WILEY | Hindawi

Research Article

Traffic Flow Prediction with Rainfall Impact Using a Deep Learning Method

Yuhan Jia,[1,2] **Jianping Wu,**[1] **and Ming Xu**[1]

[1]*Department of Civil Engineering, Tsinghua University, Beijing 100084, China*
[2]*Investment Banking Department, Profitability Units, Head Office of Industrial and Commercial Bank of China, Beijing 100140, China*

Correspondence should be addressed to Ming Xu; mxu@tsinghua.edu.cn

Accurate traffic flow prediction is increasingly essential for successful traffic modeling, operation, and management. Traditional data driven traffic flow prediction approaches have largely assumed restrictive (shallow) model architectures and do not leverage the large amount of environmental data available. Inspired by deep learning methods with more complex model architectures and effective data mining capabilities, this paper introduces the deep belief network (DBN) and long short-term memory (LSTM) to predict urban traffic flow considering the impact of rainfall. The rainfall-integrated DBN and LSTM can learn the features of traffic flow under various rainfall scenarios. Experimental results indicate that, with the consideration of additional rainfall factor, the deep learning predictors have better accuracy than existing predictors and also yield improvements over the original deep learning models without rainfall input. Furthermore, the LSTM can outperform the DBN to capture the time series characteristics of traffic flow data.

1. Introduction

Traffic flow prediction is an important component of traffic modeling, operation, and management. Accurate real-time traffic flow prediction can (1) provide information and guidance for road users to optimize their travel decisions and to reduce costs and (2) help authorities with advanced traffic management strategies to mitigate congestion. With the availability of high resolution traffic data from intelligent transportation systems (ITS), traffic flow prediction has been increasingly addressed using data driven approaches [1]. In this regard, traffic flow prediction is a time series problem to estimate the flow count at a future time based on the data collected over previous periods from one or more observation locations. The prediction models in the literature can be broadly divided into parametric and nonparametric models.

For parametric analysis, linear time series models have been widely applied. An early study in 1979 investigated the application of the autoregressive integrated moving-average (ARIMA) model for short-term traffic data forecasts [2]. The results suggest that the ARIMA model could be used for one-time interval prediction. Levin and Tsao proposed

that ARIMA (0, 1, 1) is the most statistically significant for both local and express lanes prediction [3]. Since then many scientists have focused on the ARIMA model and its extensions. A hybrid method of combined Kohonen maps and ARIMA model was developed to forecast traffic flow at horizons of half an hour and one hour, and its performance was found to be superior to ARIMA [4]. In addition, many other modified ARIMA models have been examined, such as subset ARIMA [5], ARIMA with explanatory variables [6], and seasonal ARIMA [7]. Another widely used method to predict traffic volume is the filtering approach, such as the Kalman Filter model. Okutani and Stephanedes introduced Kalman filtering theory into this field and results indicated improved performances [8]. Based on Kalman filter theory, there have also been some modifications [9] and hybrid models [10].

Due to the stochastic and nonlinear characteristics of traffic flow, it is difficult to overcome the limitations of parametric models. Nonparametric machine learning methods have become increasingly popular. Artificial neural networks (NN) have been commonly utilized for this problem, which can be regarded as the general pattern of machine learning

application in traffic engineering [11]. Smith and Demetsky developed a NN model which was compared with traditional prediction methods and their results suggest that the NN outperforms other models during peak conditions [12]. Dougherty et al. studied back-propagation neural network (BPNN) for the prediction of traffic flow, speed, and occupancy and the results show some promise [13]. Since then, NN approaches have commonly been used for traffic flow forecasting [14–16]. In addition, many hybrid NN models have been proposed to improve performance [17–20]. Other nonparametric models have also been studied, such as k-nearest neighbor (k-NN) models [21] and support vector regression [22].

With the fast development of ITS, it is possible to get access to huge amount of traffic data as well as multisource environmental data. However, both the typical parametric and nonparametric models tend to make assumptions to ignore additional influencing factors, due to the shallow architecture and inability to deal with big data as well as inefficient training strategies. The deep learning method, a type of machine learning, has been well developed recently to address this problem. Deep learning has shown its advantages in pattern recognition [23, 24] and classification [25, 26]. For traffic flow prediction, Huang et al. proposed a deep belief network (DBN) architecture for multitask learning and experiments results show that the DBN could achieve about 5% improvements over the state of the art [27]. Another study used a stacked autoencoder model (SAE) to implement prediction considering explicitly the spatial and temporal correlations [28]. These papers are the pioneers in applying deep learning in traffic flow prediction.

