Forecasting Spatially-Distributed Urban Traffic Volumes via Multi-Target LSTM-Based Neural Network Regressor

Alessandro Crivellari * and Euro Beinat

Department of Geoinformatics—Z_GIS, University of Salzburg, 5020 Salzburg, Austria; euro.beinat@sbg.ac.at
* Correspondence: alessandro.crivellari@sbg.ac.at

check for
updates

Abstract: Monitoring the distribution of vehicles across the city is of great importance for urban traffic control. In particular, information on the number of vehicles entering and leaving a city, or moving between urban areas, gives a valuable estimate on potential bottlenecks and congestions. The possibility of predicting such flows in advance is even more beneficial, allowing for timely traffic management strategies and targeted congestion warnings. Our work is inserted in the context of short-term forecasting, aiming to predict rapid changes and sudden variations in the traffic volume, beyond the general trend. Moreover, it concurrently targets multiple locations in the city, providing an instant prediction outcome comprising the future distribution of vehicles across several urban locations. Specifically, we propose a multi-target deep learning regressor for simultaneous predictions of traffic volumes, in multiple entry and exit points among city neighborhoods. The experiment focuses on an hourly forecasting of the amount of vehicles accessing and moving between New York City neighborhoods through the Metropolitan Transportation Authority (MTA) bridges and tunnels. By leveraging a single training process for all location points, and an instant one-step volume inference for every location at each time update, our sequential modeling approach is able to grasp rapid variations in the time series and process the collective information of all entry and exit points, whose distinct predicted values are outputted at once. The multi-target model, based on long short-term memory (LSTM) recurrent neural network layers, was tested on a real-world dataset, achieving an average prediction error of 7% and demonstrating its feasibility for short-term spatially-distributed urban traffic forecasting.

Keywords: deep learning; LSTM; neural networks; traffic forecasting; urban traffic distribution

1. Introduction

The growing data availability, and widespread monitoring systems, have boosted the focus on urban traffic analysis and vehicle movement observation. The study of spatial-temporal transit volume distribution in urban environments is a central aspect for guiding decision-making processes according to the current traffic situation, and can be approached in several different modalities and in a wide variety of contexts [1,2]. Typical applications include traffic congestion warnings [3,4], car accident risk assessments [5,6], and pollution measurement estimations [7,8].

In this context, the focus on predictive analytics is prominent, and a number of works dealing with traffic forecasting can be enumerated [9–14]. Vehicle-related predictions have been widely approached under different concepts, definitions, and assumptions, becoming a very active research area in the big picture of location-based services and smart city features. A general forecasting direction addresses the inference of the amount of active vehicles located in a specific area at some point in the future, based on the present volume, past trend variations, and additional side information, if available. Depending on

the background, this may embody into very different problems and approaches, comprising the identification of congestion locations [15–17], traffic estimations of road tracts [18–20], and individual vehicle movement predictions [21–23].

Our paper tackles a particular case, focusing on the problem of predicting the number of vehicles moving across several urban regions. In contrast with many works analyzing traffic flow in a single selected geographic point (e.g., [24–27]), we open to concurrent predictions on multiple flows, considering different reference points (and directions) connecting city areas, therefore providing an estimated future distribution of vehicles across different locations of interest over the territory. Moreover, we concentrate upon short-term predictions, namely forecasting the transit volume according to an hourly time step. Since traffic volume is subject to rapid changes in time and space, the main challenge lies in grasping abrupt variations beyond the general trend.

Predicting short-term distribution of vehicles across the city is indeed of great importance for urban traffic control, particularly in the context of smart city services for nearly real-time decision-making outcomes and intelligent information retrieval [28,29]. Estimations on the number of vehicles entering and leaving a city, or moving between urban areas, gives a valuable indication on potential localized bottlenecks and congestions. Inserted in the big wave of mobility mining [30,31] heading to providing valuable knowledge on human mobility and urban travel behaviors, the possibility of predicting urban vehicle flows is highly beneficial for timely traffic management strategies, targeted congestion warnings, and any required general forestalling action.

Our case study targets urban toll plazas in correspondence with the Metropolitan Transportation Authority (MTA) tunnels and bridges in New York City, which, serving over 900,000 vehicles per day, carry more traffic than any bridge and tunnel complex in the USA [32]. Focusing on large urban contexts in the presence of high transit volumes, we propose a geographically-oriented short-term forecasting methodology to predict vehicles' distribution over the territory, leveraging sequential modeling and deep neural networks.

The underlying assumption is built on successfully mining temporal patterns of traffic volumes, embodying the hypothesis that the current number of registered moving vehicles exhibits some dependences on past measured quantities. The characteristic features in the series are therefore assumed to carry essential information for anticipating future vehicles' distribution. While typical time series forecasting focus on learning the general trend via purely statistical methods [33–36], the natural characterization of our short-term predictive analysis is defined by rapid and steep variations that need to be properly detected in order to provide accurate estimations. Smoothing the phenomena with regressive statistical models tends to lead to ineffective predictions, when in the presence of intense oscillations and shifts to boundary conditions [37–39]. Moreover, since multiple traffic flows are taken into account, multiple predictions are to be performed, whereby each reference point (and direction) on the territory is associated to different traffic volumes and different sequential patterns over time. Our goal is to collectively generate accurate predictions based upon a single model (and a single training process) comprising the totality of reference points, therefore following an efficient strategy and avoiding singularly fitting each selected traffic direction with its corresponding distinctive model (as it is required for statistical approaches). Furthermore, this way leverages a combined global view of the city trend, not only analyzing patterns of single geographic directions, but also detecting inter-location relations in the traffic variation.

In an attempt to satisfy the above-mentioned requirements in a straightforward and effective modality, we propose a multi-target deep learning regressor, trained on location-specific time series, for simultaneous predictions of traffic volumes in multiple entry and exit points among city neighborhoods. By leveraging a single training process for all location points and an instant one-step volume inference for every location at each time update, our sequential modeling approach is able to grasp rapid variations in the time series and process the collective information of all reference points, whose distinct predicted values are outputted at once. Firstly, sequences of numbers of vehicles are considered, whereby each selected geographic point on the territory is represented by a series of

values unfolding in a fixed time step. Then, the sequences are stacked together as a combined input to a long short-term memory (LSTM) recurrent network-based regressor, which is jointly trained on the block of time series to learn the underlying patterns of urban traffic flows. Finally, a multi-target dense layer provides a number of output values equal to the number of reference locations (and directions). The suggested approach is purely data-driven, capturing variability patterns directly from traffic volume sequences, without any manual feature extraction.

