions [28].

2.2 Approaches towards Reliable Sensor Networks

To date, many approaches to tackle sensors failures based on optimization of traffic sensor locations have been followed [16]. Moreover, other methods [17][26] suggest to compensate for the missing data by using previously collected data from the same sensor [7].

Many deep learning networks have been widely used for traffic prediction purposes. However, to the best of our knowledge, such learning models have not been employed to address the problem of data missing due to sensor systems malfunctions. Since this problem in essence consists of a traffic forecasting problem, for our study we consider state-of-the-art neural networks for this task as recurrent and temporal convolutional neural networks.

An approach for reducing the complexity of the networks for short-term traffic forecasting by means of decreasing the number of sensors used as input to the models was presented by [1]. In this case, the networks show to still be able to reconstruct the vehicle counts along the entire road segment without directly getting the data from sensors placed at the center of the lane. In this paper we follow these ideas with the main focus being to investigate the possibility of replacing a faulty sensor. Additionally to the recurrent models considered in [1], we examine and analyze the performances of another state-of-the-art neural network model for traffic forecasting, the temporal convolutional network.

Typically, traffic prediction methods make use of data with narrow time intervals, from 10 to 30 seconds, which result to be insignificant for traffic pattern recognition [19]. However, for the specific purpose of this research, which is to determine (long-term) tendencies in the traffic flow of a road, longer time intervals are preferred [20].

3 Methodology

The approach followed in our research is presented in this section. This is divided in two parts: a general overview of the problem layout, and the structure of the prediction models.

3.1 General Problem Formulation

In this study we take the same approach to the sensor arrangement as in the traffic forecasting scenario described by [1]. They divide all sensors on a road stretch into three main categories: sensors on the entry and exit points, and sensors on the centre lane.

This classification allows for recognizing the link between the different sensors. Clearly, all the vehicles detected by the sensors on the main lane have been previously counted by the sensors in the first category, and are later detected by a sensor in one of the exit points. Hence, intuitively, the information collected by the center sensors can be directly inferred by the other two types of sensors. This is the main principle behind this input data scheme: in a traffic network any sensor for which a high failure rate has been registered, is surrounded by other sensors. Therefore, thanks to the presence of neighboring sensors, any sensor can be considered as part of the central lane by simply selecting an appropriate layout and identifying the closest sensors as entry or exit points of the chosen road segment.

3.2 Traffic Flow Predictions

The forecasting process consists of three distinct steps. These are the same independently on how many (or which) sensors are used as inputs and outputs of the models.

Initialization The first step consists of creating LSTM and TCN models, initialize them and pre-process the sensors data according to the network structure. These models were generated by using the parameters presented in Table 1.

Table 1. Network initialization parameters.

	Hidden Dimension	*Learning Rate*	*Number of Layers*	*Kernel Size*	*Dropout*
LSTM	400	0.0005	1	-	-
TCN	250	0.0001	1	20	0.2

Network Training The models are trained with 80% of the collected data for several epochs. All time sequences are ordered beforehand in a completely random manner, in order to prevent the network to recognize patterns. The error is then computed by comparing the network outputs to their corresponding targets for each epoch by using the Root Mean Squared Error (RMSE), shown in Equation 1. Here, N, $\hat{y}_i$, and y_i correspond to the number of future steps to be outputted, the predicted number of vehicles, and the actual traffic flow at that time step, respectively. Then, this error is used as loss function to update the parameters on the network.

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

Model Testing The accuracy of the models is measured by feeding the networks with the remaining 20% of the data, and by again using the RMSE to compare the outputs to the actual values. This is done every epoch so as to track the improvements of the network.

The last two steps are repeated for three iterations, and then the results of these are averaged. This procedure ensures a more reliable estimation of the models accuracy.

4 Case Study

In order to determine the accuracy and effectiveness of the presented method, in this section we evaluate it in an experimental setup. The aim of this research is to present a scaled-down case study as a demonstration example. However, as potential further research direction, this methodology can be applied to larger network segments to investigate whether the promising results scale well with the size of the traffic network.