Furthermore, to capture the time series characteristics in the training and prediction steps, another branch of the deep learning models is introduced as the recurrent neural network (RNN), which is designed to cope with time series data prediction problem [29, 30]. In terms of traffic data, a kind of time series data, the RNN can use memory cells to save the temporal information from previous time intervals [31, 32]. One of the most famous RNN is the long short-term memory (LSTM) model, which can automatically adjust some hyperparameters and can capture the long temporal features of the input data [33]. The LSTM has been introduced into the traffic information prediction field and experiments have shown some promising results [34, 35]. The DBN and LSTM will be introduced in detail to represent the basic deep learning method and the advanced recurrent NN, respectively.

However, in most studies, the predictions have not considered weather factors as a multisource input. Whether the deep learning can improve prediction accuracy based on more realistic conditions with additional data sources, like rainfall intensity, is an important issue that warrants investigation.

With regard to the impact of rainfall, there is general consensus that it significantly affects traffic flow characteristics and leads to congestion and accidents. Without a comprehensive understanding of the weather influence on traffic flow, traffic management authorities cannot consider relevant factors in related operational policies to improve

traffic efficiency and safety [36]. For rainfall-integrated traffic flow prediction using machine learning methods, Dunne and Ghosh combined stationary wavelet transform and BPNN to develop a predictor that could choose between a dry model and a wet model depending on whether rainfall is expected in the prediction hour [37]. The results show that rainfall-integrated predictor could improve the prediction accuracy during rainfall events. Deep learning tools provide a promising avenue to incorporate the impacts of rainfall in traffic flow prediction.

In this paper, a systematic procedure to predict traffic flow considering rainfall impact using the deep learning method is presented. The main contribution of this work is to investigate whether the deep learning models can outperform typical methods with multisource data. This study also aims to compare the difference between advanced recurrent NN and basic deep learning NN for the prediction of time series traffic flow. In addition, the temporal and spatial characteristics can also be included with multiply sets of traffic detectors. In this regard, the features of traffic flow under various rainfall conditions can be learned after training using local arterial traffic and rainfall intensity data from Beijing, China. The methodology and results can facilitate improving the traffic modeling, operation, and management performance. The findings of this paper can be significant by investigating the prediction accuracy with additional nontraffic data and can be regarded as reference for further consideration of more environmental factors.

Section 2 reviews the literature relating to the DBN, LSTM, and rainfall impact analysis. Section 3 develops the rainfall-integrated DBN (R-DBN) and the rainfall-integrated LSTM (R-LSTM) models with multisource data input considering the additional rainfall intensity. In Section 4, experiments are presented and the results are discussed. Finally, Section 5 provides a summary, concluding remarks, and future research directions.

2. Research Background

2.1. Deep Belief Network.
The DBN is a type of deep neural network with many hidden layers and large amounts of hidden units in each layer. The typical DBN is equivalent to a stack of Restricted Boltzmann Machine (RBM) models with an output layer on the top. The DBN uses a fast, greedy unsupervised learning algorithm to train RMBs and a supervised fine-tuning method to adjust the network by labeled data [38].

Each RBM has a visible layer **v** and a hidden layer **h**, connected by undirected weights. For the stack of RBMs in the DBN, the hidden layer of one RBM is regarded as the visible layer of the next RBM. The parameter set of a RMB is $\boldsymbol{\theta} = (\mathbf{w}, \mathbf{b}, \mathbf{a})$, where w_{ij} is the weight between v_i and h_j. b_i and a_j are the bias of the layers. A RBM defines its energy as

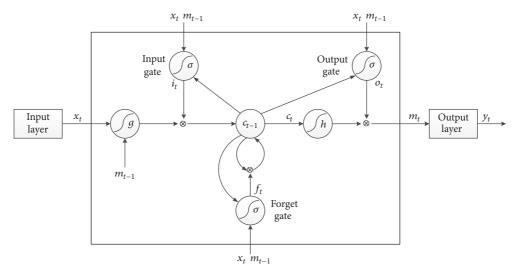$$E(\mathbf{v}, \mathbf{h} \mid \boldsymbol{\theta}) = -\sum_i b_i v_i - \sum_j a_j h_j - \sum_i \sum_j w_{ij} v_i h_j \qquad (1)$$