The main contribution of the paper therefore stands out as an extensive exploration of multi-target artificial neural networks for promisingly handle multiple traffic series belonging to spatially-related reference points, arranging a collective processing of multiple input sequences into corresponding multiple forecasted output values. Falling in the context of deep learning applications on human mobility analysis, the underlying purpose is to extend a purely data-driven perspective on vehicle number forecasting for an improved quality of urban management and organization.

The structure of the paper is arranged as follows: Section 2 describes the methodological workflow supporting the prediction task, specifically illustrating the data pre-processing (Section 2.1), the model architecture (Section 2.2), and the model training (Section 2.3); Section 3 concerns the experimental arrangement, presenting the dataset (Section 3.1), the general set-up (Section 3.2), and the achieved results (Section 3.3); Section 4 finally reports the discussion and conclusion.

2. Methodology

The proposed forecasting method aims to collectively learn sequential patterns of spatial-temporal traffic volume variation, specifically predicting the future number of vehicles crossing each of many selected reference locations. Given a stacked input block of multiple time series (each referring to a specific reference location) sampled at a fixed coordinated time step, the solution of our model consists of inferring the vehicles' distribution over the territory during the next time step in the future. The current section reports the overall methodological description, from collective time series definition to sequential multi-target deep learning modeling.

2.1. Time Series Data Structure

The methodology follows a sequential modeling approach, leveraging a collection of historical time series, each referring to a specific reference location over the territory.

A time series is defined as a sequence of real values unfolding into a predefined time step. If data are not originally sequentially structured (e.g., still in the raw form of individual vehicle recordings), they require to be reshaped into a series of vehicle counts by transforming the continuity of time into fixed time steps and counting the total number of vehicles crossing the reference location within each time span t. Each location p is therefore described as a series of chronologically ordered vehicle counts $S_p = \left\{ \#vehicles(t)_p \,\middle|\, t = 1, 2, 3, \ldots \right\}$.

The sequences of all locations are then stacked together, synchronized at the same time step. The input to the prediction model becomes a time series of multiple dimensions, where the dimensionality is equal to the number of reference locations. In particular, given a number of P locations, the portion of the sequence identified by a specific time span t' is made of a vector of P values reporting the number of vehicles within the time span t' in each location p, namely $S(t') = [\#vehicles(t')_p \,|\, p = 1, 2, \ldots, P]$. Therefore, the same longitudinal position along the stacked sequence identifies locations' count values that are referring to the same time span, as shown in Figure 1. The length of the time step is case specific, its choice depends on the data source and the prediction problem; our focus specifically relies on short-term forecasting of rapid sequential variations. Analogously, different datasets leverage different numbers of reference locations within the same general phenomenon.

...	7–8 am May 5th	8–9 am May 5th	9–10 am May 5th	10–11 am May 5th	...
LOC_ID = 1	...	1000	1238	1531	2033	...
LOC_ID = 2	...	979	1512	1717	2021	...
LOC_ID = 3	...	1199	1604	1824	2008	...
LOC_ID = 4	...	1114	1566	1700	1821	...
...	...	...	...	...	...	...

Figure 1. Input block of longitudinally stacked locations' time series. Each value refers to the number of recorded vehicles crossing a certain location within a defined time span.

In any case, the data configuration should consist of a set of reference points over the territory, each of them described by a specific time series made of sequential counts of recorded vehicles in consecutive time spans. The final input to the model is then arranged as a block of these series, stacked together on the longitudinal dimension, unfolding in the same identical time steps.

2.2. Multi-Target Deep Learning Model for Distributed Traffic Prediction

The collection of stacked location-specific time series is defined as the input data to the deep neural network model, whose processing workflow can be depicted according to its three structural components: a multi-dimensional input layer, a block of LSTM recurrent network layers, and a multi-target output layer. A high-level visual overview of the modeling conceptual architecture is reported in Figure 2. The following subsections singularly focus on each of the three parts composing the model.
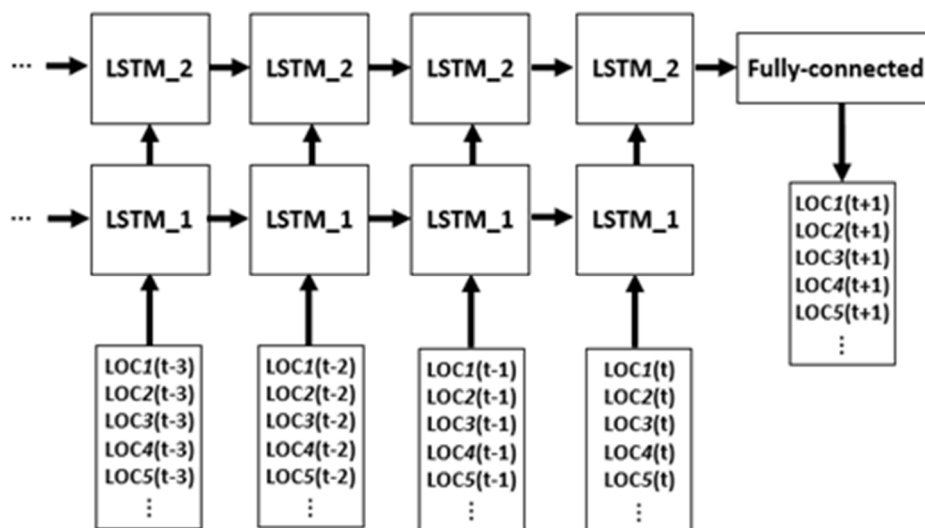

Figure 2. Visual overview of the modeling conceptual architecture.

2.2.1. Input Layer

To collectively handle multiple time series with a single model, while still retaining individual characteristics, we make use of a multi-dimensional input layer, whereby each input neuron refers to a specific reference location. Sequences are therefore processed simultaneously but initially addressed through separate channels. A number of P reference vehicle flows identify P time series, which are combined into a sequence of P dimensions, input for a layer of P neurons. At each time step, the model is fed with a vector of P values, one for each reference location; over new consecutive steps, the values

referring to a same location always occupy the same position in the array. This means to automatically get the collection of traffic volumes for all the areas at once, but heading to separate input neurons, as reported in Figure 3. In general, two input processing perspectives are present: the sequential time-dependent evolution of the series, and the horizontal traffic distribution over multiple location points. During the training process, the model receives consecutive chronologically-ordered vectors of size *P*, and accordingly defines the updates to the parameters of the recurrent and output layers.
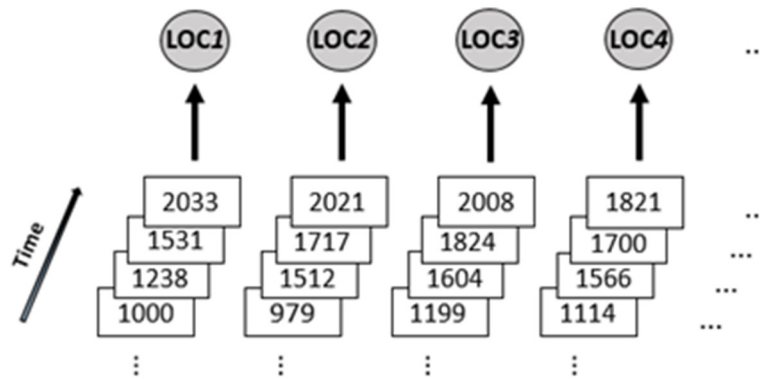

Figure 3. Input layer representation: each neuron receives a series referring to a specific reference location.

