Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 207 (2022) 552–562

Procedia Computer Science

www.elsevier.com/locate/procedia

26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Neural Network-based Model for Traffic Prediction in the City of Valencia

Cristian Villarroya[1,1], Carlos T. Calafate[1,], Eva Onaindia[1], Juan-Carlos Cano[1], Francisco J. Martinez[1]

[a]*Departament of Computer Engineering (DISCA), Universitat Politècnica de València, Camino de Vera S/N, 46022 Valencia, Spain*
[b]*Departament of Computer Science (DSIC), Universitat Politècnica de València, Camino de Vera S/N, 46022 Valencia, Spain*
[c]*Computer Science and System Engineering Department (DIIS), University of Zaragoza, Ciudad Escolar S/N 44003 Teruel, Spain*

Abstract

There are many models that attempt to predict vehicular speed in urban and interurban roads, the noise pollution caused by traffic in cities, or even the traffic flow based on historical data from cameras or from people's mobile phones. Such information can be useful for administration authorities, and for researchers attempting to improve the living conditions of citizens. In this context, the aim of the present study is to design a model capable of predicting the traffic flow in the city of Valencia, Spain, based on data collected by electromagnetic loops distributed throughout the city. With a good traffic prediction, it will be possible to foresee possible traffic jams, and also to trigger countermeasures to mitigate them. Therefore, two models based on two recurrent neural networks of Long Short-Term Memory (LSTM) type have been designed to predict the traffic flow in the different streets of Valencia at the different hours of the day. We also study the influence of the specific characteristics used on the accuracy of the model. The results of our experiments show that, despite the high heterogeneity in terms of per-street traffic behaviour, it is possible to reach useful prediction models with low errors.

Keywords: Forecasting; traffic flow; time series; deep learning; LSTM.

1. Introduction

With the development of smart cities, intelligent transportation systems (ITS) have been deployed around the world to fix or alleviate the problems associated with traffic [18]. Urban traffic management systems collect data about the state of traffic from the urban traffic network, such as volume, occupancy and speed. This data can come from a variety of sources, such as GPS systems, loop detectors, radio frequency identification (RFID), or other systems. In the case of

* Corresponding author. Tel.: +34-96-387-7007
E-mail address: calafate@disca.upv.es

10.1016/j.procs.2022.09.110

this work, the traffic flow information comes from electromagnetic loops distributed throughout the city of Valencia, counting the number of vehicles per hour on the city's streets.

Traffic flow prediction is a fundamental issue for researchers and also a challenging one, as traffic flow is often highly non-linear and complex in its patterns [2]. In addition, out-of-the-ordinary events, such as this year's pandemic, or something more common, including mass events like a concert or a football match, can drastically alter the normal behaviour of traffic flow [26]. Being able to predict traffic flow in large cities provides a great benefit in multiple areas, such as vehicle routing or traffic congestion management. Good traffic flow prediction can help to avoid traffic congestion by being able to reroute a vehicle to a less congested route.

A particularly useful technique for predicting traffic flow is to rely on long short-term memory (LSTM) networks [9]. LSTM networks are a type of recurrent neural networks that are suitable for modelling serial data such as time series. One of their main advantages is that they are able to process long sequences of information, such as traffic flow. For this reason, in the scope of this work, we find that they can be particularly useful when targeting the changing behaviour of traffic flow levels in different streets along the day. Traffic prediction with LSTM has gained attention lately due to the ability of LSTM to deal with time long series and calculate optimal time lags [17]. Most of the works apply traffic prediction to limited areas such as a road network with four expressways [16], a few locations representing several entry and exit reference points between neighbourhoods in New York City [3], or twenty-one loop-detectors installed on a northbound section of an interstate in Chicago [19]. In contrast, in this work we propose the application of a Deep Learning approach to predict the traffic flow of an entire city with a population of almost 1,5 million people.

This paper proposes two solutions that allow predicting, for each hour of the day, the traffic flow in the streets of the city of Valencia, Spain. To this end, two novel short-term traffic flow prediction models are defined based on recurrent neural networks; specifically, we rely on LSTMs to create our models, and on metrics like the Mean Absolute Percentage Error (MAPE) [5] and the Symmetric Mean Absolute Percentage Error (sMAPE) [12] to validate the quality of our models. Notice that we discard using the MSE metric as it is scale dependant, and in our case we have streets with very heterogeneous traffic flow levels, and so we have decided to use MAPE and sMAPE only.

The remainder of this paper is organised as follows: a brief overview of related work, and an explanation of time series, take place in Section 2. Then, a characterisation of available data is presented in Section 3. Section 4 details the methodology followed in this work. The evaluation of the proposed models is then performed in Section 5, including the results' discussion. Finally, Section 6 presents the main conclusions, and refers to future research works.