FIGURE 1: The structure of the LSTM memory block.

and its joint probability distribution of **v** and **h** as

$$p\left(\mathbf{v}, \mathbf{h} \mid \boldsymbol{\theta}\right) = \frac{\exp\left(-E\left(\mathbf{v}, \mathbf{h} \mid \boldsymbol{\theta}\right)\right)}{\sum_{\mathbf{v},\mathbf{h}} \exp\left(-E\left(\mathbf{v}, \mathbf{h} \mid \boldsymbol{\theta}\right)\right)} \tag{2}$$

and the marginal probability distribution of **v** as

$$p\left(\mathbf{v} \mid \boldsymbol{\theta}\right) = \frac{\sum_{\mathbf{h}} \exp\left(-E\left(\mathbf{v}, \mathbf{h} \mid \boldsymbol{\theta}\right)\right)}{\sum_{\mathbf{v},\mathbf{h}} \exp\left(-E\left(\mathbf{v}, \mathbf{h} \mid \boldsymbol{\theta}\right)\right)}. \tag{3}$$

To obtain the optimal $\boldsymbol{\theta}$ for a single data vector **v**, the gradient of log-likelihood estimation can be calculated based on equation, given as

$$\frac{\partial \log p\left(\mathbf{v} \mid \boldsymbol{\theta}\right)}{\partial w_{ij}} = \left\langle v_i h_j \right\rangle_{\text{data}} - \left\langle v_i h_j \right\rangle_{\text{model}},$$

$$\frac{\partial \log p\left(\mathbf{v} \mid \boldsymbol{\theta}\right)}{\partial a_j} = \left\langle h_j \right\rangle_{\text{data}} - \left\langle h_j \right\rangle_{\text{model}}, \tag{4}$$

$$\frac{\partial \log p\left(\mathbf{v} \mid \boldsymbol{\theta}\right)}{\partial b_i} = \left\langle v_i \right\rangle_{\text{data}} - \left\langle v_i \right\rangle_{\text{model}},$$

where $\langle \cdot \rangle$ denotes the expectations under the distribution specified by the subscript that follows.

Because there are no connections between the units in the same layer, $\langle \cdot \rangle_{\text{data}}$ is easy to obtain by calculating the conditional probability distributions, given as

$$p\left(h_j \mid \mathbf{v}, \boldsymbol{\theta}\right) = \frac{1}{1 + \exp\left(-\sum_i w_{ij} v_i - a_j\right)},$$

$$p\left(v_i \mid \mathbf{h}, \boldsymbol{\theta}\right) = \frac{1}{1 + \exp\left(-\sum_j w_{ij} h_j - b_i\right)}. \tag{5}$$

The activation function is sigmoid function.

For $\langle \cdot \rangle_{\text{model}}$ the contrastive divergence (CD) learning method could be utilized by reconstruction to minimize the difference of two Kullback-Leibler divergences (KL). CD learning is proved to be efficiently practical and also can reduce computational cost compared with typical Gibbs sampling method [38–40].

The weights in DBN layers are trained by unlabeled data using the aforementioned fast and greedy unsupervised method. For prediction purpose, a supervised layer should be added above the DBN to adjust the learned features by labeled data using an up-down fine-tuning algorithm. In this study, the fully connected layer is selected to perform as the top layer, also using the sigmoid activation function.

2.2. Long Short-Term Memory. The RNN models can take the time series characteristics into consideration, using a recurrent mechanism to save the temporal and spatial states of previous intervals. However, some early models, such as time-delay neural network (TDNN) [31], fail to describe the long-term correlation of data. The training and application steps of RNN to predict traffic information usually contain data with long time lags. In this regard, the LSTM model is developed to overcome this shortcoming.