2.2.2. LSTM Block

The automatic pattern learning activity is conducted by a recurrent network block, leveraging LSTM layers. LSTM [40] is a complex recurrent module, composed of four distinctive neural networks interacting between each other. Due to the characteristic capability of handling sequential data, it has been utilized in a variety of applications dealing with motion trajectory prediction [41–43] and mobility-related time series forecasting [44–46]. A few adoptions to studies on traffic analysis are also present, but has been mainly modeled for singularly predicting flows on a selected target location [47–50].

The functioning of an LSTM cell consists of processing sequential data one element at a time, receiving, at each step, two input sources: the current input data element and the processed network output of the previous time step. The encoded information, repeatedly modified by the four internal neural network structures, is defined in the form of a cell state vector. In the last processing step preceding prediction, this evolves into an output vector carrying the overall compressed characterization of the series, subsequently used for disclosing the explicit prediction outcome.

Examining the process from our perspective, the recurring input to the LSTM is represented, at each time step, by the current distributed vector representation along the multi-dimensional sequence of traffic volume, concatenated with the output vector of the LSTM module at the previous step. As a result of the network structure, the cell state contains the pattern information of all reference locations together, encoded in a single vector representation. Therefore, even if different locations initially targeted different entry channels, their single peculiarities are then mixed up in the LSTM state vector and, consequently, in the final output characterization. If the LSTM block contains multiple LSTM layers, the first layer is fed with the input data, the second layer is fed with the output of the first layer, and so on; the final output vector refers to the last step of the last layer. Figure 4 visually reports the last two steps of a sequence of vehicle flows, processed by a recurrent block of two LSTM layers.

Equations (1)–(6) describe a repeating module of LSTM in further detail. Considering a sequence *X* and an input vector x_t at time *t*, the forget gate (1) is intended to selectively erase information; the input gate (2) is oriented to indicate which values are required to be updated; the tanh network (3) aims to provide novel information in the form a vector of new candidate values; the cell state (4) is therefore updated following a filtering action through the forget gate and adding up with the combined outcome of the input gate and the tanh network; the output gate (5) is then arranged to select which

parts of the cell state to output; and finally, the LSTM outcome (6) derives from the multiplication of the output gate with the tanh of the updated cell state.
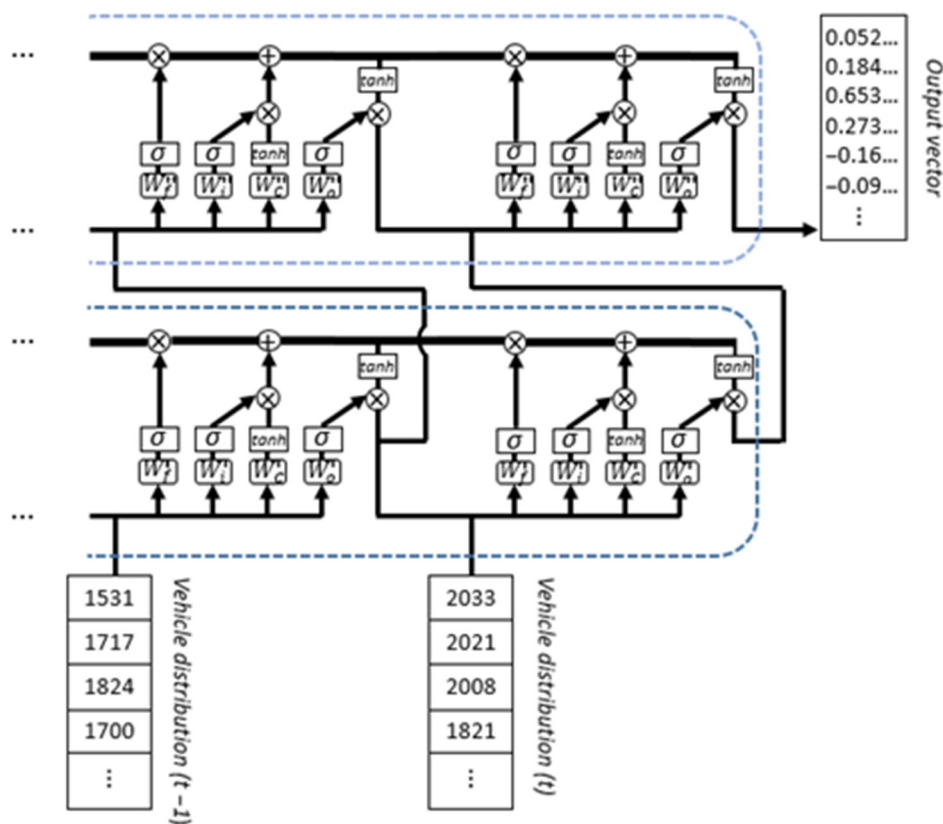

Figure 4. Visual schema reporting the last two steps of a sequence of vehicle flows, processed by a recurrent block of two LSTM layers.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\widetilde{C}_t = \tan h\ (W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \tag{4}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * \tan h\ (C_t) \tag{6}$$

2.2.3. Output Layer

The output layer is intended to sustain the function of translating the LSTM final vector representation into the explicit traffic volume distribution across the different reference locations. The task relies on providing a number of output values equal to the number of locations. In particular, each output neuron is required to predict the future number of vehicles crossing a single specific reference point, therefore depicting the forecasting process as a multi-target regression analysis.

The layer consists of a multi-output fully-connected neural network on top of the LSTM block. Its purpose is to transform the compressed encoded information, contained in a single dense vector, into multiple simultaneous predictions conforming to the distributed geographic perspective, as reported in Figure 5.
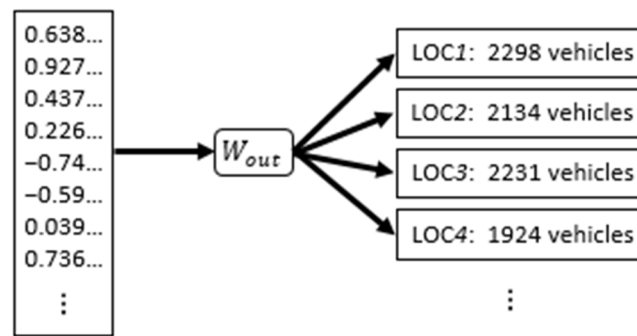
Figure 5. Output layer information reshaping process from the LSTM final vector representation to the explicit traffic volume distribution across the different reference locations.

