Alexandria Engineering Journal (2023) 65, 151–162

HOSTED BY

ELSEVIER

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

ORIGINAL ARTICLE

Multivariate machine learning-based prediction models of freeway traffic flow under non-recurrent events

Fahad Aljuaydi [a,b,*], Benchawan Wiwatanapataphee [a], Yong Hong Wu [a]

[a] EECMS, Curtin University, Perth WA, Australia
[b] Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

Received 4 September 2022; revised 26 September 2022; accepted 4 October 2022
Available online 26 October 2022

KEYWORDS

Non-recurrent events;
Traffic prediction;
Multivariate model;
Machine Learning

Abstract This paper concerns multivariate machine learning-based prediction models of freeway traffic flow under non-recurrent events. Five model architectures based on the multi-layer perceptron (MLP), convolutional neural network (CNN), long short-term memory (LSTM), CNN-LSTM and Autoencoder LSTM networks have been developed to predict traffic flow under a road crash and the rain. Using an input dataset with five features (the flow rate, the speed, and the density, road incident and rainfall) and two standard metrics (the Root Mean Square error and the Mean Absolute error), models' performance is evaluated.

© 2022 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Engineering, Alexandria University. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The Intelligent Transportation System (ITS) relying heavily on traffic flow prediction enables traffic stakeholders to use transportation networks safer and more efficiently [1,2]. Individual travelers, traffic managers, policymakers, and road users are all traffic stakeholders.

The quality of traffic data determines the effectiveness of these systems, and only then can an ITS be successful. According to the World Health Organization's Global Status Report on Road Safety, road traffic deaths continue to rise, with 1.35 million recorded in 2016. This makes traffic forecasting a vital tool in reducing congestion and making travel safer and more cost-effective [3,4].

The big question is whether or not traffic patterns, queuing patterns, and time can be accurately predicted. Can forecasting traffic flow assist decision-makers in avoiding potential traffic jams? As a result, these questions have emerged as critical areas of study, particularly in urban settings. The urban fleet of vehicles has recently advanced to the point where filtered sensor data can be uploaded directly to the cloud [5].

Predicting traffic flow is improved when a vehicular cloud provides services to autonomous vehicles. However, some prediction models find it challenging to estimate traffic with accuracy due to the integration of many variables from various

* Corresponding author.
E-mail addresses: f.aljuaydi@postgrad.curtin.edu.au, f.aljuaydi@psau.edu.sa (F. Aljuaydi), b.wiwatanapataphee@curtin.edu.au (B. Wiwatanapataphee), Y.Wu@curtin.edu.au (Y.H. Wu).

Peer review under responsibility of Faculty of Engineering, Alexandria University.

https://doi.org/10.1016/j.aej.2022.10.015
1110-0168 © 2022 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Engineering, Alexandria University.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

road segments and time-varying traffic patterns. It is challenging to accurately estimate traffic parameters in prediction models since traffic is unpredictable and dynamic [6]. Prediction models are severely constrained by physical infrastructure and traffic laws, and external events like accidents, weather, and road closures significantly impact the model. Understanding the various architectures created to handle this issue is crucial.

The transportation industry has recently transitioned into the big data age [7]. The collection of relevant data, such as traffic speed, weather, and accident data, heavily relies on sensors and other equipment in traffic congestion forecasting. While traditional models employ shallow networks, vehicles on the road have grown exponentially in recent years, and traditional models are no longer applicable in current scenarios. Traditionally, the models utilise parametric methods [8]. For predicting the short-term traffic flow, well-known standards and frameworks are employed, e.g., ARIMA (Auto-Regressive Integrated Average) model [9]. The results confirm that ARIMA model was adapted and altered for better performance which is confirmed [10–12]. The inherited nature of the traffic flow is non-deterministic (stochastic) and non-linear, which is why the parametric model's predictions are not accurate [13]. Hence, the non-parametric model has been chosen. Artificial Neural Network (ANN) performed well, but accuracy decreased with big data reported by Smith et al. [14]. As a reason, various prediction models have been developed based on deep learning models such as Restricted Boltzmann Machines (RBM), Recurrent Neural Networks (RNN), Short-Term Memory (LSTM), Deep Belief Network (DBN), and Convolutional Neural Networks (CNN) [15–17].

The core idea behind Artificial Neural Network (ANN) is that of a brain which functions with the help of a large collection of interconnected units for communicating between the units [18]. A Multilayer perceptron (MLP) is a straightforward example of a feed-forward ANN with input, hidden, and output layers as its three structural components. The depth of ANN, as the procedure is known, allows for more than one hidden layer. An increase in the number of hidden layers results in a higher resolution of information characterization. In fact, including many hidden layers in an MLP is also referred to as a deep learning algorithm. Each hidden layer constitutes a certain number of inter-linked neurons with other neurons, each connection having a certain weight. In ANN, these weights are updated based on the information to determine an input–output relationship. The weights are trained using a supervised machine learning technique known as backward propagation [19,20]. However, unlike other deep learning architectures, including feedback loops, MLP only comprises forward connections between two neurons. Despite this, neurons in the same layer are not linked. The information flows from the input layer to the output layer, hence the name feed-forward. MLP uses the Backpropagation training process. A supervised learning algorithm in which the MLP learns the desired output from varied entry data. It uses the input/output pair of data train algorithms while unsupervised learning algorithms uses unlabeled data for clustering and analysis.

CNN is a deep learning algorithm mainly applied in the image processing domain [21]. Other applications of CNN include speech recognition [22], natural language processing [23] and object recognition [24]. The idea behind CNN is that of the animal visual cortex, characterized by connections between neurons [25–27]. A CNN using its neuron organization can identify features from higher dimensional data, e.g. the pixels in an image. The input and output layers of CNN, like those of MLP, are separated by a number of hidden layers that can be combined, pooled, or completely coupled. These convolution layers employ filters to capture characteristics using various input information settings. A pooling layer can give an abstract representation of the data by reducing the dimensionality of the data. The prediction of the non-linearity of traffic flow from the deep learning methods showed promising results. The popular CNNs models have incorporated the 3-dimensional traffic dependencies [28,29]. The time-based (short-term and long-term) traffic flow relationships are accurately predicted by the RNNs model, especially with LSTM designs [30,31]. Along with the advantages of the models mentioned above, there are many disadvantages, since the implemented transportation system is highly based on data, and any missing data can lead to inaccurate results [32]. The current study [33] tries to bridge the gap which is created due to the missing data by using data imputation and implementing the unsupervised tensor completion method. Deep learning systems, e.g., RNNs or CNNs, need a large amount of data which is quite an easy task nowadays as large datasets are available, but sometimes these models can over-fit the model because of the large fluctuations in the traffic flow within a small interval of time [34].