The success of LSTM relies on the ability to learn the time series characteristics and to determine several hyperparameters automatically from data [33–35]. The LSTM model is composed of one input layer, one hidden layer, and one output layer. The hidden layer is regarded as a memory block with memory cells to memorize the long and short temporal features. Three gates are developed, including the input gate, output gate, and forget gate, to control the input, output, and continual time series characteristics, respectively. The state of the memory cell is recorded and recurrently self-connected, as shown in Figure 1. The input data, output data, and memory information of hidden layer at interval t are denoted as x_t, y_t, and m_t, respectively. The i_t, o_t, f_t, and c_t represent the output of input gate, output gate, forget gate, and the memory cell state, respectively.

The LSTM model can be conducted by the equations shown below.

$$
\begin{aligned}
i_t &= \sigma\left(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i\right), \\
f_t &= \sigma\left(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f\right), \\
c_t &= f_t \otimes c_{t-1} + i_t \otimes g\left(W_{cx}x_t + W_{cm}m_{t-1} + b_c\right), \\
o_t &= \sigma\left(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o\right), \\
m_t &= o_t \otimes h\left(c_t\right), \\
y_t &= W_{ym}m_t + b_y,
\end{aligned}
\tag{6}
$$

where $\sigma(\cdot)$ means the sigmoid activation function, $g(\cdot)$ and $h(\cdot)$ are the centered logistic sigmoid functions, W means the weigh matrices, and b means the bias vectors.

Unlike DBN training, labeled data is utilized to train the neural network of LSTM in a BP manner using the gradient descent optimization method. Due to the mathematical framework of recurrent mechanism, the time series information can be learned by the memory cell in the LSTM without specifying the time lags from previous intervals [34].

2.3. Rainfall Impact. The reason we chose to consider rainfall intensity as the additional data source is that there is a general consensus that it significantly affects traffic flow characteristics. Conclusions from an early study implied that rainfall weather has negative impacts on the traffic flow system [41]. Since then, many papers have focused on studying the rainfall impact on principal road traffic characteristics such as capacity and operating speed. It is noted that both road capacity and vehicle speed will influence the traffic flow count. Thus, it is reasonable to consider this weather factor when implementing prediction using a deep learning method.

Previous studies examining the influence of weather have been limited to the lack of data. However, in recent years with the access to high resolution data from traffic agencies and weather stations, an increasing number of investigations have focused on the quantitative rainfall impacts on the traffic system. The speed-flow-occupancy relationship under adverse weather conditions was examined by using dummy variables within a multiple regression model and the results of capacity loss are marginal in light rain but about 15% in heavy rain are indicated [42]. A study estimated the rainfall impact on freeway road capacity and operating speed under different rainfall categories and found that the capacity reductions in heavy rain are greater than that recommended by the HCM2000 [43, 44]. More specifically, an investigation found that rainfall conditions of 0–0.01 in/h, 0.01–0.25 in/h, and >0/25 in/h cause 2%, 7%, and 14% capacity reductions and 2%, 4%, and 6% speed reductions, respectively [45]. For urban arterial flow count reductions, using the data from the Tokyo Metropolitan Expressway, a study found that rainfall decreases flow ranging from 4–7% in light rain to a maximum of 14% during heavy rain [46]. For Belgium, a paper focused on the effect of weather conditions on daily traffic flow. The heterogeneity of the weather effects

between different traffic count locations and the homogeneity of the weather effects on upstream and downstream traffic at specific locations were observed [47]. More recent works have focused on the development and calibration of generalized rainfall impact models, which have yielded good results for traffic management and simulation systems [48, 49].

3. Methodology

For accurate traffic flow prediction, we propose two deep learning methods to learn the effective features of traffic flow and rainfall data. From a previous study based on local data in Beijing, it was concluded that rainfall weather significantly affects the traffic flow characteristics, including road capacity and operating speed [50]. The reduced road capacity and slower vehicle speed yield the differences of flow count between different weather conditions. In other words, despite other factors remaining the same except for rainfall condition, the traffic flow counts will differ. In most previous studies, this factor was not included in the input vector, resulting in inappropriate features learned by predictors and inaccurate prediction outcomes.

In this study, together with the traffic flow from previous intervals, the rainfall intensity data for corresponding previous intervals is added to both the training and the testing data set. In the training stage, the feature of rainfall impact on traffic flow can be learned to reflect the reductions of flow under various rainfall conditions. Consequently, in the test stage, the model will have better prediction compared to those that neglect this weather feature [37].