Equation (7) formally describes the output layer, where P indicates the total number of reference locations and h_{fin} refers to the final vector representation of the LSTM block.

$$\#orders_j = W_{(out)j}h_{fin} + b_{(out)j} \qquad \forall j \in (0, P] \tag{7}$$

2.3. Model Training

The procedure of identifying, at each time step, the input data elements and the corresponding target value for training the model is accomplished by scanning the multi-dimensional sequence of vehicle distribution with a sliding window, which, consecutively moving forward by one step until the end of the sequence, produces multiple input segments of a fixed length. The segment length defines the extent of the continuous sequential dependencies grasped through pattern mining, and is strongly influenced by the data characteristics. Its choice may vary according to different experimental setups and time resolutions, even within the same domain analysis.

The sequential learning of traffic evolution is driven by such collection of segments, characterizing a training process based on input sequences paired with their corresponding future target values. The parameter optimization focuses on minimizing the mean squared error between the predicted amounts of vehicles and the actual recorded numbers. The weights within every layer of the network are therefore updated through a backpropagation process and mini-batch stochastic training, gradually adjusted towards the direction of the gradient to provide more and more accurate prediction outputs, with respect to the real target volumes.

At testing time, the learned parameter configuration, based on historical pattern mining, allows for predicting the traffic distribution in the future. The instant forecasting of the distributed number of vehicles across the territory is therefore grounded on the model's detection, in the time span preceding prediction, of sequential patterns that resemble past traffic variability tendencies.

3. Experiment

We hereby provide a description of the experimental settings and the evaluation dataset, followed by an overview of the results and findings on the forecasted traffic volume distribution within a real-world environment. The deep learning model was implemented and trained using TensorFlow, and subsequently evaluated in relation to multiple baseline performances for a proper measure of its feasibility.

3.1. Dataset

The experimental analysis on future traffic distributions across urban regions leverages a real-world dataset of hourly counts of vehicles crossing several entry and exit reference points between city neighborhoods. Specifically, the data refer to the urban toll plazas in correspondence of the nine MTA

tunnels and bridges in New York City [32], depicting an exemplifying case study of large urban context in the presence of high transit volumes. In particular, we analyzed the time span from January 2012 to December 2015, targeting a total of 19 different vehicle flow directions, as represented in Figure 6. A single observation includes the plaza identifier, the transit direction, the date and hour, and the number of observed vehicles.

Figure 6. Output Geographic representation of the 19 traffic flows under study, in correspondence to the nine MTA bridges and tunnels.

The model input consists of a 19-dimensional sequence, describing the traffic volume variation over time for each of the reference locations. The sequence unfolds in hourly time steps, whereby a single step reports the total number of vehicles in the last hour, distributed over multiple geographic points. This multi-location perspective therefore identifies each new input element as a vector of 19 values, representation of the current hourly distribution of traffic volume across city neighborhoods' connecting routes.

3.2. Experimental Settings

The neural network model architecture was designed with a block made of two LSTM layers, characterized by a hidden size of 128 neurons each. The sliding window progressively collecting the scaled input values was set to include the 12 h before prediction, and the forecasting output format was defined as a vector of 19 dimensions, reporting the predicted number of vehicles crossing each of the reference locations in the following hour. The mini-batch stochastic training process featured a parameter configuration update based on the mean squared error loss function and Adam optimizer [51]. The outcome was finally assessed on a data portion that was not involved in the training phase; specifically, three consecutive years were chosen as a training set (2012–2014), and the successive year (2015) was selected as testing data for evaluating the model.

Global measures, to quantify the overall performance, were based on the prevalent metrics of mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE). Separate results on the single reference locations were also taken into account, to assess some possible geographic inhomogeneity in the forecasting outcomes.

For a meaningful reasoning on the quality of our methodology, the performance evaluation was compared to baseline approaches comprising cyclical mining of historical recurrent trends and regressive statistical modeling, typical methods for time series analysis. The next subsection reports the experimental results and examines the findings deriving from the predicted vehicle distribution across the multiple selected reference directions.

3.3. Results

The global scores in terms of MSE, RMSE and MAE are reported in Table 1. The multi-target LSTM model is compared with baselines built on the concept of cyclical repeating patterns or steady behaviors in time, aiming to quantify its performance advantage over rule-based approaches highlighting particular specific characteristics of the sequential phenomena. Specifically, four baselines were selected, each focusing on a particular aspect of traffic flow dynamics. A first approach ("Last hour") models steady behaviors, assuming little changes on consecutive hours, therefore predicting the number of vehicles as the same of the number registered in the previous hour. A second baseline ("Daily cycle") depicts daily cyclical recurrences, assuming the prediction outcome as equal to the number of vehicles observed at the same hour of the previous day. A third baseline ("Weekly cycle") focuses instead on weekly cycles, expecting the future volume to be equal to the number of vehicles of the previous week, in the same day of the week and at the same time, hypothesizing a general weekly repetitiveness of the trends. The fourth baseline ("Weekly cycle (4 weeks avg)") is a generalization of the third one, applying the same principle but, instead of only focusing on the previous week, it takes into account the whole previous month, averaging the corresponding weekly values (e.g., the traffic volume on the next Monday at 6pm is forecasted as the average of the past four Mondays at 6 pm) and, therefore, reducing fluctuation noises along the sequence. Finally, we set two further baselines in the form of regressive statistical methods for time series analysis, fitting each series with the commonly utilized ARIMA and FBprophet models, widely-adopted performance thresholds for series-related comparative analysis, due to their widespread and efficient automatic implementations of parameter tuning.

The results report the best baseline to be the monthly-averaged weekly cycle-based approach, enhancing a weekly recurrent characteristic pattern. On the other hand, our model is shown to substantially outperform the baselines, pointing out a remarkable forecasting capability, disclosing a MSE of 26K versus the best baselines' value of 72K.

Referring only to a global mean score may not be fully revealing because of the diverse quantities of traffic volume in correspondence of different reference locations, which consequently leads to various distinctive degrees of prediction error. Figure 7, therefore, provides a complete overview of the MAE distribution over the multitude of reference points. Specifically, the distributed outcome of the LSTM model is compared to the baselines with respect to each of the 19 traffic flows under study, revealing a superior performance in all of them, regardless of any difference in the local traffic amounts.

The best baseline is confirmed to be the same for the great majority of reference directions, whereas the sequential statistical processes are not able to catch the rapid variations in the series, ending up in a widespread poor fitting performance (combined with the inefficiency of requiring a separate fit for each single flow, not handling any multi-output option).