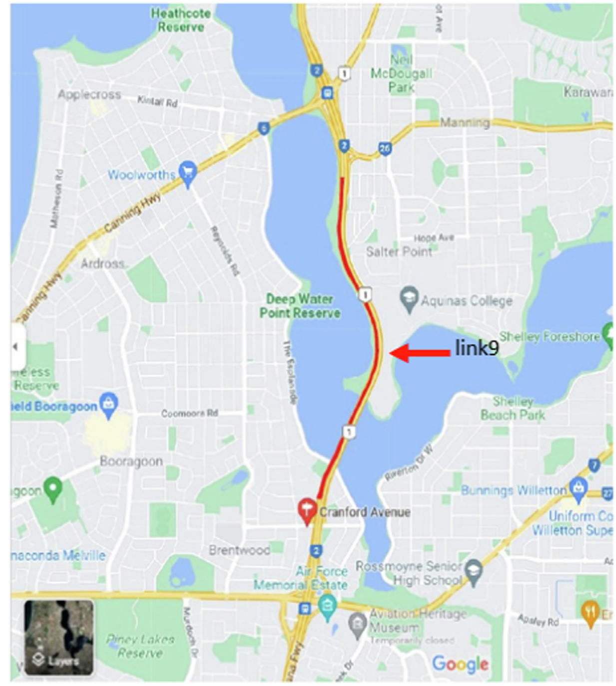
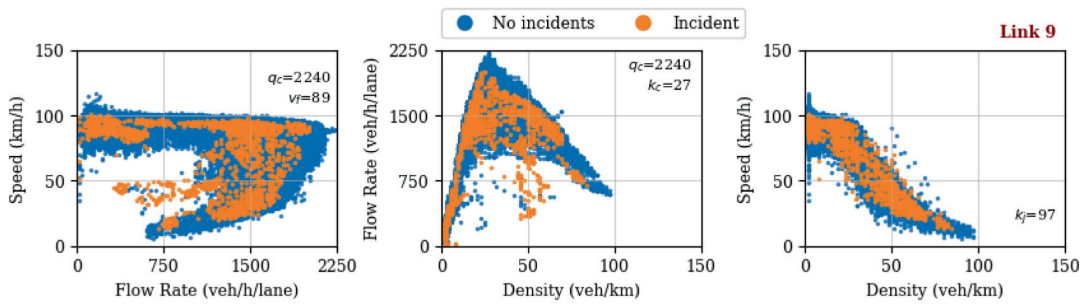


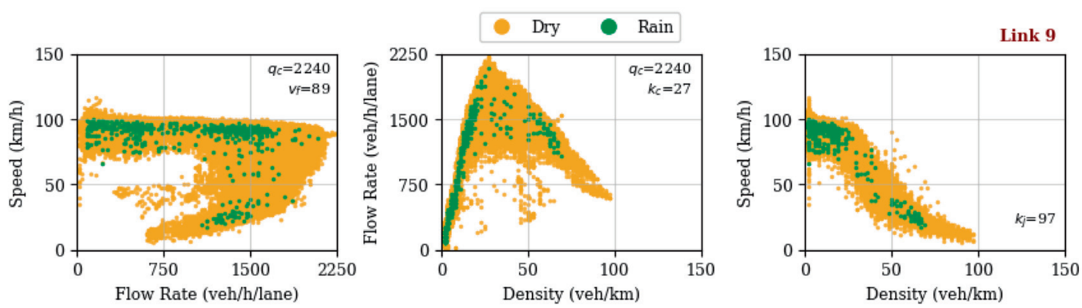
Fig. 1 Study region (red curve), Link9: between the Cranford Avenue on-ramp and the Canning Highway northbound off-ramp [47].

Recent review articles discussed the traffic flow prediction details regarding the urban flow prediction models, machine learning-based methods, and statistical models [35,36,8,2]. They also discussed the many available deep learning architectures and looked at the increasing popularity of numerous hybrid methods [37]. Due to this, researchers have been adapting unsupervised methods with a hybrid approach instead of a deep learning architect [34,38]. The particular correlation was developed by applying the Deep CNN model to random subspace learning [28]. Complex computations and multidimensional inputs can be treated very easily with typical DNN. The temporal and spatial characteristics can be estimated by employing the self-learned filters. The model generalization is improved by the CNN, and the overall computational time was reduced. Due to these benefits, CNNs are considered helpful in predicting traffic flow.

A Recurrent Neural Network (RNN) takes the outputs from the previous models and sometimes gets information from the past and acts as a new type of neural network [39,40]. RNN is extensively implemented by many researchers in speech recognition and Natural Language Processing (NLP). LSTM is also one of the architectures of RNN and is widely used for processing time-series data [41]. It uses hidden components like memory cells. LSTM is useful for analyzing long-time series data and forecasting with autocorrelations



(a) Road incident effect



(b) Rainfall effect

Fig. 2 Relationships of traffic variables under road incidents and rainfall effects: (a) road incident; (b) rainfall between medium and heavy level.

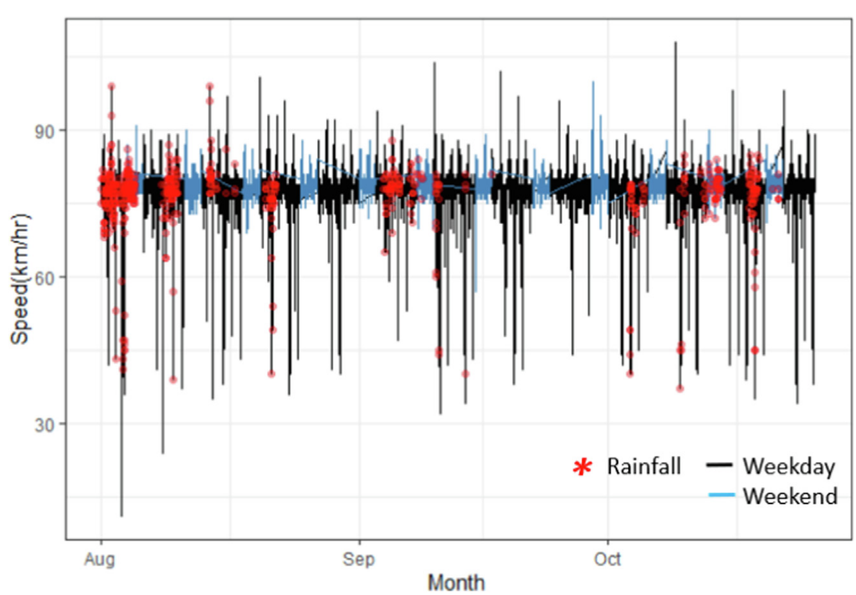


Fig. 3 Effect of rain on traffic speed from 1 August to 1 November 2018.

[42]. The LSTM model [43] is an effective recurrent neural system created expressly to prevent the exploding or vanishing gradient issues encountered when learning long-term dependencies, even when the smallest time lags are fairly substantial [44]. A constant error carousel (CEC), which retains the error signal inside each unit's cell, is typically used to prevent this. Such cells are recurrent networks, having an intriguing design of how the CEC is augmented with new components, particularly the input and output gates, to create the memory cell. The feedback is indicated by the self-recurrent connections with a one-step latency. A cell, an input gate, an output gate, and a forget gate make up a standard LSTM unit. The forget gate was not initially included in the LSTM network, but Gers et al. [45] proposed it as a way for the network to reset its state. The three gates control the flow of information associated with the cell, and the cell remembers values across arbitrary time intervals. In short, the LSTM architecture comprises memory blocks, a collection of recurrently connected sub-networks. The memory block's goal is to maintain its state over time while using nonlinear gating devices to control information flow.

For the temporal characteristic of sequential data, the fundamental problem of gradients has vanished and helped them to train against the Long-term dependencies. Deep neural network modelling skills are enhanced by utilizing the benefits of RNN and CNN [17]. T.Sainth et al. [17] combined the LSTM with the CNN for the temporal sequence. They observed an improvement of 4–6% when implementing the different vocabulary tasks. Various research studies have also combined the LSTM and CNN model for the extraction of spatial and temporal characteristics and it is found that the results were better when compared to one model data [41].