Based on the input layer, a stack of RBMs are developed to extract the features of data using a fast, greedy unsupervised learning method. The hidden layer of RBM is regarded as the visible layer of the next RBM. Each RBM is trained separately using unlabeled data by the CD learning method to reconstruct the input. A fully connected output layer is developed on top of the RBMs to fine-tune the entire architecture using labeled data by an up-down algorithm. Unlike typical pattern recognition or classification task, the traffic flow data in this study are from one arterial, collected by four sequent detector sets which have spatial-temporal relationships. Thus, it is feasible to put the four related prediction tasks together to share the trained features and weights to avoid unreasonable bias. This rainfall-integrated DBN is denoted as R-DBN and its structure is shown in Figure 2.

Similar to the structure of R-DBN, the input data of R-LSTM also contains the rainfall data for corresponding previous intervals with the flow data. The model has one hidden layer as the memory cell to recurrently save the time series characteristics of the traffic data. The network can be trained by labeled data in a BP manner using the gradient descent optimization method. Similarly, the structure of R-SLTM is shown in Figure 3.

The deep learning methodology adopted in this paper is as follows.

Step 1. Use traffic data and rainfall data to prepare training and test data set for 10-minute and 30-minute prediction.
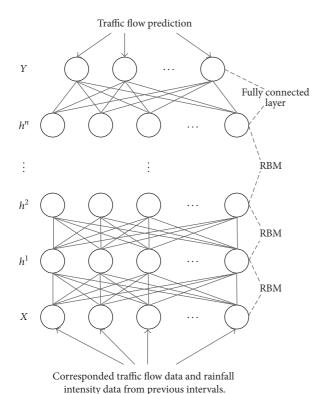
FIGURE 2: The structure of the R-DBN predictor.

Step 2. Use architecture optimization data set to decide optimal R-DBN architecture parameters, including input dimension, layer size, hidden units size per layer, and epochs.

Step 3. Use training data set to train the R-DBN and R-LSTM based on optimal architectures.

Step 4. Use test data set to examine prediction performance. Compare the results with rainfall-integrated BPNN (R-BPNN) and rainfall-integrated ARIMA (R-ARIMA). The regular DBN, LSTM, BPNN, and ARIMA without rainfall consideration are also considered as benchmarks.

4. Experiments

4.1. Data Description. An arterial segment from Deshengmen to Madian connecting the 2nd and 3rd Ring Road in Beijing, China, is selected as the test objective. The graph of the test area is shown in Figure 4, with three lanes for each direction. This 2.1-kilometer long arterial has the commonality to represent the corridors in Beijing, which ensures the model to be transferable to other areas. Four sets of detectors are distributed in both directions. Traffic flow count data are recorded in 2-minute intervals and archived by the Beijing Traffic Management Bureau (BTMB). For weather data in the corresponding urban area, hourly rainfall intensity is recorded by the National Meteorological Center (NMC) of the China Meteorological Administration (CMA). The flow count data collected by detectors from three lanes are aggregated as one data point. The study period is from June

to August, 2013, and June to August, 2014. The data from July 8 to July 14 of 2013 are selected as the test set. The 2013 June data are used for model architecture optimization. And the rest of the 2013 and 2014 data are regarded as the training set. Most of the available data is utilized to train the models to ensure accuracy. The test week is selected as it contains four rainfall days, 8th, 9th, 10th, and 14th. It is noted that this study only uses one direction of the arterial as an experiment demonstration. The aggregated traffic flow count data for the four detector sets of the test week are shown in Figure 5. The horizontal axis is time in 10-minute unit and the vertical axis is traffic flow count. The patterns of daily traffic flow count can be observed.

To evaluate the performance of the predictor, three performance measurements are selected, which are the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the root mean square error (RMSE), shown below.

$$
\begin{aligned}
\text{MAE} &= \frac{1}{N} \sum_{i=1}^{N} |x_i - y_i|, \\
\text{MAPE} &= \frac{1}{N} \sum_{i=1}^{N} \left| \frac{x_i - y_i}{x_i} \right|, \\
\text{RMSE} &= \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2},
\end{aligned}
\tag{7}
$$

where x_i and y_i are the observed and predicted flow data for interval i and N is the test sample size. The units of MAE and RMSE are both vehicle per hour (veh/h).