2. Background and Related work

In this section we provide some background on relevant areas for this work: (i) time series, (ii) Autoregressive integrated moving average (ARIMA) models, (iii) Neural network-based models, and (iv) mixed models.

2.1. Time series

A time series is defined as a collection of observations of a variable collected sequentially over time; these observations are usually collected at equidistant time intervals. This is intended not only to explain past events, but also to be able to predict the future by looking at past samples. Time series are usually modelled by a stochastic process, i.e, a sequence of random variables. Typically, by knowing the value of a variable at a time t, the value of that variable at a time $t + 1$ can be predicted.

There are three important aspects to take into account when defining a time series, and these are stationarity, seasonality, and autocorrelation. A time series is said to be stationary if its statistical properties do not change over time, i.e, it has a constant mean and variance. Furthermore, within stationarity, we study the trend of the data, which can be defined as a long-run change that occurs relative to the mean of the data, or the long-run change in the mean. In other words, the trend shows whether the data is increasing, decreasing, or remaining constant over time. On the other hand, seasonality refers to periodic fluctuations in the data; for example, electricity consumption is normally high during the day, and low at night. These fluctuations can occur at different time levels, and can be hourly, daily, weekly, monthly, etc. Finally, autocorrelation refers to the similarity between observations as a function of the time lag between them, i.e, how closely related a time series is to a time lagged version of itself.

Another important part is the decomposition of the time series. Two types of decomposition models can be done: (i) multiplicative models where the time series is the result of multiplying the elements, and (ii) additive models, which are the result of adding the element of the series. The additive model follows the equation:

$$Y_t = T_t + S_t + e_t \tag{1}$$

and the multiplicative model:

$$Y_t = T_t \cdot S_t \cdot e_t \tag{2}$$

where Y is the time series, T the trend, S the seasonality, and e the residual.

Trend and seasonality have been discussed above, while the residual is the result of subtracting the seasonality and trend from the time series or dividing them, depending on the type of decomposition model used.

2.2. ARIMA models

The ARIMA model is a dynamic model using time series data developed by Box and Jenkins in 1970 [23]. The ARIMA model allows a value to be described as a linear function of past data and random errors and, in addition, allows for the inclusion of a cyclical or seasonal component.

ARIMA models have been widely applied in traffic forecasting problems like predicting the traffic flow in a road of Beijing, China [7], where authors point that most of the state-of-the-art work trains the models with the full time series, without splits. However, they think that the time series should be split, so they have divided the time series into different days. In addition, within the days, they difference three periods: the morning peak period, the evening peak period and the normal period (non-peak period). They conclude their work by comparing their model with a model without splitting the time series, where it can be seen that they have been able to improve the results.

The authors in [13] explain that one of the problems of using ARIMA models is that they require a solid database for the construction of the model. Therefore, they propose a solution to this problem by using a seasonal ARIMA model (SARIMA) [11] for short-term traffic flow prediction with limited input data. SARIMA models are an extension of ARIMA models which capture the purely seasonal behaviour of a series, and do not have to be removed from the data, as with ARIMA models.

2.3. Neural network-based models

One of the most widely models used nowadays are Neural Networks (NN) due to their high efficiency. In [22], the authors propose a city-wide short-term traffic flow prediction model based on a deep convolutional NN, called TFFNET (Traffic Flow Forecasting NETwork). The model takes into account the spatial and temporal dependencies of traffic flow, and it is also able to take into account external factors such as accidents, holiday periods, or other events. In [20], the authors propose a model based on a deep neural network to predict traffic flow for a full day in the city of Seattle, USA. The model uses a multi-layer supervised algorithm to extract the potential relationship between traffic flow data, and a combination of key contextual factors. The authors in [4] introduce a short-term traffic flow prediction model that combines seasonal analysis with a gated recurrent unit (GRU) neural network. The model uses a spatio-temporal feature detection algorithm to define the optimal input time interval and spatial data volume; the GRU network then processes the feature information to make predictions.

2.4. Mixed models

Much of the research in traffic flow prediction proposes hybrid models. For example, in [24] the authors propose a traffic flow prediction model for a freeway based on a combination of the multilayer perceptron (MLP) and the

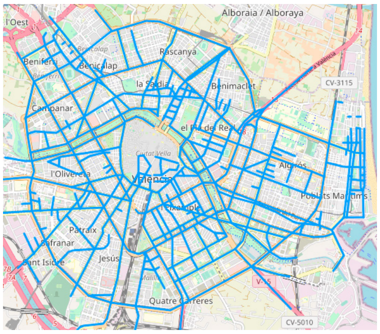


Fig. 1: Streets of Valencia with traffic flow data availability.