The main issue in ML models is overfitting the training set and lengthy training times. Thus, training using additional data, employing early stopping to cease the network's training at the appropriate point, and using the proper number of epochs are all necessary to prevent the overfitting of the model. Recent studies have concentrated on predicting numerous outcomes from multiple inputs utilizing a data set with different classifications (features). Road incident is one of these features that has a substantial impact on road capacity. The vehicle movement and traffic accidents within the tunnels were being detected and studied using the deep learning method [46]. Recently, multivariate deep learning model is the most challenging techniques for traffic flow prediction under non-recurrent event. In this study, multivariate prediction models based on a deep-learning model and a combined deep-learning model have been developed to predict freeway traffic flow under non-recurrent events. The work is structured as follows. Section 2 presents the study region and the input dataset with 5 features including the traffic volume, speed, density, boolean road incident and rainfall. Section 3 concerns the methodology for traffic prediction using proposed models.

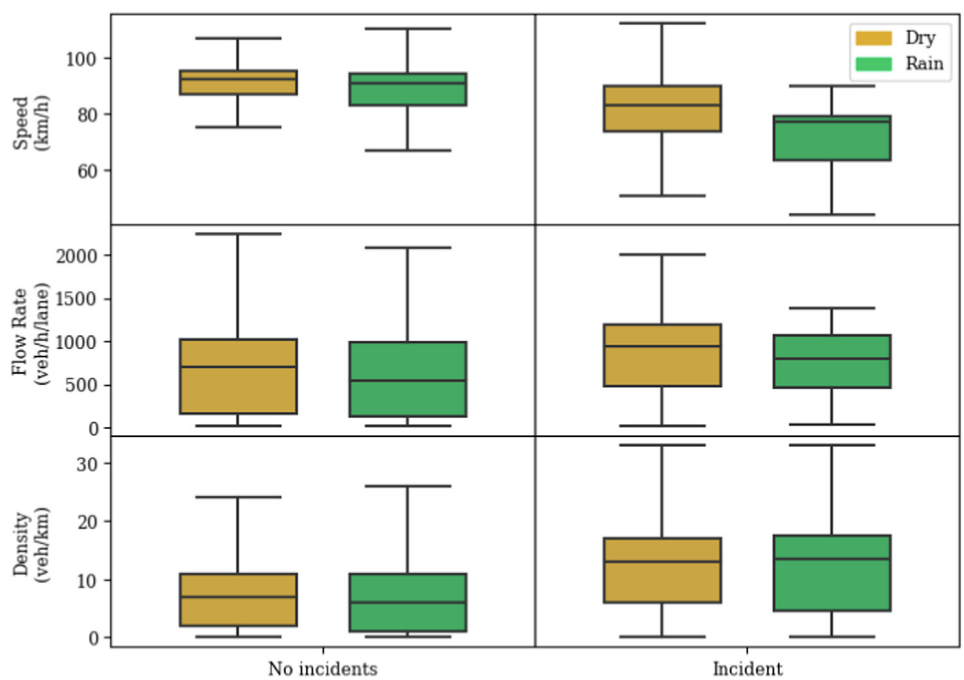


Fig. 4 Box plot of traffic variables with and without road incident and rain.

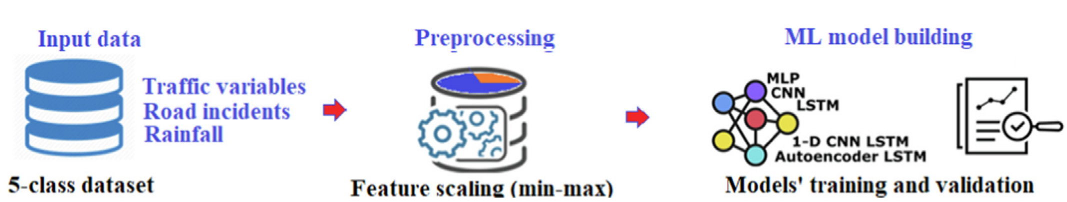
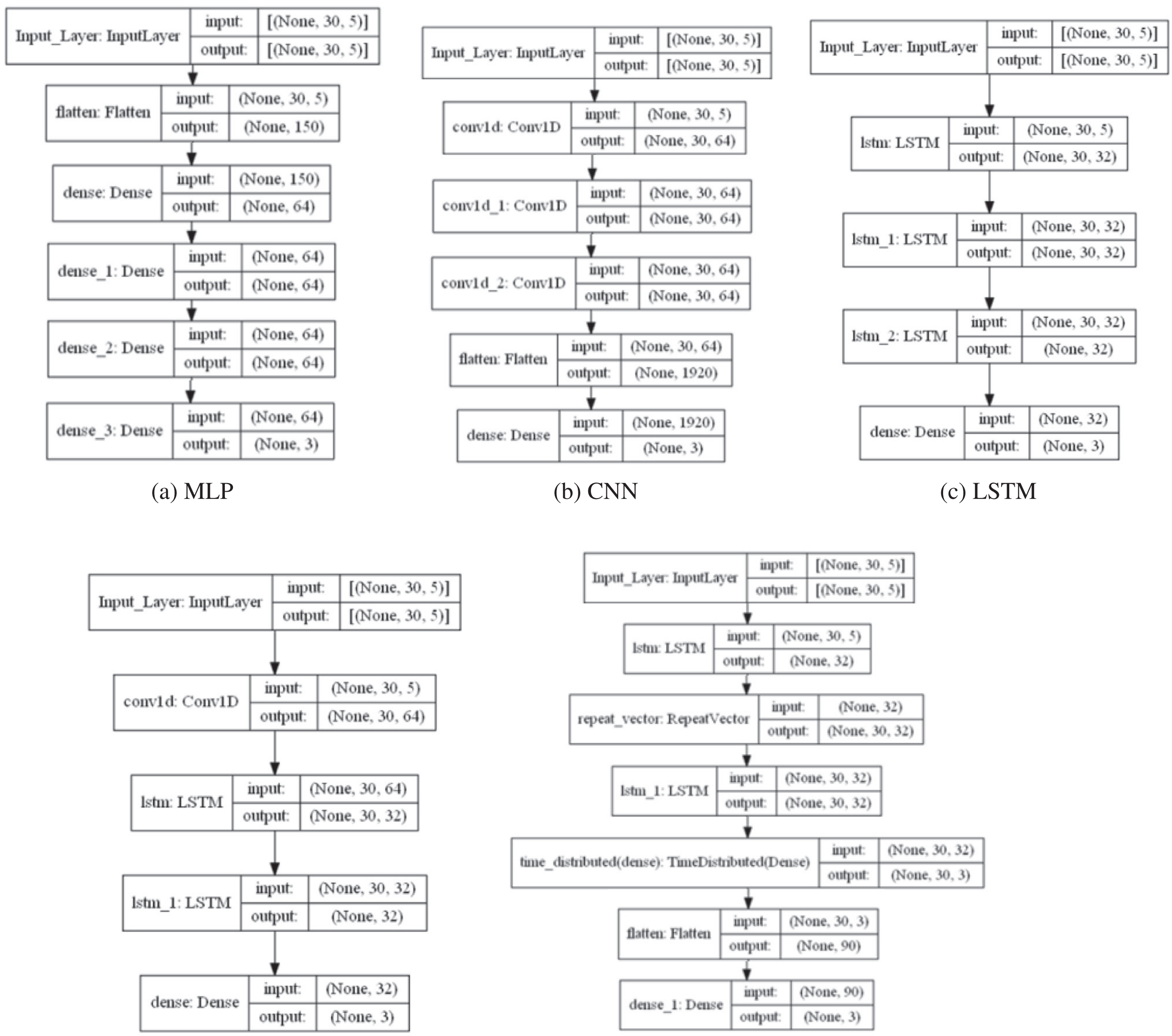


Fig. 5 Machine learning workflow.