Table 1. Overall comparison results of the LSTM model and the baseline approaches in terms of MSE, RMSE, and MAE.

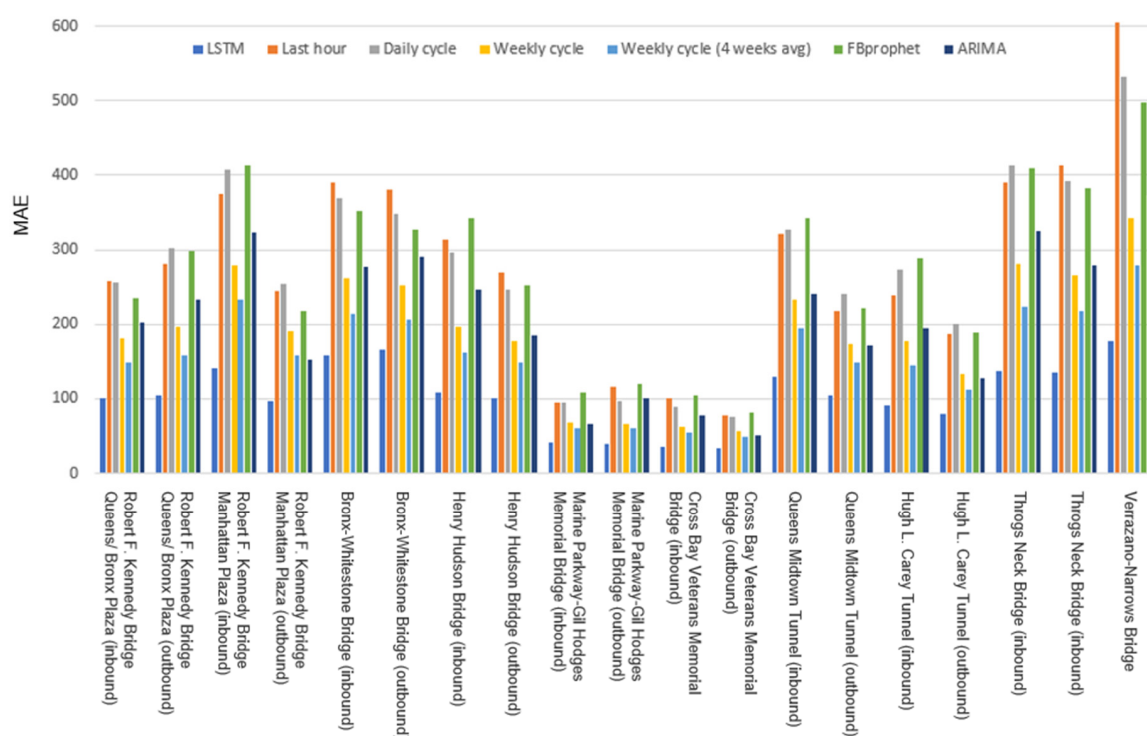|	LSTM	Last Hour	Daily Cycle	Weekly Cycle	Weekly Cycle (4 Weeks Avg)	FBprophet	ARIMA
MSE	26,793	176,388	233,077	121,596	72,799	167,293	93,791
RMSE	163.7	420.0	482.8	348.7	269.8	409.0	306.3
MAE	104.2	277.6	274.5	189.0	156.5	272.9	206.2

Figure 7. MAE distribution over the 19 reference traffic flows.

Figure 8 enriches the information, after normalizing by the actual local number of recorded vehicles, presenting the results in terms of percentage of predicted volume error, with respect to the real volume. The LSTM outcomes are reported to range between 4% and 9% of prediction error.

To have a deeper insight into single reference traffic flows, a further analysis may take into account the time variable. It is worth noticing, however, that different flows imply different characteristic behaviors, and different traffic volume variations over time. To acknowledge the presence of various series' profiles, we report the average daily traffic pattern of two exemplifying reference locations in Figure 9. The volume variation is depicted together with the overall predictability trend over time, disclosing the forecasting variability over the 24 h. Comparing our model with the best baseline reveals distinctive performance outlines in terms of average hourly prediction error; the maximum improvement is particularly registered during the late afternoon hours.
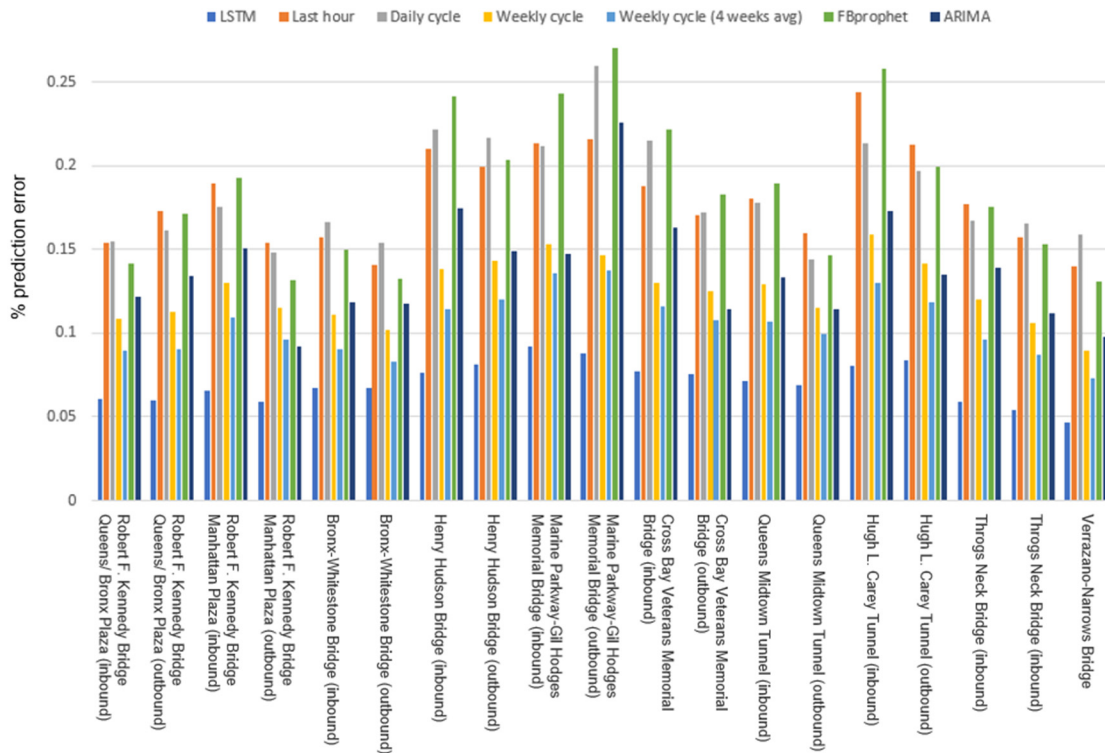
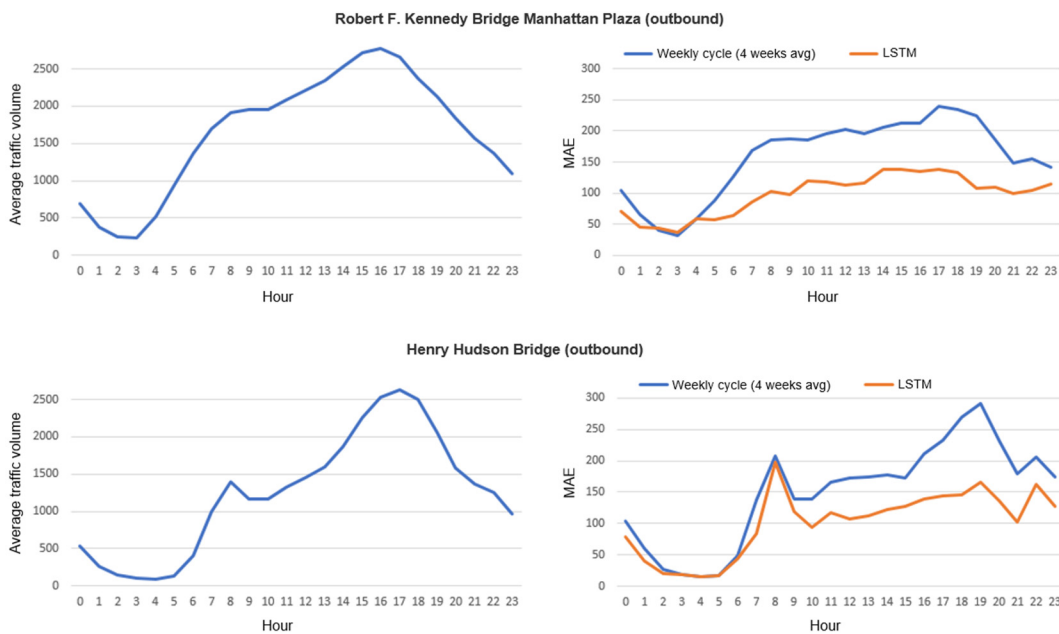**Figure 8.** Percentage of predicted volume error with respect to the real volume.