4.1 Problem Layout

In this study we analyze traffic on a road stretch from The Hague, Netherlands: the S200[1]. The section comprises 16 detector loops, counting all vehicles going above them by detecting field disturbances. Their location is shown in Figure 1.

[1] Traffic data was provided by the municipality of The Hague.

They are subdivided into two main groups: the white sensors correspond to the ones situated on the main lane, whereas the black ones refer to sensors located at ingoing/outgoing points of the road.

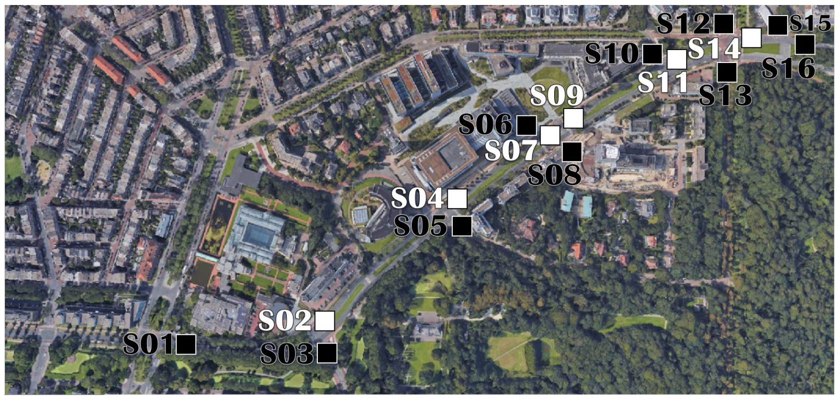

Fig. 1. Layout of the sensors on the S200 road stretch (aerial view).

We focus on the traffic flow forecasting of the sensors on the central lane (white) by means of deep learning models and using the rest of the neighboring sensors (black) as predictors. The potential scenarios that we intend to consider involve failures of one or multiple white-colored central sensors. In this case, the data collected by the black-colored sensors can be used by learning models to replace the information lost.

4.2 Data Collection

In order to train and test the models, and compare their predictions to the real vehicle counts, traffic flow information is required from all sensors. To this end, data corresponding to the entire month of November 2019 was retrieved from the 16 different sensors in intervals of 15 minutes. This month was chosen in order avoid the influence that the COVID-19 pandemic had on vehicle flow numbers, which started to affect the Netherlands in March 2020.

4.3 Results

Once the data is structured, a time sequence up to and including time step x^t is fed to the deep learning models. For this specific case study, a one-layer LSTM with a hidden dimension of 400 and a TCN with hidden dimension and dropout values of 250 and 2, respectively, are used. After giving the sensor data as input, these are then trained for several epochs (30 for the TCN and 100 for the LSTM, as the former requires less time to converge) to predict the next time step of target sequence y^{t+1}. This is done for three iterations, and the resulting

RMSE per epoch obtained from comparing the outputs to the actual traffic flow values are then averaged for allowing to obtain more reliable accuracy estimates.

In order to investigate the main research question, two different problem settings are analyzed. This helps in comparing the results and finally determining whether there is any advantage in only using the external sensors (black in Figure 1). In all the settings, only the data from the sensors on the main lane (white) is predicted.

10-6 Setting This is the main input scheme to examine in the case study. The networks only receive as input the data from the entry and exit points and use them to predict the traffic flow along the main road (white sensors), while the traffic flow predicted in this case corresponds to the data that is not fed to the model, the sensors on the main lane (white).

16-6 Setting Here data from all 16 sensors is given as an input to the models. This setting is mainly a baseline for comparing the results from the first scheme to a situation in which previous time steps of the traffic flow at the points on the central lane are also given as an input to the network.

In Table 2, the final RMSE values computed over independent test samples for both networks are shown for the two settings presented.

Table 2. Total RMSE for the LSTM and the TCN.

	Total RMSE	
	10-6 Setting	*16-6 Setting*
LSTM	131.77	198.46
TCN	115.57	117.91

5 Evaluation

In this section we analyze the performances of the predictive models over the proposed input-output scheme, which allows us to deduce whether external sensors are actually sufficient to predict internal sensors data in case of malfunctions. Moreover, LSTM and TCN results are compared, deriving conclusions on the effectiveness of the deep learning models.