random forest (RF) methods. For this purpose, they use traffic flow data on a freeway in China and, in addition, they add weather information, local holiday data, toll data on the freeway, and other data sources. The final prediction will be a combination of both, MLP and RF, by solving the weights corresponding to the models using the minimum error as the objective function. Li et al. [15] consider combining an ARIMA model and a radial basis function neural network (RBF-ANN) to predict short-term traffic flows. The authors use the ARIMA model to model the linear component of the traffic flow time series, and then use the RBF-ANN model to capture the non-linear component of the ARIMA model residuals modelling. The final prediction will be a combination of the prediction results from both models. Du et al. [8] present a hybrid multi-modal deep learning method for short-term traffic flow forecasting using a dataset of real traffic on UK freeways. The model consists of a convolutional network and, on the other hand, GRU neural networks with an attention model. The convolutional layer is responsible for capturing the non-linear characteristics of the traffic data, and the GRU with the attention model is responsible for capturing the long-term temporal dependencies. In [25], the authors suggest a hybrid model which is able to deal with spatial and temporal dependencies of the data, and also of considering the influence of indirect neighbour nodes on current nodes, by combining its graph-based neural network, called ST-ChebNet, with an LSTM network.

To conclude this section, in this work we propose two models based on LSTM networks, a type of network that is widely used in traffic flow prediction. Despite plenty of work has been done in traffic flow prediction, nearly all of them focus on highways or motorways, not in urban traffic. From the related work presented in this paper, only one of them deals with city-wide traffic flow prediction, evidencing its complexity compared to highway predictions. Therefore, the novelty of this paper resides in being able to predict traffic flow in the entire city of Valencia, which is a city with very heterogeneous traffic patterns that complicate the analysis.

Our aim is to obtain models that can predict the traffic flow with an acceptable accuracy, and to study the influence of the specific characteristics used in the model on the accuracy achieved.

3. Data exploration

For our work in this paper, we have collected data from a web API of the city council of Valencia [1] that provides the number of vehicles per hour in 386 streets of Valencia. The data set is composed of 3 months, from April to June 2021, and for every street with data available in Valencia, measuring, for each of the days, the number of vehicles per hour. The streets for which information is available are highlighted in Figure 1. In order to guide the experiments to the best configuration, the stationarity, seasonality and autocorrelation properties of the time series have been analysed. To study the stationarity of the time series, the augmented Dickey-Fuller test (ADF) [6] has been used. The idea of this test is that it allows us to determine the intensity with which a time series is defined by a trend.

The null hypothesis of the test consists in that the time series can be represented by a unit root, which is non-stationary, i.e, it has some time-dependent structure. The alternative hypothesis, which rejects the null hypothesis, implies that the time series is stationary. The result is interpreted using the p-value of the test. A p-value below a threshold, usually 5% or 1%, indicates that the null hypothesis is rejected, and, therefore, the series is stationary; conversely, a value above the threshold indicates that the null hypothesis is not rejected, and, therefore, the series is non-stationary. There is also the ADF statistic value where, the more positive its value is, the more likely is it to accept

Table 1: Results of the ADF test.

ADF Parameter	Value
p-value	0
ADF statistic	-16.422
critical value 1%	-3.433
critical value 5%	-2.863
critical value 10%	-2.567

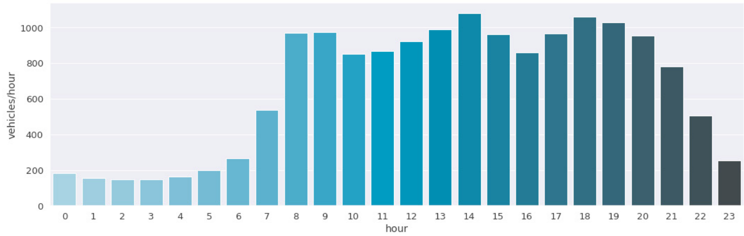


Fig. 2: Average number of vehicles/hour per hour in Valencia during a day.

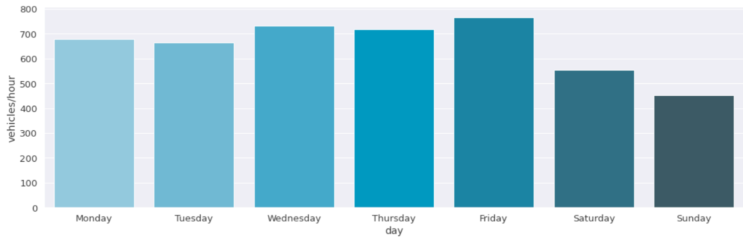


Fig. 3: Average number of vehicles/hour per week day in Valencia.

the null hypothesis, and the more negative it is, the more likely it is that the series is stationary. Finally, the critical values refer to the confidence intervals by which the hypothesis will be rejected or not. For example, the critical value of 1% allows the null hypothesis to be rejected or accepted with a 99% confidence. When applying this test to our time series, we obtained the results shown in Table 1. The results show that our time series is stationary.