Figure 9. Average traffic volume variation and predictability trend, over the 24 h, of two exemplifying reference locations.

4. Discussion and Conclusions

The paper proposed a deep neural network approach for predicting traffic volumes, distributed over multiple reference locations across the city. We particularly focused on movements between urban regions, targeting crowded crossing points between contiguous neighborhoods. The task assumed the trait of a short-term forecasting process in a distributed geographic context, grounding its implementation in a sequential modeling perspective. The process relied on a recurrent LSTM-based

network model, set up as a multi-target regression problem, involving location-specific sequences of vehicle counts unfolding in a fixed time step. The collection of all reference locations' input series were concatenated on the time axis, defining each input step as a multi-dimensional vector containing the current distributed amounts of vehicles across the neighborhood connections under study. The prediction outcome, following the same principle, was the product of a multi-target dense layer outputting the time-stamped vector of the expected vehicle distribution across the totality of reference locations. Our case study involved a real-world dataset in New York City, assessing the feasibility of predicting the number of vehicles moving through each of the MTA bridges and tunnels. The scope was therefore enclosed in a short-term distributed traffic volume estimation process, in the context of real-time decision-making strategies and urban management services.

The presented methodology was proposed as an effective solution for capturing rapid variations and abrupt changes beyond the general trends, essential aspect for a proper short-term traffic prediction. Compared to a variety of baselines embodying particular characteristics of the series (steady behaviors and repetitive cycles), our approach reported a superior performance, effect of a significant capability of modeling sequential characteristics of traffic evolution over time in correspondence of multiple distinctive locations. Besides the baselines' evidence of a prominent weekly-based recurrent cycle, the substantially better results of the deep learning model presumed a presence of more sophisticated inherent patterns in the series, which, when properly mined, revealed information on future traffic changes as a result of more complex properties than a pure cyclical repetitiveness. Contrary to regressive statistical methods, which fit the general trend but were not able to catch sudden variations, our model better identified quick time evolutions of traffic distribution for anticipating short-term rapid leaps in the future.

The experimental results reported outperforming measures over all comparison approaches with respect to each of the 19 selected reference traffic flows in the city, revealing a widespread effectiveness on the whole territory, regardless of geographic locations and local traffic volumes. This indicates a general performance stability, demonstrating the feasibility of collectively modeling a variety of different location-specific characteristic patterns and different traffic profiles over time.

From a practical utilization perspective, it is worth noticing some implicit advantages and facilitations, especially regarding the deployment phase. The proposed approach was designed for processing the collection of reference locations together within one single global training procedure, in contrast to traditional statistical models requiring a separate fit for each different flow. The prediction outcome relies on an instant automatic one-step inference of distributed numbers of vehicles at each time update, releasing the output values of all locations together in a single forecasting block. Furthermore, the neural network does not require a refit for each prediction update, but only a periodical retraining is suggested, depending on the performance degrade caused by evolving behaviors in time.

This work is inserted in the wide field of smart city research [52], on the side of predictive analytics for advanced decision-making and user-oriented services, leveraging deep learning-based artificial intelligence advancements within a purely data-driven perspective. The broad objective refers to the traffic monitoring and control in an urban environment, implying a variety of applications related to future traffic volume estimations. A straightforward direction leads to predictions of bottlenecks and congestions, enabling real-time urban strategies by local organizations, including warning information messages to drivers about the level of expected traffic distribution in the next hour. Another subject of interest focuses on exploratory studies of inter-neighborhood motion characteristics, portraying a future depiction of human mobility across the city by forecasting entry and exit flows within urban areas; local crowding estimates can be derived from the number of vehicles entering and leaving the city or moving between specific neighborhoods, leading to designing appropriate preventive strategies. Reliable predictions on the spatial distribution of people are important for various tasks, ranging from supply adjustment of location-based services to real-time crowd control related management and countermeasures.