5.1 Methodology Assessment

Here the goal is to focus on the input scheme applied, and determine whether it can act as a solution for the traffic sensors malfunctioning problem. For this, the results of the presented settings are compared. These two main cases differ in set of input variables for the deep learning models. As aforementioned, for the main scheme (10-6 setting), the historical data of central sensors that we want to predict are not given as input, while in the 16-6 setting, past information from all traffic sensors is fed to the networks. Hence, in the first setting, the models do not foresee any information from the sensors to be predicted but rely only on the ones at the entry and exit points of the main lane.

By analyzing the outcomes of the networks for both settings, we observe that not only the accuracy of the predictions is not reduced due to the removal of part of the input data, but it even improves. A reduction in the amount of input data results on a lower dimension of the input space of the networks, which accelerates their learning process. Moreover, from this behavior, it can be concluded that the history of the sensors on the centre lane (white) does not add essential information, given the previous history of the inflow/outflow sensors (black). Notice that especially in the case of the LSTM, the total RMSE values decrease by more than 60 points when excluding the internal sensors from the set of input sensors. This can also be seen by looking at the prediction plots which confirm the substantial improvement.

The results presented in Table 2 suggest that the LSTM in fact benefits substantially from only including the information of the external sensors, and gains no additional insight from the past history of the sensors on the central lane. On the other hand, the RMSE of the predictions made by the TCN is similar, showing that not inputting information from the white sensors does not affect the results of this network.

Therefore, these results show that it is possible to make accurate predictions on the central road segment by making use of the only history of sensors placed at the intersections with crossing lanes. This suggests that, in case of malfunctions, the data generated by a deep learning model previously trained using only crossing roads input data, can be effectively used to replace the temporarily missing or incomplete data from the failing sensors.

This research may also establish a starting point for continuous traffic forecasting, in which the future traffic flow at any point on a road could be estimated without requiring sensors on those locations. Moreover, this would mean using small amounts of input data, which would highly reduce the complexity of the networks.

5.2 Model Comparison

Firstly, the performance of both deep learning models is compared by means of analyzing their predictions and accuracy.

As it has been shown in the previous section, the RMSE values suggest that the predictions made by the TCN for the baseline setting 16-6 are more accurate than the ones from the LSTM. In fact, the predictions for one day in the 16-6 settings depicted in Figure 2 show that the LSTM does not follow closely the target curve. On the contrary, the TCN seems to provide better approximations. This might be the mere effect of too early stopping in training the LSTM which suffers more from the high dimensionality of input space, hence requiring more time to converge. As a result, the LSTM fails in reaching the peaks of the target curves, thus achieving lower performances compared to the TCN.

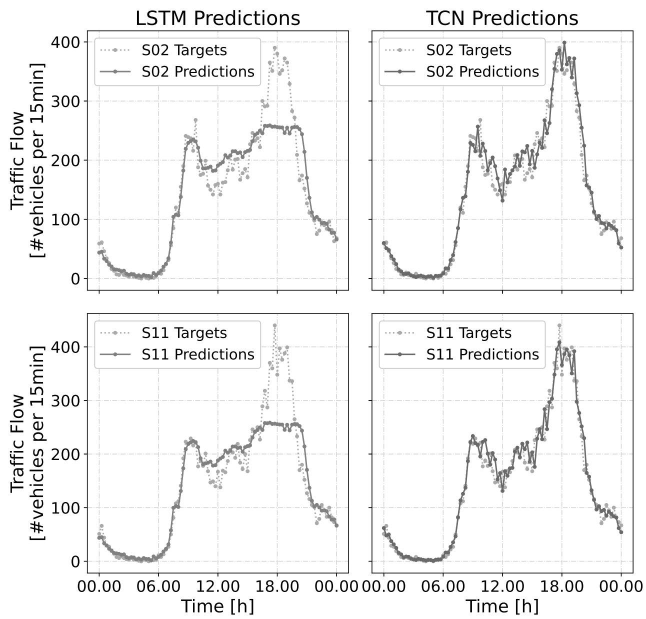

Fig. 2. Targets vs. predictions plots for the 16-6 setting.