4.2. Prediction Architecture. The architecture of the DBN predictor depends on several principal parameters, including the number of previous time intervals utilized for prediction, the hidden layer size, the units in each hidden layer, and the number of epochs for training a RBM. To decide the model architecture, the previous time intervals for prediction are set from 1 to 15 (2 minutes to 30 minutes), and the hidden layer size is from 1 to 10. The units in each RBM layer are chosen from 100 to 1000 with gap value 100, and the epochs are set from 10 to 50 with gap value of 10. In this paper, the DBN predictor is utilized to predict the next 10-minute and 30-minute traffic flow. For each of the three prediction horizons, grid search has been implemented to decide the most effective architecture for predictor based on the training set. The parameter choices of R-DBN and DBN are listed in Table 1.

From Table 1, it is surprisingly observed that for all the cases the optimal hidden layer size should not exceed 3. In other words, a DBN with three RBM is sufficient for Beijing traffic flow analysis. For previous data set intervals, it is noted that more intervals are needed to predict traffic flow with longer horizon. For example, more than 20 previous intervals are required to implement a 30-minute prediction, while only 12 or 14 intervals are needed for a 10-minute horizon. Regarding the units in each hidden layer, the number should be neither too small nor too large. Furthermore, the epoch time ranges from 10 to 40. Using a grid search, the optimal
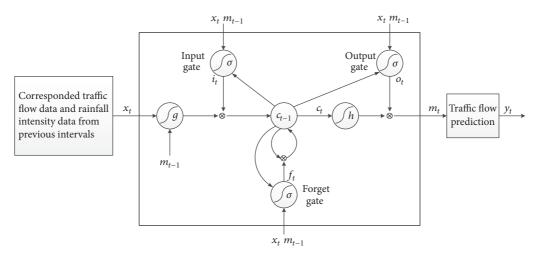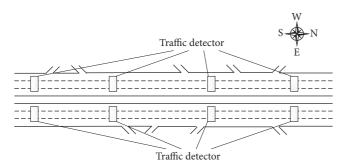
Figure 3: The structure of the R-LSTM.

Figure 4: Graph of the experiment area.

Table 1: The most effective architecture.

	Previous intervals	Hidden layer	Hidden units	Epochs
10-minute R-DBN	12	2	300	40
10-minute DBN	14	2	400	20
30-minute R-DBN	25	2	400	10
30-minute DBN	24	3	300	10

DBN architecture is determined for arterial traffic flow prediction in Beijing. The results could provide references for future research and application.

In terms of the R-LSTM and LSTM, only one hidden layer is utilized in this paper, and the optimal previous time intervals can be decided automatically. Thus, it is unnecessary to consider the determination of the hyperparameters.

4.3. Result Discussion. The test data set is from July 8 to July 14, which as noted earlier includes four rainy days. For comparison, several other predictors are selected, including R-BPNN and R-ARIMA. The regular DBN, BPNN, and ARIMA without rainfall consideration are also implemented as benchmarks. For BPNN and ARIMA, grid search is also implemented due to the structure uncertainty with the additional rainfall input. The hidden layer number for BPNN

Table 2: Comparisons between different predictors.

Model	Measurement	10-minute prediction	30-minute prediction
R-DBN	MAE (veh/h)	178.90	325.58
	MAPE (%)	13.14	19.75
	RMSE (veh/h)	255.79	356.49
R-LSTM	MAE (veh/h)	166.17	305.04
	MAPE (%)	11.69	17.88
	RMSE (veh/h)	240.98	296.91
R-BPNN	MAE (veh/h)	192.70	337.34
	MAPE (%)	15.59	21.80
	RMSE (veh/h)	298.22	377.86
R-ARIMA	MAE (veh/h)	236.24	409.75
	MAPE (%)	18.47	28.14
	RMSE (veh/h)	331.82	447.06
DBN	MAE (veh/h)	186.15	327.06
	MAPE (%)	13.58	20.19
	RMSE (veh/h)	265.16	374.33
LSTM	MAE (veh/h)	172.63	315.15
	MAPE (%)	12.84	19.55
	RMSE (veh/h)	255.04	346.83
BPNN	MAE (veh/h)	200.44	352.21
	MAPE (%)	16.76	22.81
	RMSE (veh/h)	302.74	375.32
ARIMA	MAE (veh/h)	227.85	369.54
	MAPE (%)	17.37	25.95
	RMSE (veh/h)	313.65	413.86

is 1, and the units are 400 and 400 with rainfall consideration and 400 and 300 without rainfall consideration, for 10-minute and 30-minute horizons, respectively. For ARIMA, the best structure is ARIMA (1, 1, 2) with rainfall input and ARIMA (1, 2, 3) without rainfall input.