With the aim of analysing seasonality, the data for all streets have been aggregated for each of the hours in the time series. Then, we have used the *statsmodels* library [21], which allows us to decompose the time series into 4 graphs: (i) observed data, (ii) trend, (iii) seasonality, and (iv) residual. In this paper, an additive model has been used to decompose the time series, and we have used a daily frequency to perform this analysis.

The trend of the time series is shown in Figure 4. We can observe that the time series is increasing slightly, which is not sufficient to consider the series as non-stationary, according to the previous results shown in Table 1. We can also see that there are fluctuations in the graph. These fluctuations are due to the fact that, on weekends, the traffic flow is lower than on weekdays. To prove this, in Figure 3 we show the average amount of vehicles/hour per day in the city of Valencia.

Regarding seasonality, which is shown in Figure 5, we can see the fluctuations in the series that reflect the fact that, in the evening hours, there is a lower traffic flow than in the daytime hours. We can also see this in Figure 2, where we show the average amount of vehicles/hour on each hour in the city of Valencia.

Finally, we show the residual part in Figure 6. In general, the error remains centred around zero and does not fluctuate excessively. The fewer fluctuations in the error, the better the decomposition of the time series and, therefore, the better the model fits the data, making it possible to obtain better predictions.

Another important part deserving more scrutiny is the autocorrelation of the time series, since it represents the similarity of a time series to a time-lagged version of itself; in particular, it allows knowing with how many time steps the best prediction will be obtained The autocorrelation values as a function of lag values can be seen in Figure 7. The

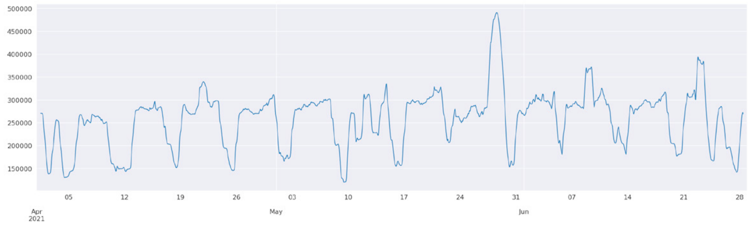


Fig. 4: Trend of the time series.

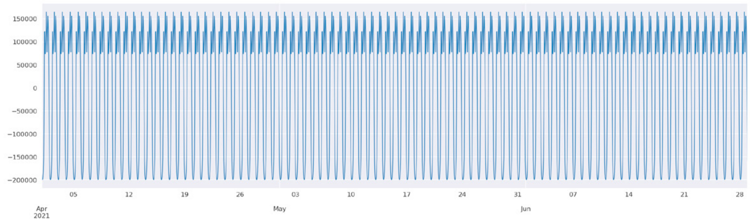


Fig. 5: Seasonality of the time series.

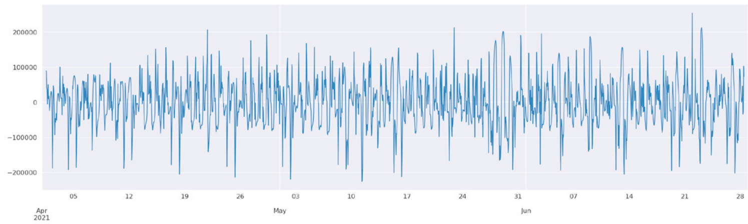


Fig. 6: Residual of the time series.

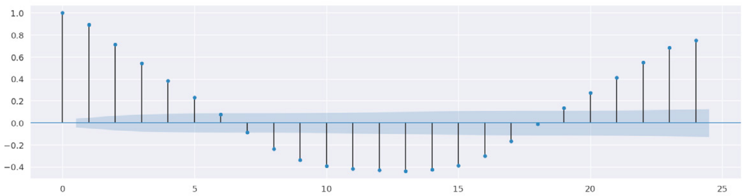


Fig. 7: Auto-correlation of the time series.

lag value equal to 0 is the highest and has an autocorrelation of 1, as it corresponds to the autocorrelation of the time series with itself and therefore has perfect autocorrelation. The next value, which corresponds to the autocorrelation of the time series with a one-hour lagged version of the time series, is approximately 0.9 which is also a high value. From here, the values go down, becoming negative at a lag value of 7. Then the values go up again until they peak again at the lag value equal to 24 and then, they go down again. This process is repeated due to the seasonal component of the series. We can conclude that the best autocorrelation value is obtained with a lag value equal to 1.

In this section, we have carried out an exhaustive analysis of the data for a better understanding of such data from a high-level perspective. This analysis has allowed us to guide the experiments towards the best configuration.