For the main input scheme, the 10-6, the traffic forecasting for the same sensors as before is shown in Figure 3. Now, the LSTM predictions start also to reach higher target values, getting closer to the peaks. This is reflected by the decrease of the RMSE for the LSTM that we observe when comparing the setting 16-6 and 10-6 as in Table 2. On the other hand, the forecasting made by the TCN behaves in a fairly similar way for both settings, which matches the closeness of the error as well.

This implies that using the past records of target sensors that need to be predicted as input for the networks, does not improve the performances and may even prevent the networks to converge faster, hence not facilitating the learning process. One explanation is that deep learning models might require extra resources, in terms of training iterations or network size, in order to effectively extract relevant information from input data with bigger dimension. Since the outgoing sensors already retain all relevant information for the problem, adding the histories of the target sensors will only slow down the network learning process.

Once the models and their predictions are analyzed, it is also interesting to look at their errors over training runtime. In this case, the RMSE for each of the networks at every epoch was recorded, together with the runtime. Here, since

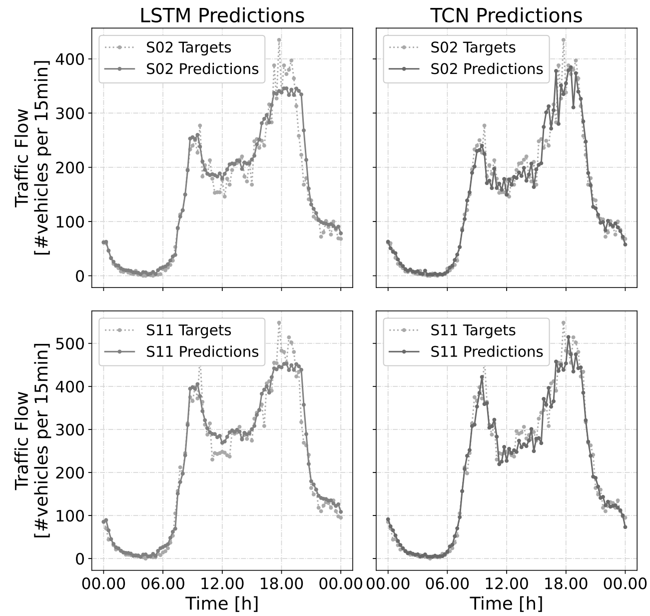

Fig. 3. Targets vs. predictions plots for the 10-6 setting.

the LSTM was trained over 100 epochs, while the TCN only over 30, it is to be expected that the total runtime for the LSTM is higher than for the TCN. Nonetheless, it is still possible to compare performances measured by the RMSE over runtime.

For the first setting, 16-6, the performance of the two networks is plotted in Figure 4. From this, it can be clearly seen that not only the TCN has a shorter convergence time, but it also achieves much lower RMSE values when compared to the LSTM. Moreover, it is observed that the accuracy of the recurrent network is not stable in this case.

Likewise, the performances as function of training runtime for the main scheme are shown in Figure 5. When compared to the 16-6 setting, the LSTM converges faster, resulting in a lower total error. On the other hand, the TCN presents an almost identical behavior when compared to the previous setting. As a conclusion, the two networks seem to achieve similar results in terms of accuracy at convergence.

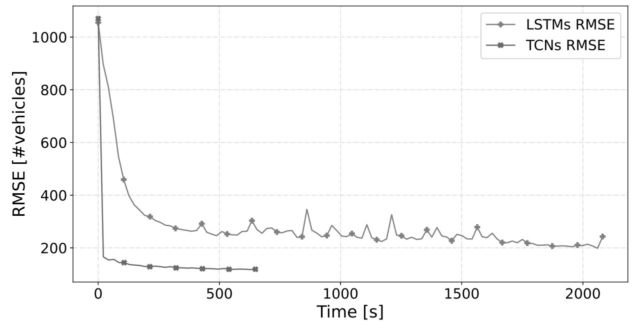

Fig. 4. Performance plot for the 16-6 setting.