(a) MLP (b) CNN (c) LSTM

(d) CNN LSTM (e) Autoencoder LSTM

Fig. 6 Architectures of various ML models: (a) MLP; (b) CNN; (c) LSTM; (d) 1D-CNN LSTM; (e) Autoencoder LSTM.

Table 1 Two standard metrics, RMSEs and MAEs, of ML for each traffic parameter.

RMSE												
Parameter	**Baseline**		**MLP**		**CNN**		**LSTM**		**1D CNN-LSTM**		**AE-LSTM**	
	Train	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**
Flow rate	11.02	11.75	3.58	3.61	3.10	3.15	3.02	3.08	2.50	2.65	2.89	2.90
Speed	10.25	10.81	5.23	5.35	5.0	5.09	4.89	5.0	4.60	4.63	4.75	4.77
Density	24.15	24.71	7.32	7.55	7.22	7.29	7.02	7.11	6.65	6.70	6.75	6.80
MAE												
Parameter	**Baseline**		**MLP**		**CNN**		**LSTM**		**1D CNN-LSTM**		**AE-LSTM**	
	Train	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**
Flow rate	6.52	6.98	4.18	4.63	4.0	4.10	3.90	3.98	3.70	3.75	3.75	3.80
Speed	5.49	6.18	2.75	2.80	2.70	2.71	2.58	2.60	2.46	2.50	2.50	2.55
Density	11.14	11.95	5.51	5.48	5.32	5.39	5.12	5.15	4.52	5.0	5.01	5.08

Results and discussion are given in Section 5. The conclusion is given in Section 6.

2. Study area and data

The study area, the Kwinana Smart Freeway in Western Australia between the Cranford on-ramp and the Canning Highway off-ramp (link 9) with a total length of 2.13 km as shown in Fig. 1, is chosen because it has a distinguished record of the highest number of road accidents. For an effective multivariate ML model for long short-term traffic prediction, a dataset of traffic variables (the flow rate, the speed, and the density), road incidents and rainfall is used in this study. Traffic data including the flow rate (volume), the speed and the density from the Main Road Western Australia (WRWA)

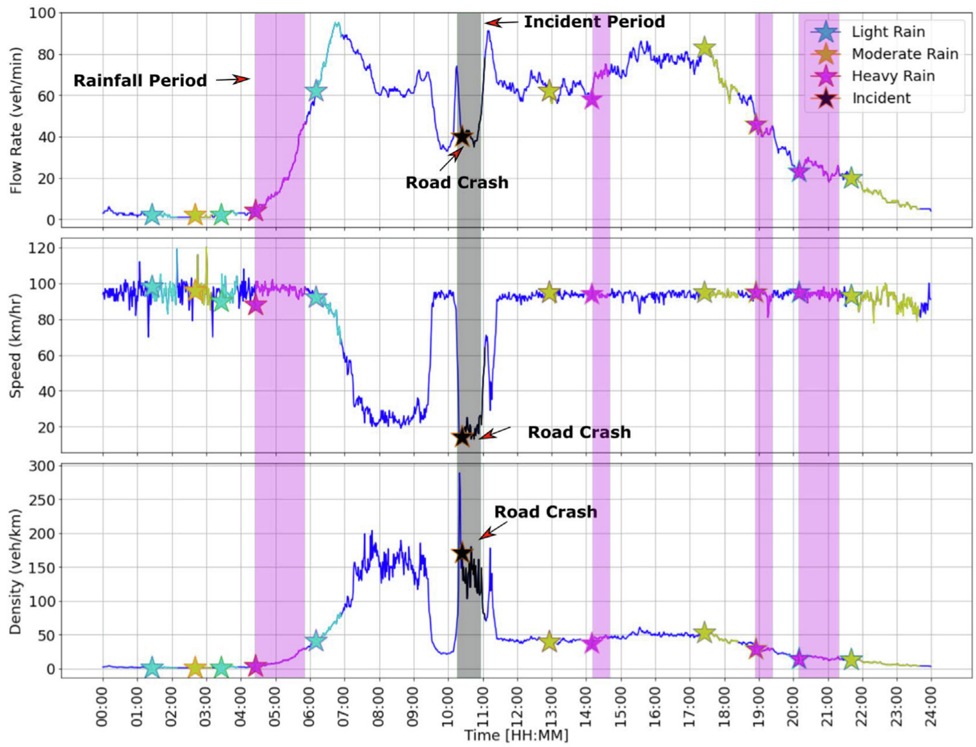


Fig. 7 Observed traffic flow rate (top), speed (middle) and density (bottom) with road crashes and rainfall during the prediction period (4 September 2018).

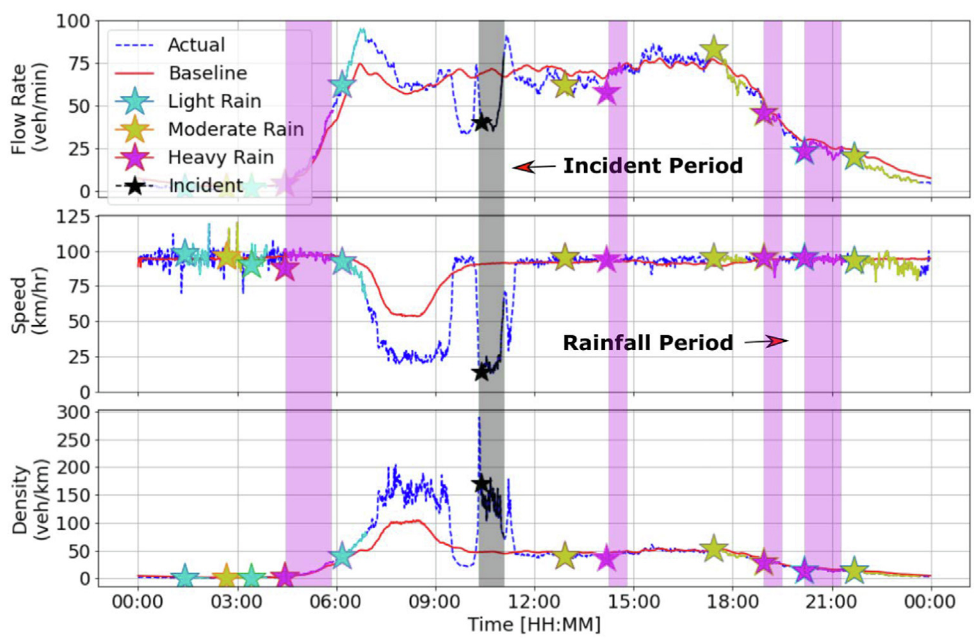


Fig. 8 Baseline predictions of traffic flow rate (veh/min), speed (km/hr) and density (veh/km) under a road incident (★) and rain.

are available between 1 January and 25 November 2018. In the same period of available traffic data, road incidents and rainfall data are obtained from the Web Emergency Operations Centre (WebEOC) and the Bureau of Meteorology (BOM) WA, respectively.