Several further extensions can be identified, subjecting the methodology to specifically targeted modifications accomplishing a wider spectrum of contents. A possible research motive may focus on adapting the model for providing a prediction output that extends over multiple steps in the future, namely forecasting multiple hours ahead and assessing the performance horizon identifying the maximum point in time that still guarantees an advantage over baselines. A further evolution may concern the possibility of handling additional input information (e.g., event-specific instances, weather and environmental factors) within the model architecture, quantifying potential improvements originating from a mixture of contributions; for a widest possible generalization, we indeed leveraged only historical time series, not resorting to any additional information. More generally, similar variants of the proposed methodology can be applied to spatial-temporal phenomena of a different nature, still leveraging geographically-distributed sequential measurements characterized by rapid variations and abrupt location-specific patterns. Finally, following the recent expanding success of deep learning methodologies in a variety of research domains involving sequential processing tasks (e.g., [53,54]), focused improvements to the current architecture can be further explored, aiming to enhance the present achievements.

In conclusion, the proposed multi-target LSTM neural network-based model was demonstrated to promisingly handle multiple traffic series belonging to spatially-related reference points, arranging a collective processing of multiple input sequences into corresponding multiple forecasted output values. Representative of an effective application of deep learning on human mobility analysis, the paper contributes to extending a purely data-driven perspective on vehicle number forecasting, tackling a geographically-distributed sequential problem by means of deep neural networks, to aim for an improved quality of urban management and organization.

Author Contributions: A.C. conceived and designed the experiments, analyzed the data and wrote the paper. E.B. supervised the work, helped with designing the conceptual framework, and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. De Souza, A.M.; Brennand, C.A.R.L.; Yokoyama, R.S.; Donato, E.A.; Madeira, E.R.M.; Villas, L.A. Traffic management systems: A classification, review, challenges, and future perspectives. *Int. J. Distrib. Sens. Netw.* **2017**, *13*, 1550147716683612. [CrossRef]
2. Djahel, S.; Doolan, R.; Muntean, G.-M.; Murphy, J. A communications-oriented perspective on traffic management systems for smart cities: Challenges and innovative approaches. *IEEE Commun. Surv. Tutor.* **2014**, *17*, 125–151. [CrossRef]
3. Jain, V.; Sharma, A.; Subramanian, L. Road traffic congestion in the developing world. In Proceedings of the 2nd ACM Symposium on Computing for Development, Atlanta, GA, USA, 11–12 March 2012; pp. 1–10. [CrossRef]
4. Arnott, R. A bathtub model of downtown traffic congestion. *J. Urban Econ.* **2013**, *76*, 110–121. [CrossRef]
5. Gakis, E.; Kehagias, D.; Tzovaras, D. Mining traffic data for road incidents detection. In Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), Qingdao, China, 8–11 October 2014; pp. 930–935. [CrossRef]
6. Abdel-Aty, M.A.; Radwan, A.E. Modeling traffic accident occurrence and involvement. *Accid. Anal. Prev.* **2000**, *32*, 633–642. [CrossRef]
7. Sun, C.; Luo, Y.; Li, J. Urban traffic infrastructure investment and air pollution: Evidence from the 83 cities in china. *J. Clean. Prod.* **2018**, *172*, 488–496. [CrossRef]
8. Fiedler, P.E.K.; Zannin, P.H.T. Evaluation of noise pollution in urban traffic hubs—Noise maps and measurements. *Environ. Impact Assess. Rev.* **2015**, *51*, 1–9. [CrossRef]

9. Min, W.; Wynter, L. Real-time road traffic prediction with spatio-temporal correlations. *Transp. Res. Part C Emerg. Technol.* **2011**, *19*, 606–616. [CrossRef]

10. Li, Y.; Zheng, Y.; Zhang, H.; Chen, L. Traffic prediction in a bike-sharing system. In Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 3–6 November 2015; pp. 1–10. [CrossRef]

11. Lv, Y.; Duan, Y.; Kang, W.; Li, Z.; Wang, F.-Y. Traffic flow prediction with big data: A deep learning approach. *IEEE Trans. Intell. Transp. Syst.* **2014**, *16*, 865–873. [CrossRef]

12. Ma, X.; Dai, Z.; He, Z.; Ma, J.; Wang, Y.; Wang, Y. Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction. *Sensors* **2017**, *17*, 818. [CrossRef]

13. Polson, N.G.; Sokolov, V.O. Deep learning for short-term traffic flow prediction. *Transp. Res. Part C Emerg. Technol.* **2017**, *79*, 1–17. [CrossRef]

14. Li, Y.; Yu, R.; Shahabi, C.; Liu, Y. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv* **2017**, arXiv:1707.01926.

15. Kong, Q.-J.; Li, Z.; Chen, Y.; Liu, Y. An approach to urban traffic state estimation by fusing multisource information. *IEEE Trans. Intell. Transp. Syst.* **2009**, *10*, 499–511. [CrossRef]

16. Shankar, H.; Raju, P.L.N.; Rao, K.R.M. Multi model criteria for the estimation of road traffic congestion from traffic flow information based on fuzzy logic. *J. Transp. Technol.* **2012**, *2*, 50–62. [CrossRef]

17. Kong, X.; Xu, Z.; Shen, G.; Wang, J.; Yang, Q.; Zhang, B. Urban traffic congestion estimation and prediction based on floating car trajectory data. *Future Gener. Comput. Syst.* **2016**, *61*, 97–107. [CrossRef]

18. Andre, H.; Ny, J.L. A differentially private ensemble kalman filter for road traffic estimation. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 6409–6413. [CrossRef]

19. He, Z.; Cao, J.; Li, T. Mice: A real-time traffic estimation based vehicular path planning solution using vanets. In Proceedings of the 2012 International Conference on Connected Vehicles and Expo (ICCVE), Beijing, China, 12–16 December 2012; pp. 172–178. [CrossRef]

20. D'Acierno, L.; Cartenì, A.; Montella, B. Estimation of urban traffic conditions using an automatic vehicle location (avl) system. *Eur. J. Oper. Res.* **2009**, *196*, 719–736. [CrossRef]

21. Phillips, D.J.; Wheeler, T.A.; Kochenderfer, M.J. Generalizable intention prediction of human drivers at intersections. In Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, USA, 11–14 June 2017; pp. 1665–1670. [CrossRef]

22. Houenou, A.; Bonnifait, P.; Cherfaoui, V.; Yao, W. Vehicle trajectory prediction based on motion model and maneuver recognition. In Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 3–7 November 2013; pp. 4363–4369. [CrossRef]

23. Zernetsch, S.; Kohnen, S.; Goldhammer, M.; Doll, K.; Sick, B. Trajectory prediction of cyclists using a physical model and an artificial neural network. In Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV), Gothenburg, Sweden, 19–22 June 2016; pp. 833–838. [CrossRef]

24. Vlahogianni, E.I.; Karlaftis, M.G.; Golias, J.C. Statistical methods for detecting nonlinearity and non-stationarity in univariate short-term time-series of traffic volume. *Transp. Res. Part C Emerg. Technol.* **2006**, *14*, 351–367. [CrossRef]