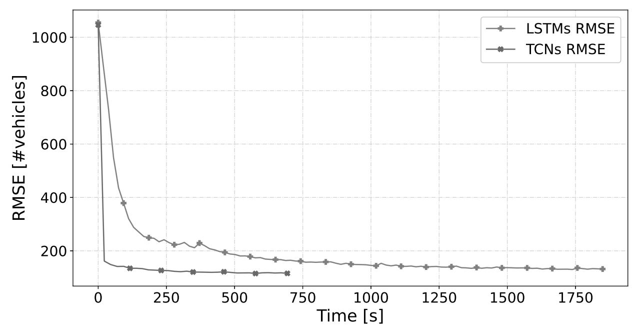

Fig. 5. Performance plot for the 10-6 setting.

From this analysis we conclude that the TCN presents overall a higher accuracy and lower convergence time when compared to the LSTM. In case the input space is not too large as in the 10-6 setting, the accuracies of both networks are similar, although the LSTM requires longer training time to reach convergence than the TCN.

5.3 Model Robustness against Noisy Data

Robustness of our method can be assessed by taking a look at how the deep learning models behave when there is a malfunction in the sensors, and, as a result, the data is noisy or incorrect. In our care, the data used to evaluate performances contained an instance of such malfunction. That is, no vehicle was registered by any of the sensors for a period of 30 minutes, corresponding to two consecutive data points as Figure 6 shows. This issue implies that the networks receive incorrect traffic counts which may potentially affect the quality of the predictions. Here the scenario examined is the 10-6 setting. All the 16 sensors presented the same type of dropping behaviour as represented for one specific sensor by Figure 6. Therefore, the input sequences were also affected by this issue.

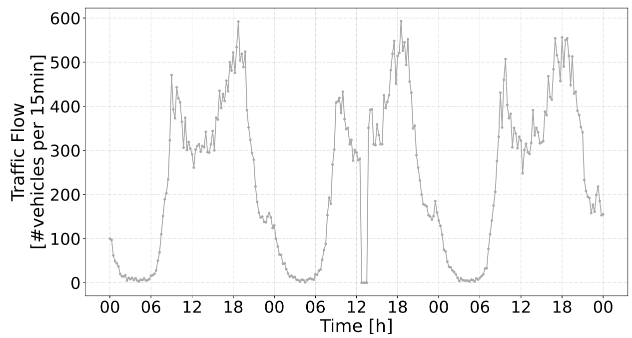

Fig. 6. Malfunctioning in the sensor data for 30 minutes.

The traffic forecasting of both networks (LSTM and TCN) for the specific day presenting the malfunctioning is shown in Figure 7. Note that the traffic flow drops to zero in two consecutive data points. This may potentially affect the model outputs.

The intention is to evaluate to what extent the networks are affected and whether in this kind of situations the deep learning models are still able to provide useful predictions. Looking at the results for the TCN, we observe that its predicted vehicle number also falls off right after the detected malfunction. On the other hand, this does not occur for the LSTM. This network keeps producing sound predictions and seems to be less prone to error than the TCN in this situations. As a conclusion, the LSTM demonstrates a higher level of robustness towards inaccurate input data compared to TCN.

6 Conclusion

In this paper we evaluate the application of deep learning models to cases of sensors failure or malfunctioning. Following previous studies, sensors are classified into two main categories: those on the main lane, and on entry/exit points. The key idea is that the information collected by the sensors on the main lane is already recorded by sensors placed at crossing points. Therefore the traffic flow on the main lane can be predicted by exploiting advances in deep learning for sequence prediction from the temporal records of the neighbouring entry/exit points. As a consequence, the data stream from a faulty sensor can be replaced by deep learning model's predictions.