2.1. Input data

Road incident data is converted into Boolean data by considering the non-existence as zero and the existence of the incident as one. For rainfall message, rain rate intensity, is considered in three categories including the light rain with precipitation less than 0.1 inches per hour (*iph*), the moderate rain with precipitation between 0.1 and 2.5 *iph* and the heavy rain with rainfall greater than 2.5 *iph*. The input data is then obtained by matching the timestamp of the traffic data, the boolean incident data and the rainfall data. In this study, we have a 1-min input dataset with 429,120 observations and five features.

2.2. Road incidents and rain effects

Road incidents including crashes, vehicle breakdowns and debris commonly affect the flow of traffic. Fig. 2 shows the relationships of traffic variables under road incidents and rainfall effects. It is noted that the road incidents have an impact on the traffic flow on the roadway. It may reduce 10–25% of traffic capacity. Fig. 3 presents speed profile with rain's effect on Link 9 between 1 August and 1 November 2018. It indicates that speed will drop on the roadway when it rains, as a driver driving in the wet commonly reduces the speed to allow the car's tyres to grip to the road at all times.

Fig. 4 presents a box plot showing traffic flow with and without the impact of the road incidents and the rain. It demonstrates that the road incidents have a negative impact on the speed and the density of traffic, and the rain seems to magnify the effect of road incidents on the flow of traffic, indicating by the significant higher density and lower speed. When road incidents occur, the flow capacity decreases while the density increases.

3. Methodology

This section concerns building model architectures of multivariate prediction models based on the MLP, CNN, LSTM, CNN-LSTM and Autoencoder LSTM networks.

The machine learning workflow, as illustrated in Fig. 5, consists of data preparation, pre-processing, model training, and model testing and evaluation.

3.1. Preparing data and pre-processing

Gathering, sorting and cleaning all datasets are needed for development of the predictive learning model as any discrepancies in the data will lead to the failure analysis of the predictive model. The input dataset with n observations and five features (classes) is in the form of

$$\mathbf{X} = (X_i^1, X_i^2, X_i^3, X_i^4, X_i^5)_{i=1}^n,$$

where traffic parameters X^1, X^2, X^3 denote respectively traffic volume, speed and density, X^4 is the boolean road incident data, and X^5 is the rainfall data.

The traffic parameters, road incident and rainfall data are homogenised by feature scaling. By implementing the min–max normalisation, x_i^c is transformed to ξ_i^c, provided ξ_i^c greater than zero and less than 1,

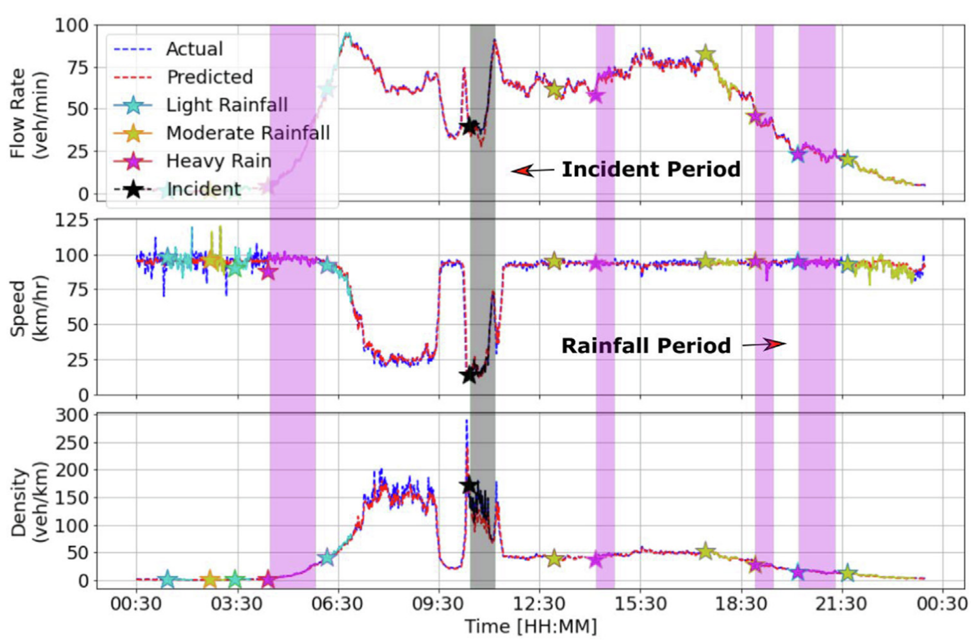


Fig. 9 Long-term traffic prediction of the flow rate (veh/min), speed (km/hr) and density (veh/km) obtained from the best multivariate ML model based on the 1D CNN-LSTM network.