4. Methodology

We introduce a method to build a traffic flow prediction model based on LSTM neural network, and use this model to predict the future mid-term traffic flow. LSTM are a type of recurrent neural networks capable of learning long-

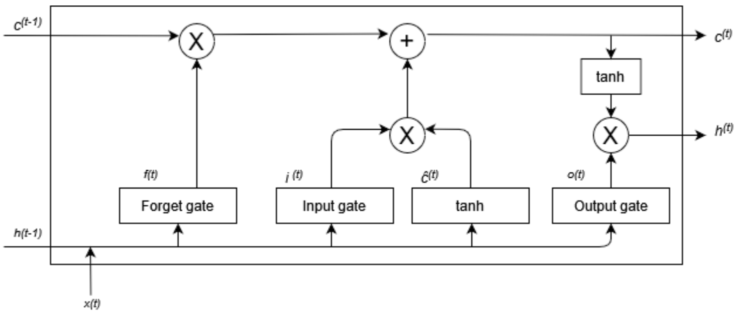


Fig. 8: Architecture of an LSTM cell.

term dependencies that was introduced by Hochreit y Schmindhuber [10] in 1997, and subsequently popularized and improved by the scientific community.

LSTM networks were designed to overcome the long-term dependency problem faced by recurrent neural networks due to the vanishing gradient problem. The LSTM network overcomes this problem by using an architecture which enforces a constant error flow through each repeating cell (see Figure 8). This property enables LSTMs to process entire sequences of data, for example time series, without treating each element of the sequence independently. The repeating module has three interacting layers, a forget gate layer, an input gate layer, and a cell state. The cell state carries the information along the module, and it can be considered as the network memory. The forget gate decides which bits of the cell state will be remembered/forgotten. The input gate determines what new information should be added to the cell state, given the previous hidden state and new input data. Finally, the hidden state corresponds to the output at the previous point in time.

Given an input sequence $x = (x_1, ..., x_t)$, the feed forward process of an LSTM is implemented with the following equations:

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \tag{3}$$

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \tag{4}$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{5}$$

$$h_t = o_t \odot W_{om}h_{t-1} + b_o \tag{6}$$

$$\hat{c}_t = \tau(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \tag{7}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \hat{c} \tag{8}$$

$$y_t = \phi(W_{yh}h_t + b_y) \tag{9}$$

where the W terms are the weight matrices, b terms are the bias vectors, σ is the sigmoid function, τ is the hyperbolic tangent function, i_t is the input gate, o_t is the output gate, f_t is the forget gate, c_t is the output of the memory cell, $\hat{c}_t$ is the cell input activation vector, and y_t is the output of the system with ϕ as the output activation function.

5. Experimental results

Two different experiments were carried out in this paper, each with an LSTM network adjusted to the experiment, in order to compare the performance of both approaches. In order to evaluate the performance, we selected a representative subset of streets in the city of Valencia: Avenida del Cid, Santos Justo y Pastor, Paseo Ciudadela, Avenida del Mediterráneo, Blasco Ibañez, Ausias March, Avenida Aragón, and Puente de Aragón.

5.1. Traffic flow on all streets

The first experiment consists of predicting the traffic flow of all the streets of Valencia at the same time according to the traffic flow in each of them. To this end, after the model fitting process, we use one layer of LSTM, two dense layers, and also we added dropout on hidden layers, to avoid overfitting. We used ReLU as the activation function in all layers, except for the output layer, where a linear activation function has been used. To facilitate the training of the network, the traffic flow values have been scaled to the range (0, 1]. In the left part of Table 2 (model A) we show the values for the MAPE and sMAPE metrics for the streets mentioned above. The results show that the model obtains good predictions for many streets, where the MAPE obtained is close to 15; however, there are others, such as Paseo Ciudadela or Avenida del Cid, where the MAPE is very high, meaning that the prediction is not acceptable. In general, values under 10 represent a highly accurate forecasting, values between 10 and 20 represent a good forecasting, and values between 20 and 50 remain acceptable; values greater than 50 are considered inaccurate [14].

To intuitively illustrate the prediction results of the model, we show the real traffic flow compared with the forecasted traffic flow on each of the above-mentioned streets on the 8th of June 2021 (see Figure 9). From these results, we can visually observe what we stated before: the forecasting results of this model remain close to the actual traffic in most of the streets, as evidenced by figs. 9a to 9e, but there are some streets, see figs. 9f to 9h, where the model is unable to obtain acceptable predictions. In fact, we find that there are more streets facing this issue. Hence, in Figure 10 we show the 10 streets with the highest MAPE, where the red line represents the average MAPE for all streets pertaining to this study. We can see that there are several streets with an error well above the average. This can be due to the fact that the traffic pattern in one street does not depend on what is occurring on the streets of the whole city, but only on those adjacent to it. We therefore propose another model where the traffic flow on a specific street is predicted as a function of the traffic flow on that street and its entrance streets (i.e, streets acting as traffic inputs).

5.2. Traffic flow on a street and its entrances

As mentioned before, this second experiment consists of predicting the traffic flow of a street according to the traffic flow on that street, and on the streets providing incoming traffic. After the model configuration fitting process, the structure of this model consists of 2 stacked LSTM layers, and 2 dense layers, using ReLU as the activation function despite the output layer having a linear activation function.