The main focus of this work is to demonstrate the effectiveness of such approach. To this end, such method has been tested in a real-world case study

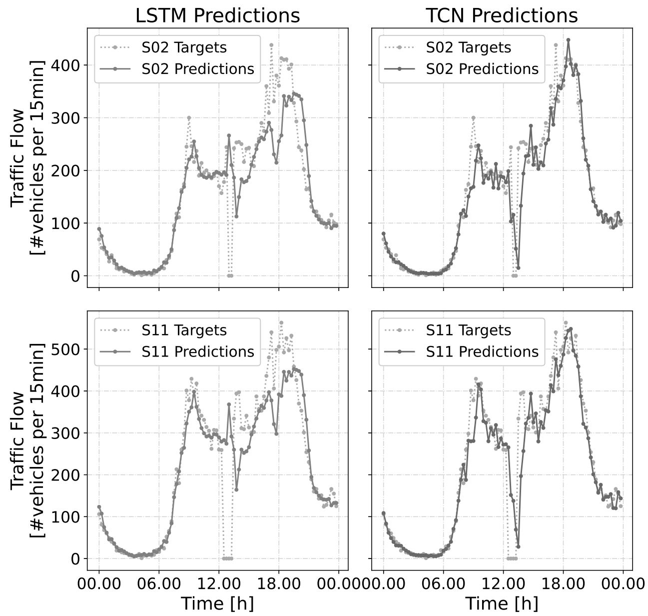

Fig. 7. Targets vs. predictions plots in the event of a short-term sensor malfunction.

using traffic data from a road stretch in The Hague (Netherlands). The performances of two different deep learning models, an LSTM and a TCN, have been analyzed and compared for this task.

From this preliminary study, we can conclude that the deep learning models act as good proxies for the potential malfunctioning sensors. Moreover, the quality of the learning model predictions remain unaffected or even improve when the past history of the malfunctioning sensors is not used as predictor for the future data.

Furthermore, a comparison between the two networks is carried out in terms of accuracy and runtime. From this analysis, we conclude that the TCN converges to its final error much faster than the LSTM. However, the latter can reach the same level of accuracy and it shows robustness towards generalized breakdowns of the sensors network.

Future research could build up from this paper by applying the methodology here presented in other, potentially larger, scenarios to further analyze its effectiveness for diverse traffic networks topologies or different sensors configuration and scalability to larger and more complex traffic networks. Moreover, different networks, such as Graph Neural Networks (GNNs) which would account for spatial dependencies between sensors, could be tested. Additionally, external data sources, such as weather forecasts, weekday/end labels, or seasonal impact, could be used to augment the accuracy of forecasting [10].

Acknowledgment

We would like to thank Rogier Hoek, from the municipality of The Hague (Netherlands), who supplied the data used in this study.

This project had received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No.758824 —INFLUENCE).

References

1. Z. Abbas, A. Al-Shishtawy, S. Girdzijauskas, and V. Vlassov, "Short-term traffic prediction using long short-term memory neural networks", in *2018 IEEE International Congress on Big Data (BigData Congress)*, pp. 57-65, Jul 2018.
2. M. S. Ali, M. Adnan, S. M. Noman, & S. F. A. Baqueri, "Estimation of traffic congestion cost-A case study of a major arterial in Karachi", in *Procedia Engineering*, vol. 77, pp. 37–44, 2014.
3. N. Aljeri, and A. Boukerche, "Movement prediction models for vehicular networks: an empirical analysis", in *Wireless Netw 25*, pp. 1505–1518, 2019.
4. S.H. An, B.H. Lee, and D.R Shin, "A Survey of Intelligent Transportation Systems", in *2011 IEEE Third International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN)*, pp. 332-337, Jul 2011.
5. A. Azari, P. Papapetrou, S. Denic, and G. Peters, "Cellular Traffic Prediction and Classification: A Comparative Evaluation of LSTM and ARIMA", in *Discovery Science*, pp. 129-144, Oct 2019.
6. S. Bai, J. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling", in *ArXiv: Learning*, Mar 2018.
7. L. Beretta and A. Santaniello, "Nearest neighbor imputation algorithms: a critical evaluation", in *BMC Medical Informatics and Decision Making*, vol. 16, no. 74, Jul 2016.
8. D. Brockmann, L. Hufnagel, T. & Geisel, "The scaling laws of human travel", in *Nature*, vol. 439, pp. 462-465, Feb 2006.
9. A. Danczyk, X. Di, and H. X. Liu, "A probabilistic optimization model for allocating freeway sensors", in *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 378-398, 2016.
10. S. Dunne and B. Ghosh, "Weather adaptive traffic prediction using neurowavelet models", in *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 370-379, Mar 2013.
11. Federal Highway Administration, "Traffic Detector Handbook: Third Edition", vol. 2, 2006.
12. J. C. B. Gamboa, "Deep learning for time-series analysis", in *ArXiv*, 2017.
13. H. Hewamalage, "Recurrent neural networks for time series forecasting: Current status and future directions" in *International Journal of Forecasting*, vol. 37, pp. 388-427, Aug 2020.