The MAE, MAPE, and RMSE between observed and predicted data are presented in Table 2. For all the predictors,

(a) Traffic detector set number 1

(b) Traffic detector set number 2

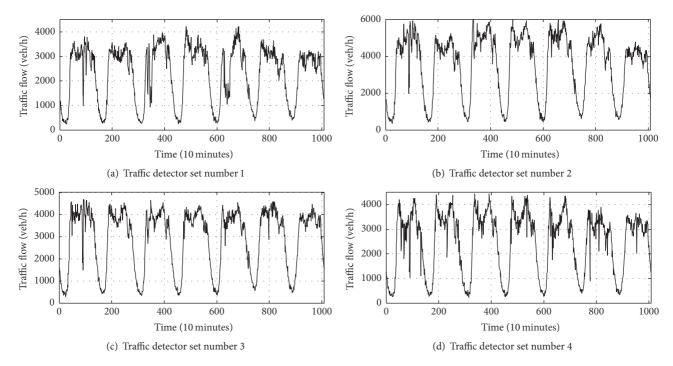(c) Traffic detector set number 3

(d) Traffic detector set number 4

FIGURE 5: Traffic flow count data of test week for the four detector sets.

performance results show that it works better for a 10-minute horizon than 30-minute horizon, for all the measurements. This finding is expected and consistent with other studies wherein larger errors are observed for longer prediction horizon. In comparison with the predictors without rainfall input, the R-DBN and R-LSTM perform better for both time horizons, indicating that the accuracy of deep learning could be improved with multisource data inputs. This emphasizes the significance of using the rainfall-integrated model for traffic flow prediction. The rainfall feature can be obtained by deep learning and then utilized for forecasting, indicating that the R-DBN and R-LSTM have the ability to handle multisource data and learn the underlying patterns of traffic flow under various rainfall conditions. Furthermore, the LSTM family performs better than DBN, with/without considering rainfall factor, which proofs the advantage of LSTM to capture the time series characteristics of traffic data.

For the comparison with other existing methods, we choose BPNN and ARIMA to represent a shallow machine learning model and a typical parametric model, respectively. From Table 2, it is noted that these two rainfall-integrated shallow models are even less effective than the LSTM and DBN without rainfall input. Furthermore, the BPNN has better accuracy than ARIMA whether with or without rainfall consideration. Another interesting thing is that the R-BPNN works better than regular BPNN, while the R-ARIMA is worse than the regular one. This phenomenon indicates that the ability to conduct multisource data is limited for the time series model like ARIMA. This work implies that with additional data such as weather factor, deep learning could also outperform shallow machine learning and typical predictors.

The performance of R-LSTM is shown in Figure 6. The horizontal axis is time in 10-minute unit and the vertical axis is traffic flow count. It can be observed that the predicted data can describe the real situation with/without rainfall influence.

5. Conclusion

In this paper, we try to investigate whether the deep learning model can outperform traditional methods based on traffic flow data considering the rainfall factor. This study also aims to compare the difference between advanced recurrent NN and basic deep learning NN for the prediction of time series traffic data. For this purpose, we propose two deep learning methods, the R-DBN and the R-LSTM models. The DBN model consists of a stack of RBMs, which utilizes unsupervised and supervised training. The LSTM model uses memory block with memory cells to memorize the long and short temporal features, which yields better performance for the prediction of time series data. The data used is from an arterial in Beijing, with rainfall intensity data for the corresponding period. Test experiments show that the R-DBN and R-LSTM can effectively discover the features of traffic flow under various rainfall conditions, and the new models can outperform R-BPNN and R-ARIMA and also yield improvements over the original deep learning models without rainfall consideration. In addition, the LSTM family can outperform the DBN family to describe the time series characteristics of traffic data. In summary, the results highlight the importance of incorporating weather impacts and the significant improvements attainable in prediction accuracy. This work suggests that deep learning is promising with multisource data inputs.