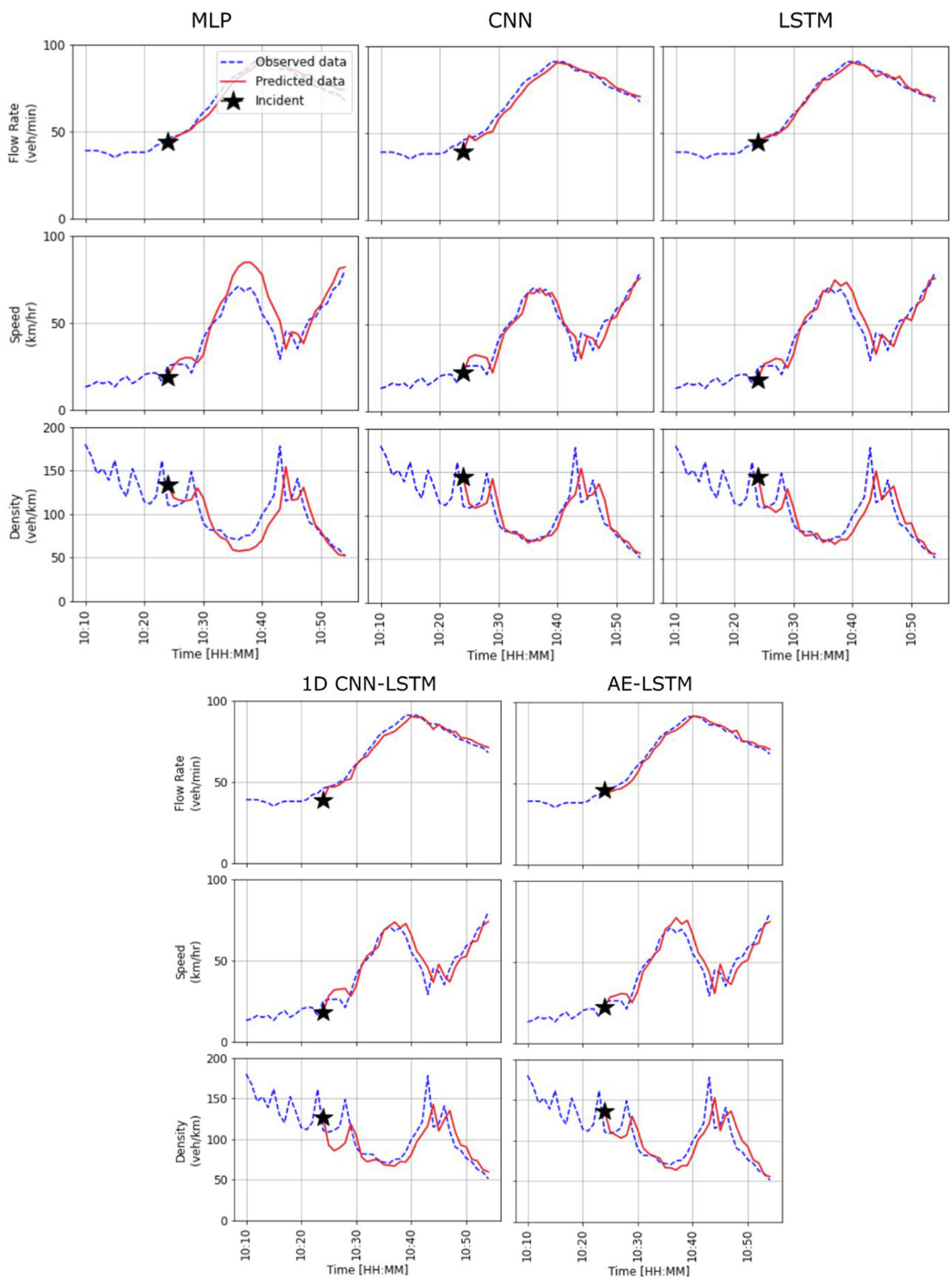


Fig. 10 Short-term predictions under a road incident on 4 September 2018 between 10:25 and 10:55 obtained from five prediction models based on various ML networks: the MLP, the CNN, the LSTM, the 1D CNN-LSTM and the Autoencoder LSTM networks.

$$\xi_i^c = \frac{x_i^c - \min_c}{\max_c - \min_c}, \quad c = 1, \ldots, 5, \tag{1}$$

where X^c is the values of the observed set of x_i^c, $\max_c$ and $\min_c$ represent the maximum and minimum values of X^c, respectively.

4. Model architecture

The normalised input data with n observation and five features are split into the test set (30%) and the training set (70%). For each ML model, its optimal hyperparameters are found by grid searching, the Adam optimiser [48] and an early stopping technique. The accuracy and efficiency of each trained model are assessed by two standard metrics, the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE). These accuracies and efficiency give information about the goodness of the learning model. The deviation from the mean value estimated through the RMSE and MAE can be calculated using the following equations.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{2}$$

and

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{3}$$

Fig. 6 demonstates five model architectures based on the MLP, the CNN, the LSTM, the 1D-CNN LSTM and the Autoencoder LSTM networks.

Table 1 presents model validation using two standard metrics, the RMSEs and MAEs. A lower value of the MAE or the RMSE signifies a better model fit. The results indicate that all ML models give better prediction than the baseline model and the 1D CNN-LSTM model outperforms other ML models including the MLP, the CNN, the LSTM and the Autoencoder LSTM models. Comparing the RMSEs and the MAEs of other models, the 1D CNN-LSTM model gives the lowest values of the RMSEs and MAEs for all cases, i.e., RMSEs of 2.65, 4.63 and 6.70 and the MAEs of 3.75, 2.50 and 5.0 for the flow rate, speed and density predictions, respectively.

5. Results and discussion

For long and short-term predictions of traffic flow under a road crash and the wet road with rain intensity between medium and high rate, the Link-9 observed traffic flow under non-recurrent events on 4 September 2018 is chosen in this study because there was a long-period road crash between 10:24 and 13:09 (black star with dark solid line) and three periods of heavy rain, i.e., 4:30–6:00, 14:00–14:30 and 19:00–21:30 (purple star with purple solid line) as shown in Fig. 7. Using all proposed ML models, traffic variables including the flow rate, the speed and the density under non-recurrent events on Link 9 of the Kwinana Freeway are predicted and compared to find the optimal prediction model.

Fig. 8 shows long-term predictions of traffic parameters using the baseline model. As the baseline model gives average values of each traffic parameters, it thus cannot capture traffic patterns during non-recurrent events. Here, we present the performance of various ML models for predicting traffic pattern under the road crash and the rain. It is found that the best long-term prediction model is based on the 1D CNN-LSTM networks. Its prediction performance is illustrated in Fig. 9.

For short-term prediction of traffic flow under non-recurrent events, we consider separately the effect of a road crash and the heavy rain on traffic flow within 30 min after the occurrence of the incident (solid line).

In this study, observed data of traffic flow under a road crash (a dark star) between 10:25 and 10:55, and the heavy rain (a purple star) between 20:10 and 20:40 were compared with the predicted data. Fig. 10 presents the short-term prediction of traffic flow under a road crash. It illustrates that two ML models based on the LSTM and the 1D CNN-LSTM networks performed better than other ML models as they attained the low values of the RMSEs of 1.75, 2.85 and 2.50, and MAEs of 2.30, 2.18 and 2.75 for the flow rate, speed and density predictions, respectively, as shown in Table 2.

Fig. 11 shows the short-term prediction of traffic flow under the heavy rain. The results indicate that the MLP model gives the worst prediction with RMSEs of 4.03, 5.30 and 7.09 for the flow rate, speed and density predictions, respectively. The 1D-CNN LSTM model performed better than other models as it attained RMSEs of 1.18, 3.45 and 1.15, and MAEs of 1.67, 1.73 and 1.87 for the flow rate, speed and density predictions, respectively, as shown in Table 3.

Table 2 RSMEs and MAEs of ML models' performance for short-term prediction under **a road crash**.

RMSE						
Parameter	**Baseline**	**MLP**	**CNN**	**LSTM**	**1D CNN-LSTM**	**AE-LSTM**
Flow rate	9.65	1.85	1.82	1.80	1.75	1.78
Speed	14.26	3.57	2.98	2.92	2.85	2.89
Density	10.38	3.59	2.65	2.63	2.50	2.55
MAE						
Parameter	**Baseline**	**MLP**	**CNN**	**LSTM**	**1D CNN-LSTM**	**AE-LSTM**
Flow rate	5.93	2.38	2.38	2.33	2.30	2.35
Speed	7.14	3.15	2.27	2.24	2.18	2.22
Density	5.09	3.65	2.80	2.85	2.75	2.79