25. Lin, L.; Li, Y.; Sadek, A. A k nearest neighbor based local linear wavelet neural network model for on-line short-term traffic volume prediction. *Procedia Soc. Behav. Sci.* **2013**, *96*, 2066–2077. [CrossRef]

26. Xu, D.-w.; Wang, Y.-d.; Jia, L.-m.; Qin, Y.; Dong, H.-h. Real-time road traffic state prediction based on arima and kalman filter. *Front. Inf. Technol. Electron. Eng.* **2017**, *18*, 287–302. [CrossRef]

27. Qu, L.; Li, W.; Li, W.; Ma, D.; Wang, Y. Daily long-term traffic flow forecasting based on a deep neural network. *Expert Syst. Appl.* **2019**, *121*, 304–312. [CrossRef]

28. Benevolo, C.; Dameri, R.P.; D'Auria, B. Smart mobility in smart city. In *Empowering Organizations*; Springer International Publishing: Cham, Switzerland, 2016; pp. 13–28. [CrossRef]

29. Torre-Bastida, A.I.; Ser, J.D.; Laña, I.; Ilardia, M.; Bilbao, M.N.; Campos-Cordobés, S. Big data for transportation and mobility: Recent advances, trends and challenges. In *IET Intelligent Transport Systems*; Institution of Engineering and Technology: London, UK, 2018; Volume 12, pp. 742–755. [CrossRef]

30. Han, J.; Li, Z.; Tang, L.A. Mining moving object, trajectory and traffic data. In *Database Systems for Advanced Applications*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 485–486. [CrossRef]

31. Alam, I.; Ahmed, M.F.; Alam, M.; Ulisses, J.; Farid, D.M.; Shatabda, S.; Rossetti, R.J.F. Pattern mining from historical traffic big data. In Proceedings of the 2017 IEEE Region 10 Symposium (TENSYMP), Cochin, India, 14–16 July 2017; pp. 1–5. [CrossRef]

32. Metropolitan Transportation Authority. Bridges and Tunnels Committee Meeting, November 2019. Available online: https://new.mta.info/document/12071 (accessed on 5 November 2020).

33. Aslanargun, A.; Mammadov, M.; Yazici, B.; Yolacan, S. Comparison of arima, neural networks and hybrid models in time series: Tourist arrival forecasting. *J. Stat. Comput. Simul.* **2007**, *77*, 29–53. [CrossRef]

34. Rotela, P.R., Jr.; Salomon, F.L.R.; de Oliveira Pamplona, E. Arima: An applied time series forecasting model for the bovespa stock index. *Appl. Math.* **2014**, *5*, 3383–3391. [CrossRef]

35. Chmieliauskas, D.; Guršnys, D. Lte cell traffic grow and congestion forecasting. In Proceedings of the 2019 Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 25 April 2019; pp. 1–5. [CrossRef]

36. Samal, K.K.R.; Babu, K.S.; Das, S.K.; Acharaya, A. Time Series based Air Pollution Forecasting using SARIMA and Prophet Model. In Proceedings of the 2019 International Conference on Information Technology and Computer Communications (ITCC 2019), Singapore, 16–18 August 2019; pp. 80–85. [CrossRef]

37. Williams, B.M.; Durvasula, P.K.; Brown, D.E. Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models. *Transp. Res. Rec.* **1998**, *1644*, 132–141. [CrossRef]

38. Williams, B.M. Multivariate vehicular traffic flow prediction: Evaluation of arimax modeling. *Transp. Res. Rec.* **2001**, *1776*, 194–200. [CrossRef]

39. Stathopoulos, A.; Karlaftis, M.G. A multivariate state space approach for urban traffic flow modeling and prediction. *Transp. Res. Part C Emerg. Technol.* **2003**, *11*, 121–135. [CrossRef]

40. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]

41. Dai, S.; Li, L.; Li, Z. Modeling vehicle interactions via modified lstm models for trajectory prediction. *IEEE Access* **2019**, *7*, 38287–38296. [CrossRef]

42. Crivellari, A.; Beinat, E. LSTM-Based Deep Learning Model for Predicting Individual Mobility Traces of Short-Term Foreign Tourists. *Sustainability* **2020**, *12*, 349. [CrossRef]

43. Cheng, B.; Xu, X.; Zeng, Y.; Ren, J.; Jung, S. Pedestrian trajectory prediction via the social-grid lstm model. *J. Eng.* **2018**, *2018*, 1468–1474. [CrossRef]

44. Trinh, H.D.; Giupponi, L.; Dini, P. Mobile traffic prediction from raw data using lstm networks. In Proceedings of the 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Bologna, Italy, 9–12 September 2018; pp. 1827–1832. [CrossRef]

45. Yang, D.; Chen, K.; Yang, M.; Zhao, X. Urban rail transit passenger flow forecast based on lstm with enhanced long-term features. *IET Intell. Transp. Syst.* **2019**, *13*, 1475–1482. [CrossRef]

46. Pasini, K.; Khouadjia, M.; Samé, A.; Ganansia, F.; Oukhellou, L. Lstm encoder-predictor for short-term train load forecasting. In *Machine Learning and Knowledge Discovery in Databases*; Springer International Publishing: Cham, Switzerland, 2020; pp. 535–551. [CrossRef]

47. Fu, R.; Zhang, Z.; Li, L. Using lstm and gru neural network methods for traffic flow prediction. In Proceedings of the 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China, 11–13 November 2016; pp. 324–328. [CrossRef]

48. Liu, B.; Cheng, J.; Cai, K.; Shi, P.; Tang, X. Singular point probability improve lstm network performance for long-term traffic flow prediction. In *Theoretical Computer Science*; Springer: Singapore, 2017; pp. 328–340. [CrossRef]

49. Kang, D.; Lv, Y.; Chen, Y. Short-term traffic flow prediction with lstm recurrent neural network. In Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; pp. 1–6. [CrossRef]

50. Zhang, D.; Kabuka, M.R. Combining weather condition data to predict traffic flow: A gru-based deep learning approach. *IET Intell. Transp. Syst.* **2018**, *12*, 578–585. [CrossRef]

51. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980 2014.

52. Lytras, M.D.; Visvizi, A.; Sarirete, A. Clustering Smart City Services: Perceptions, Expectations, Responses. *Sustainability* **2019**, *11*, 1669. [CrossRef]

53. Nguyen, G.; Dlugolinsky, S.; Tran, V.; García, Á.L. Deep learning for proactive network monitoring and security protection. *IEEE Access* **2020**, *8*, 19696–19716. [CrossRef]
54. Mammone, N.; Ieracitano, C.; Morabito, F.C. A deep CNN approach to decode motor preparation of upper limbs from time–frequency maps of EEG signals at source level. *Neural Netw.* **2020**, *124*, 357–372. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.