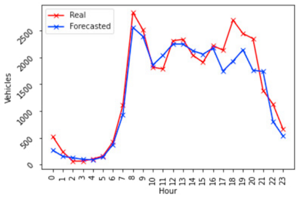
The right part of Table 2 (model B) shows the values of MAPE and sMAPE for the same streets as before. Comparing the results of this experiment (B) with the previous one (A), we notice that, despite the performance is not better in some streets, in those streets where the error was very high, this second model improves the predictions significantly. For example, in the street Blasco Ibañez, the MAPE in the second experiment is 20.26% higher than in the first experiment; on the contrary, for street Paseo Ciudadela, the MAPE in the second experiment is 71.93% lower than in the first experiment. Hence, from a global perspective, this second approach seems to return better results.

As before, Figure 11 now shows the new predictions for the 8th of June 2021, for every street being considered. We can observe that the predictions greatly resemble those of the previous experiment in most of the streets. Also, we can see that, for streets Avenida del Cid and Paseo Ciudadela, where the previous model was unable to obtain acceptable predictions, the predictions obtained with the new model are much more accurate.

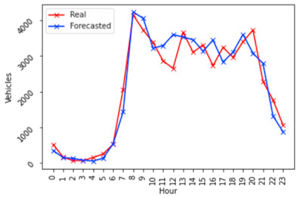
It is worth pointing out that comparing the performance of different traffic models is quite difficult due to the traffic in each location is rather different. As mentioned above, most related work focuses on motorway traffic whereas this paper deals with urban traffic. Nonetheless, from the obtained results we can affirm the predictions of our models are acceptable.

6. Conclusions and future work

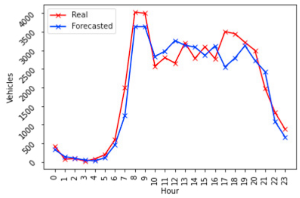
This paper addresses the problem of traffic flow prediction, which remains as a relevant challenge for the scientific community. To this end, we developed two LSTM-based models for the traffic prediction during a representative day: one that considers the full dataset to simultaneously obtain models for all streets, and a second one where each street is modelled independently, and where only those streets acting as traffic inputs are jointly considered.



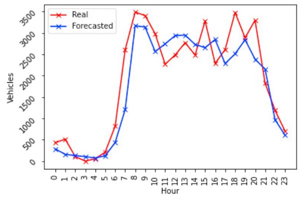
(a) Prediction for Blasco Ibañez.



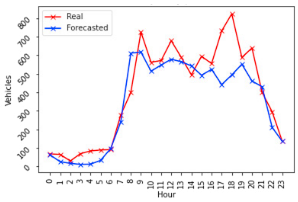
(b) Prediction for Puente de Aragón.



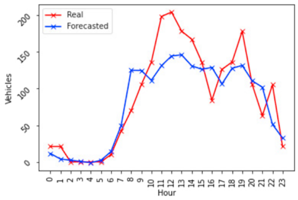
(c) Prediction for Avenida Aragón.



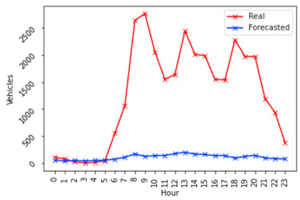
(d) Prediction for Ausias March.



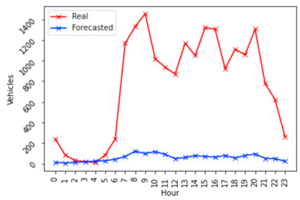
(e) Prediction for Justo y Pastor.



(f) Prediction for Avenida Mediterráneo.



(g) Prediction for Avenida del Cid.



(h) Prediction for Paseo Ciudadela.

Fig. 9: Prediction results obtained in experiment A (June 8, 2021).

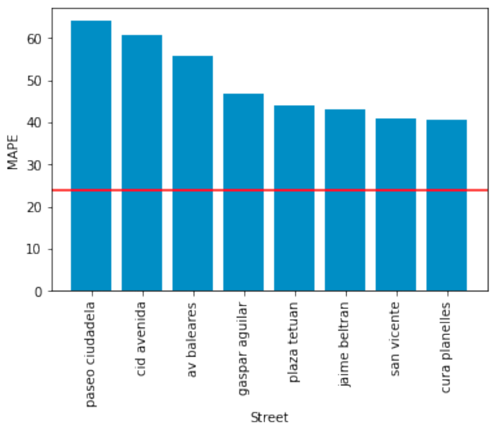
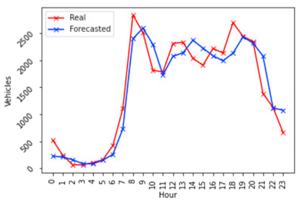
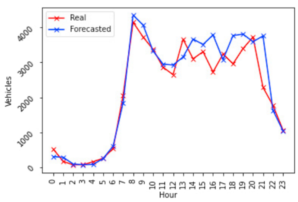


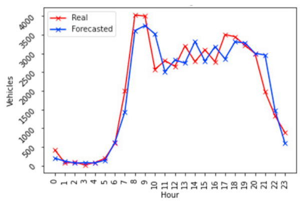
Fig. 10: Ten streets with the highest MAPE in first experiment.