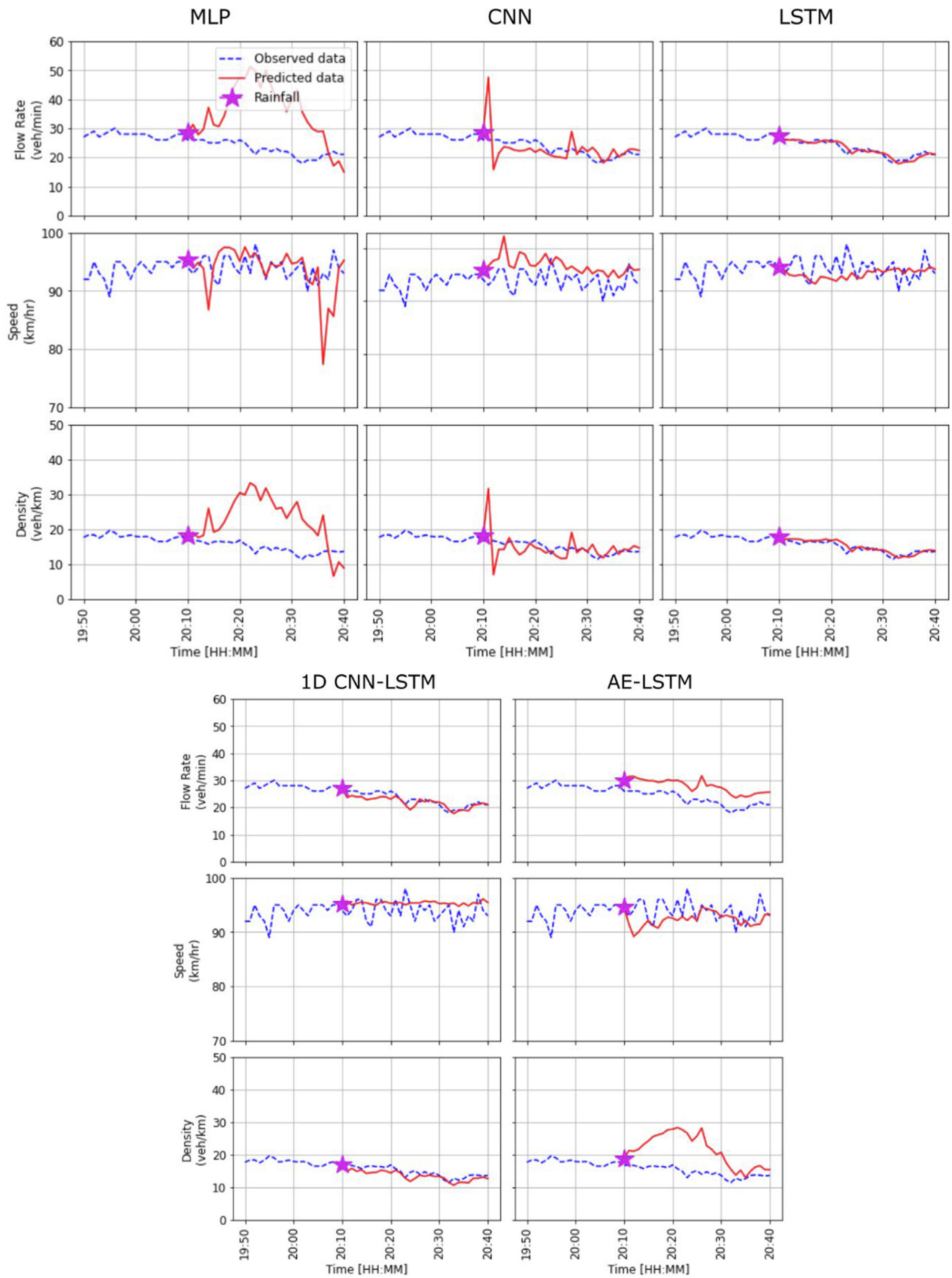


Fig. 11 Short-term predictions under the rain on 4 September 2018 between 20:10 and 20:40 obtained from five prediction models based on various ML networks: the MLP, the CNN, the LSTM, the 1D CNN-LSTM and the Autoencoder LSTM networks.

Table 3 RSMEs and MAEs of ML models' performance for short-term prediction under **the rain**.

			RMSE			
Parameter	**Baseline**	**MLP**	**CNN**	**LSTM**	**1D CNN-LSTM**	**AE-LSTM**
Flow rate	9.65	4.03	3.30	1.23	1.18	1.78
Speed	14.26	5.30	4.10	3.80	3.45	3.50
Density	10.38	7.09	4.55	1.26	1.15	3.75
			MAE			
Parameter	**Baseline**	**MLP**	**CNN**	**LSTM**	**1D CNN-LSTM**	**AE-LSTM**
Flow rate	5.93	3.38	2.79	1.78	1.67	1.81
Speed	7.14	3.15	1.89	1.82	1.73	1.79
Density	5.09	3.65	2.80	2.09	1.87	3.23

6. Conclusions

Using an input data with large observations and five features, the multivariate prediction models based on the Multilayer Perceptron (MLP), One-dimensional Convolutional Neural Network (1-D CNN), the Long Short-term Memory (LSTM) network, 1D-CNN LSTM and Autoencoder LSTM networks have been developed to predict freeway traffic under non-recurrent events. The data features include the flow rate, speed, density, the boolean incident and rainfall. From the results obtained from our proposed multivariate prediction models, we can conclude that:

- The proposed models capture traffic pattern under non-recurrent events. Few discrepancies have been observed in the traffic flow rates between the predicted and observed values. The difference is noticeable when there was an incident.
- The 1D-CNN LSTM prediction model for the density, flow rate, and speed gives more accurate results than those obtained from other ML models.

In the off-ramp and on-ramp areas, more delays and disruption occur, further research will look at the congestion that happens due to the lane change and non-recurrent events. For this purpose, the traffic flow characteristics will be predicted by the time-delay deep neural network model. A deep traffic congestion model will be developed to predict the bottleneck in the traffic flow. This will help to predict the congestion propagation for the targeted routes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research is supported by the Australian Research Council through the linkage project funding scheme (project LP170100341). The authors would like to thank the Main Roads Western Australia (MRWA) for providing traffic data for this research, and also would like to acknowledge the Road and Maritime Service NSW and the Sustainable Environment National Research Centre for their support of the project. The first author would like to thank Prince Sattam bin Abdulaziz University for its financial support.

References

[1] C. Chen, K. Li, S.G. Teo, X. Zou, K. Li, Z. Zeng, Citywide traffic flow prediction based on multiple gated spatio-temporal convolutional neural networks, ACM Trans. Knowl. Discov. from Data (TKDD) 14 (4) (2020) 1–23.

[2] P. Sun, A. Boukerche, Y. Tao, Ssgru: A novel hybrid stacked gru-based traffic volume prediction approach in a road network, Comput. Commun. 160 (2020) 502–511.

[3] W.H. Organization et al., "Global status report on road safety 2018: summary," tech. rep., World Health Organization, 2018.

[4] T. Makaba, W. Doorsamy, B.S. Paul, Exploratory framework for analysing road traffic accident data with validation on gauteng province data, Cogent Eng. 7 (1) (2020) 1834659.

[5] M. Gerla, E.-K. Lee, G. Pau, U. Lee, Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds, in: 2014 IEEE world forum on internet of things (WF-IoT), IEEE, 2014, pp. 241–246.

[6] B. Vijayalakshmi, K. Ramar, N. Jhanjhi, S. Verma, M. Kaliappan, K. Vijayalakshmi, S. Vimal, U. Ghosh, An attention-based deep learning model for traffic flow prediction using spatiotemporal features towards sustainable smart city, Int. J. Commun Syst 34 (3) (2021) e4609.

[7] Y. Lv, Y. Duan, W. Kang, Z. Li, F.-Y. Wang, Traffic flow prediction with big data: a deep learning approach, IEEE Trans. Intell. Transp. Syst. 16 (2) (2014) 865–873.

[8] H. Zheng, F. Lin, X. Feng, Y. Chen, A hybrid deep learning model with attention-based conv-lstm networks for short-term traffic flow prediction, IEEE Trans. Intell. Transp. Syst. 22 (11) (2020) 6910–6920.

[9] M. Van Der Voort, M. Dougherty, S. Watson, Combining kohonen maps with arima time series models to forecast traffic flow, Transportation Research Part C: Emerging Technologies 4 (5) (1996) 307–318.

[10] S. Lee, D.B. Fambro, Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting, Transp. Res. Rec. 1678 (1) (1999) 179–188.

[11] B.M. Williams, Multivariate vehicular traffic flow prediction: evaluation of arimax modeling, Transp. Res. Rec. 1776 (1) (2001) 194–200.