14. S. Hochreiter and J. Schmidhuber, "Long short-term memory" in *Neural Computation*, vol. 9, pp. 1735-1780, Dec 1997.
15. L. Kuang, C. Hua, J. Wu, et al., "Traffic Volume Prediction Based on Multi-Sources GPS Trajectory Data by Temporal Convolutional Network", in *Mobile Networks and Applications*, vol. 25, pp. 1405–1417, 2020.
16. X. Li and Y. Ouyang, "Reliable sensor deployment for network traffic surveillance", in *Transportation Research Part B: Methodological*, vol. 45, pp. 218-231, 2011.
17. L. Li, Y. Li, and Z. Li, "Efficient missing data imputing for traffic flow by considering temporal and spatial dependence", in *Transportation Research Part C: Emerging Technologies*, vol. 34, pp. 108-120, 2013.
18. Z. Liu, Z. Li, K. Wu, and M. Li, "Urban Traffic Prediction from Mobility Data Using Deep Learning," in *IEEE Network*, vol. 32, no. 4, pp. 40-46, Aug 2018.
19. X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction", in *Sensors*, vol. 17, 818, Apr 2017.
20. A. M. Nagy and V. Simon, "Survey on traffic prediction in smart cities", in *Pervasive and Mobile Computing*, vol. 50, pp. 148-163, Oct 2018.
21. Y. Ren, D. Zhao, D. Luo, H. Ma and P. Duan, "Global-Local Temporal Convolutional Network for Traffic Flow Prediction," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 1578-1584, Feb 2022.
22. Y. Shu, M. Yu, J. Liu, and O. W. W. Yang, "Wireless traffic modeling and prediction using seasonal ARIMA models", in *IEEE International Conference on Communications*, vol. 3, pp. 1675-1679, 2003.
23. Z. Tang and P. Fishwick, "Feedforward neural nets as models for time series forecasting", in *ORSA Journal on Computing*, vol. 5, pp. 374-385, Nov 1993.
24. D. A. Tedjopurnomo, Z. Bao, B. Zheng, F. Choudhury, and A. K. Qin. "A Survey on Modern Deep Neural Network for Traffic Prediction: Trends, Methods and Challenges," in *IEEE Transactions on Knowledge and Data Engineering*, Jun 2020.
25. H. D. Trinh, L. Giupponi, and P. Dini, "Mobile Traffic Prediction from Raw Data Using LSTM Networks," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1827-1832, 2018.
26. P. Ye and D. Wen, "Optimal Traffic Sensor Location for Origin–Destination Estimation Using a Compressed Sensing Framework," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 7, pp. 1857-1866, July 2017.
27. H. Z. Moayedi and M. A. Masnadi-Shirazi, "Arima model for network traffic prediction and anomaly detection," in *2008 International Symposium on Information Technology*, pp. 1-6, 2008.
28. W. Zhao, Y. Gao, T. Ji, X. Wan, F. Ye, and G. Bai, "Deep temporal convolutional networks for short-term traffic flow forecasting", in *IEEE Access*, vol. 7, pp. 114496-114507, 2019.
29. Z. Zhao, W. Chen, X. Wu, P. C. Chen, and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast", in *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68-75, 2017.
30. N. Zhu, S. Ma, and L. Zheng, "Travel time estimation oriented freeway sensor placement problem considering sensor failure", in *Journal of Intelligent Transportation Systems*, vol. 21, pp. 26-40, Jan 2017.