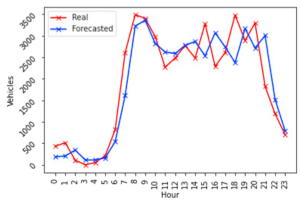
(a) Prediction for Blasco Ibañez.



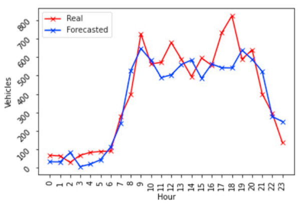
(b) Prediction for Puente de Aragón.



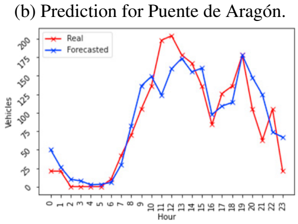
(c) Prediction for Avenida Aragón.



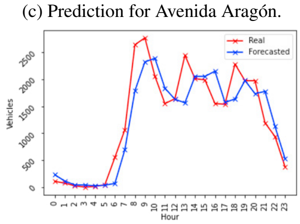
(d) Prediction for Ausias March.



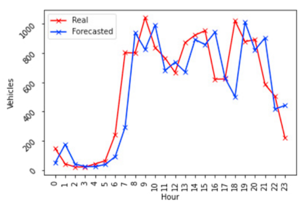
(e) Prediction for Justo y Pastor.



(f) Prediction for Avenida Mediterráneo.



(g) Prediction for Avenida del Cid.



(h) Prediction for Paseo Ciudadela.

Fig. 11: Prediction results obtained in experiment B (June 8, 2021).

Table 2: Results for the different target metrics when comparing the different streets being analysed for both experiments A and B.

	Model A		Model B	
Street	**MAPE**	**sMAPE**	**MAPE**	**sMAPE**
Blasco Ibañez	13.4001	14.3171	16.8060	15.9828
Puente de Aragón	14.1832	15.2450	13.4146	13.1506
Avenida Aragón	14.6521	15.6515	18.9181	17.3668
Ausias March	14.9857	15.5007	15.9026	15.7278
Santos Justo y Pastor	16.7251	15.8644	16.5754	15.9595
Avenida del Mediterráneo	32.3700	39.8390	25.7874	24.8818
Avenida del Cid	58.8434	90.2349	19.8860	20.4179
Paseo Ciudadela	64.4744	106.1492	18.1004	18.3678

Concerning results for the first model, we notice that, although many streets obtained an acceptable prediction, there were others that did not, achieving high MAPE results. When relying instead on the second model, we find that, despite not all streets improve their accuracy, the error improvement is very significant for those streets presenting excessively high MAPE values in the first model. Thus, this second model seems more adequate, although it involves more time and analysis to determine accurately which streets act as traffic inputs to any specific street.

As for future work, we will explore other prediction models, such as ARIMA models, in order to compare the results obtained, investigating more sophisticated feature selection methods.

Acknowledgements

This work is derived from R&D project RTI2018-096384-B-I00, funded by MCIN/AEI/10.13039/501100011033, and "ERDF A way of making Europe".

References

[1] Ajuntament de València, . Intensidad de los puntos de medida de tráfico (espiras electromagnéticas). URL: https://www.valencia.es/dadesobertes/es/dataset/?id=intensidad-de-los-puntos-de-medida-de-trafico-espiras-electromagneticas. [Online; accessed 01-April-2021].

[2] Calafate, C.T., Soler, D., Cano, J.C., Manzoni, P., 2015. Traffic management as a service: The traffic flow pattern classification problem. Mathematical Problems in Engineering 2015.

[3] Crivellari, A., Beinat, E., 2020. Forecasting Spatially-Distributed Urban Traffic Volumes via Multi-Target LSTM-Based Neural Network Regressor. Mathematics 8.

[4] Dai, G., Ma, C., Xu, X., 2019. Short-term traffic flow prediction method for urban road sections based on space–time analysis and gru. IEEE Access 7, 143025–143035. doi:10.1109/ACCESS.2019.2941280.

[5] de Myttenaere, A., Golden, B., Le Grand, B., Rossi, F., 2016. Mean absolute percentage error for regression models. Neurocomputing 192, 38–48. URL: https://www.sciencedirect.com/science/article/pii/S0925231216003325, doi:https://doi.org/10.1016/j.neucom.2015.12.114. advances in artificial neural networks, machine learning and computational intelligence.