[12] B.M. Williams, L.A. Hoel, Modeling and forecasting vehicular traffic flow as a seasonal arima process: Theoretical basis and empirical results, Journal of transportation engineering 129 (6) (2003) 664–672.

[13] K. Chen, F. Chen, B. Lai, Z. Jin, Y. Liu, K. Li, L. Wei, P. Wang, Y. Tang, J. Huang, et al, Dynamic spatio-temporal graph-based cnns for traffic flow prediction, IEEE Access 8 (2020) 185136–185145.

[14] B.L. Smith and M.J. Demetsky, "Short-term traffic flow prediction models-a comparison of neural network and nonparametric regression approaches," in Proceedings of IEEE international conference on systems, man and cybernetics, vol. 2, pp. 1706–1709, IEEE, 1994.

[15] I.J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, and V. Shet, "Multi-digit number recognition from street view imagery using deep convolutional neural networks," arXiv preprint arXiv:1312.6082, 2013.

[16] A. Graves, "Speech recognition with deep recurrent neural networks alex graves, abdel-rahman mohamed and geoffrey hinton department of computer science, university of toronto. ieee international conference 3 (2013) 6645–6649.

[17] T.N. Sainath, O. Vinyals, A. Senior, H. Sak, Convolutional, long short-term memory, fully connected deep neural networks, in: in 2015 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, 2015, pp. 4580–4584.

[18] F.B. Fitch, "Mcculloch warren s. and pitts walter. a logical calculus of the ideas immanent in nervous activity. bulletin of mathematical biophysics, vol. 5, pp. 115–133," Journal of Symbolic Logic, vol. 9, no. 2, 1944.

[19] F. Rosenblatt, Perceptions and the theory of brain mechanisms, Spartan books, 1962.

[20] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning internal representations by error propagation, California Univ San Diego La Jolla Inst for Cognitive Science, 1985, tech. rep..

[21] S. Albawi, T.A. Mohammed, S. Al-Zawi, Understanding of a convolutional neural network, in: in 2017 International Conference on Engineering and Technology (ICET), Ieee, 2017, pp. 1–6.

[22] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, et al, Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups, IEEE Signal processing magazine 29 (6) (2012) 82–97.

[23] J.P.A. Vieira, R.S. Moura, An analysis of convolutional neural networks for sentence classification, in: in 2017 XLIII Latin American Computer Conference (CLEI), IEEE, 2017, pp. 1–5.

[24] D. Kang, Y.-J. Cha, Autonomous uavs for structural health monitoring using deep learning and an ultrasonic beacon system with geo-tagging, Computer-Aided Civil and Infrastructure Engineering 33 (10) (2018) 885–902.

[25] D.H. Hubel, T.N. Wiesel, Receptive fields and functional architecture of monkey striate cortex, The Journal of physiology 195 (1) (1968) 215–243.

[26] K. Fukushima, S. Miyake, Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition, in: Competition and cooperation in neural nets, Springer, 1982, pp. 267–285.

[27] M. Matsugu, K. Mori, Y. Mitari, Y. Kaneda, Subject independent facial expression recognition with robust face detection using a convolutional neural network, Neural Networks 16 (5–6) (2003) 555–559.

[28] S. Deng, S. Jia, J. Chen, Exploring spatial–temporal relations via deep convolutional neural networks for traffic flow prediction with incomplete data, Applied Soft Computing 78 (2019) 712–721.

[29] S. Sun, H. Wu, L. Xiang, City-wide traffic flow forecasting using a deep convolutional neural network, Sensors 20 (2) (2020) 421.

[30] V. Osipov, V. Nikiforov, N. Zhukova, D. Miloserdov, Urban traffic flows forecasting by recurrent neural networks with spiral structures of layers, Neural Comput. Appl. 32 (18) (2020) 14885–14897.

[31] B. Yang, S. Sun, J. Li, X. Lin, Y. Tian, Traffic flow prediction using lstm with feature enhancement, Neurocomputing 332 (2019) 320–327.

[32] B. Bae, H. Kim, H. Lim, Y. Liu, L.D. Han, P.B. Freeze, Missing data imputation for traffic flow speed using spatio-temporal cokriging, Transportation Research Part C: Emerging Technologies 88 (2018) 124–139.

[33] Q. Li, H. Tan, Y. Wu, L. Ye, F. Ding, Traffic flow prediction with missing data imputed by tensor completion methods, IEEE Access 8 (2020) 63188–63201.

[34] B. Feng, J. Xu, Y. Lin, P. Li, A period-specific combined traffic flow prediction based on travel speed clustering, IEEE Access 8 (2020) 85880–85889.

[35] P. Sun, N. Aljeri, A. Boukerche, Machine learning-based models for real-time traffic flow prediction in vehicular networks, IEEE Network 34 (3) (2020) 178–185.

[36] J. Wang, A. Boukerche, The scalability analysis of machine learning based models in road traffic flow prediction, in: ICC 2020–2020 IEEE International Conference on Communications (ICC), IEEE, 2020, pp. 1–6.

[37] A.A. Kashyap, S. Raviraj, A. Devarakonda, S.R. Nayak K, S. KV, and S.J. Bhat, "Traffic flow prediction models–a review of deep learning techniques," Cogent Engineering, vol. 9, no. 1, p. 2010510, 2022.

[38] J. Bhatia, R. Dave, H. Bhayani, S. Tanwar, A. Nayyar, Sdn-based real-time urban traffic analysis in vanet environment, Comput. Commun. 149 (2020) 162–175.

[39] E. Choi, A. Schuetz, W.F. Stewart, J. Sun, Using recurrent neural network models for early detection of heart failure onset, J. Am. Med. Inform. Assoc. 24 (2) (2017) 361–370.

[40] W. Kong, Z.Y. Dong, Y. Jia, D.J. Hill, Y. Xu, Y. Zhang, Short-term residential load forecasting based on lstm recurrent neural network, IEEE Transactions on Smart Grid 10 (1) (2017) 841–851.

[41] Y. Jiang, K. Zhao, K. Xia, J. Xue, L. Zhou, Y. Ding, P. Qian, A novel distributed multitask fuzzy clustering algorithm for automatic mr brain image segmentation, Journal of medical systems 43 (5) (2019) 1–9.

[42] Z. Liu, Z. Li, K. Wu, M. Li, Urban traffic prediction from mobility data using deep learning, Ieee network 32 (4) (2018) 40–46.

[43] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (8) (1997) 1735–1780.

[44] S. Hochreiter, J. Schmidhuber, Lstm can solve hard long time lag problems, Advances in neural information processing systems 9 (1996).

[45] F.A. Gers, J. Schmidhuber, F. Cummins, Learning to forget: Continual prediction with lstm, Neural computation 12 (10) (2000) 2451–2471.

[46] K.B. Lee, H.S. Shin, An application of a deep learning algorithm for automatic detection of unexpected accidents under bad cctv monitoring conditions in tunnels, in: in 2019 International Conference on deep learning and machine learning in emerging applications (Deep-ML), IEEE, 2019, pp. 7–11.

[47] F. Aljuaydi, B. Wiwatanapataphee, Y.H. Wu, Deep learning-based prediction models for freeway traffic flow under non-recurrent events, in 2022 8th International Conference on Control, Decision and Information Technologies (CoDIT), vol. 1, IEEE, 2022, pp. 815–820.

[48] D.P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.