[6] Dickey, D., Fuller, W., 1979. Distribution of the estimators for autoregressive time series with a unit root. JASA. Journal of the American Statistical Association 74. doi:10.2307/2286348.

[7] Dong, H., Jia, L., Sun, X., Li, C., Qin, Y., 2009. Road traffic flow prediction with a time-oriented arima model, in: 2009 Fifth International Joint Conference on INC, IMS and IDC, pp. 1649–1652. doi:10.1109/NCM.2009.224.

[8] Du, S., Li, T., Gong, X., Horng, S.J., 2019. A hybrid method for traffic flow forecasting using multimodal deep learning. arXiv:1803.02099.

[9] Hochreiter, S., Schmidhuber, J., 1997a. Long short-term memory. Neural Comput. 9, 1735–1780. URL: https://doi.org/10.1162/neco.1997.9.8.1735, doi:10.1162/neco.1997.9.8.1735.

[10] Hochreiter, S., Schmidhuber, J., 1997b. Long short-term memory. Neural Computation 9, 1735–1780.

[11] Hyndman, R., Athanasopoulos, G., 2018. Forecasting: Principles and Practice. 2nd ed., OTexts, Australia.

[12] Kreinovich, V., Nguyen, H.T., Ouncharoen, R., 2014. How to estimate forecasting quality: A system- motivated derivation of symmetric mean absolute percentage error (smape) and other similar characteristics .

[13] Kumar, S., Vanajakshi, L., 2015. Short-term traffic flow prediction using seasonal arima model with limited input data. European Transport Research Review 7. doi:10.1007/s12544-015-0170-8.

[14] Lewis, C.D., 1982. Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. Butterworth-Heinemann.

[15] Li, K., Zhai, C., Xu, J., 2017. Short-term traffic flow prediction using a methodology based on arima and rbf-ann, in: 2017 Chinese Automation Congress (CAC), pp. 2804–2807. doi:10.1109/CAC.2017.8243253.
[16] Luo, X., Li, D., Yang, Y., Zhang, S., 2019. Spatiotemporal traffic flow prediction with KNN and LSTM. Journal of Advanced Transportation .
[17] Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y., 2015. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transportation Research Part C: Emerging Technologies 54, 187–197. URL: https://www.sciencedirect.com/science/article/pii/S0968090X15000935, doi:https://doi.org/10.1016/j.trc.2015.03.014.
[18] Makino, H., Tamada, K., Sakai, K., Kamijo, S., 2018. Solutions for urban traffic issues by its technologies. IATSS Research 42, 49–60. URL: https://www.sciencedirect.com/science/article/pii/S0386111218300293, doi:https://doi.org/10.1016/j.iatssr.2018.05.003.
[19] Polson, N.G., Sokolov, V.O., 2017. Deep learning for short-term traffic flow prediction. Transportation Research Part C: Emerging Technologies 79, 1–17. URL: https://www.sciencedirect.com/science/article/pii/S0968090X17300633.
[20] Qu, L., Li, W., Li, W., Ma, D., Wang, Y., 2019. Daily long-term traffic flow forecasting based on a deep neural network. Expert Systems with Applications 121, 304–312. URL: https://www.sciencedirect.com/science/article/pii/S0957417418308017, doi:https://doi.org/10.1016/j.eswa.2018.12.031.
[21] Seabold, S., Perktold, J., 2010. statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference. [Online; accessed 01-April-2021].
[22] Sun, S., Wu, H., Xiang, L., 2020. City-wide traffic flow forecasting using a deep convolutional neural network. Sensors 20, 421. doi:10.3390/s20020421.
[23] Tunnicliffe Wilson, G., 2016. Time series analysis: Forecasting and control,5th edition, by george e. p. box, gwilym m. jenkins, gregory c. reinsel and greta m. ljung, 2015. published by john wiley and sons inc., hoboken, new jersey, pp. 712. isbn: 978-1-118-67502-1. Journal of Time Series Analysis 37, n/a–n/a. doi:10.1111/jtsa.12194.
[24] Xu, D., Shi, Y., 2017. A combined model of random forest and multilayer perceptron to forecast expressway traffic flow, in: 2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC), pp. 448–451. doi:10.1109/ICEIEC.2017.8076602.
[25] Yan, B., Wang, G., Yu, J., Jin, X., Zhang, H., 2021. Spatial-temporal chebyshev graph neural network for traffic flow prediction in iot-based its. IEEE Internet of Things Journal , 1–1doi:10.1109/JIOT.2021.3105446.
[26] Zambrano-Martinez, J., Calafate, C., Soler, D., Lemus, L., Cano, J.C., Manzoni, P., Gayraud, T., 2019. A centralized route-management solution for autonomous vehicles in urban areas. Electronics 8, 722. doi:10.3390/electronics8070722.