(a) Traffic detector set number 1

(b) Traffic detector set number 2

(c) Traffic detector set number 3

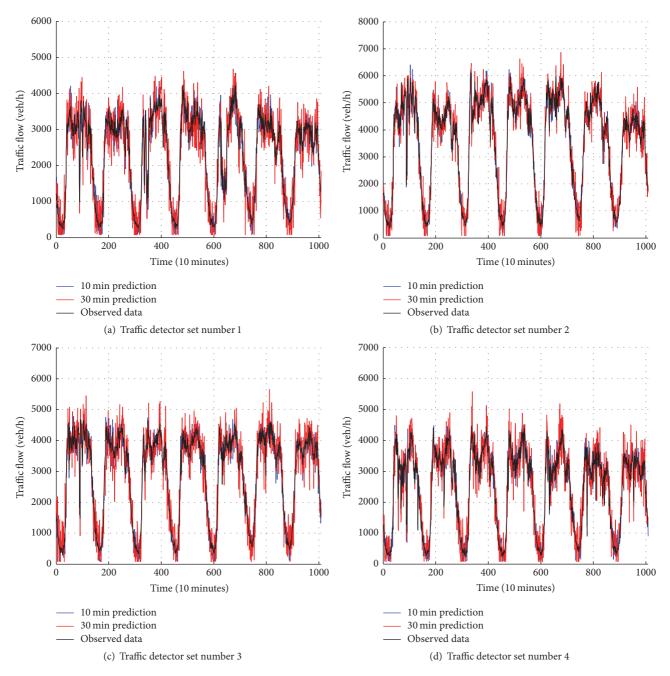(d) Traffic detector set number 4

FIGURE 6: Predicted and observed traffic flow count data for the four detector sets by using R-LSTM.

For future work, more deep learning methods should be introduced. Also, the consideration of more additional factors would be interesting.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

References

[1] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: a survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624–1639, 2011.

[2] M. S. Ahmed and A. R. Cook, "Analysis of freeway traffic time-series data by using box-jenkins techniques," *Transportation Research Record*, no. 722, pp. 1–9, 1979.

[3] M. Levin and Y. D. Tsao, "On forecasting freeway occupancies and volumes," *Transportation Research Record*, vol. 773, pp. 47–49, 1980.

[4] M. van der Voort, M. Dougherty, and S. Watson, "Combining Kohonen maps with ARIMA time series models to forecast traffic flow," *Transportation Research Part C: Emerging Technologies*, vol. 4, no. 5, pp. 307–318, 1996.

[5] S. Lee and D. B. Fambro, "Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting," *Transportation Research Record*, vol. 1678, pp. 179–188, 1999.

[6] B. M. Williams, "Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling," *Transportation Research Record*, no. 1776, pp. 194–200, 2001.

[7] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664–672, 2003.

[8] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transportation Research B*, vol. 18, no. 1, pp. 1–11, 1984.

[9] J. Guo, W. Huang, and B. M. Williams, "Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 50–64, 2014.

[10] S. Jin, D.-H. Wang, C. Xu, and D.-F. Ma, "Short-term traffic safety forecasting using Gaussian mixture model and Kalman filter," *Journal of Zhejiang University: Science A*, vol. 14, no. 4, pp. 231–243, 2013.

[11] M. Dougherty, "A review of neural networks applied to transport," *Transportation Research Part C*, vol. 3, no. 4, pp. 247–260, 1995.

[12] L. B. Smith and M. J. Demetsky, "Short-term traffic flow prediction: Neural network approach," *Transportation Research Record*, vol. 1453, pp. 98–104, 1994.

[13] M. S. Dougherty and M. R. Cobbett, "Short-term inter-urban traffic forecasts using neural networks," *International Journal of Forecasting*, vol. 13, no. 1, pp. 21–31, 1997.

[14] H. Dia, "An object-oriented neural network approach to short-term traffic forecasting," *European Journal of Operational Research*, vol. 131, no. 2, pp. 253–261, 2001.

[15] F. Jin and S. Sun, "Neural network multitask learning for traffic flow forecasting," in *Proceedings of the 2008 International Joint Conference on Neural Networks (IJCNN '08)*, pp. 1897–1901, June 2008.

[16] K. Kumara, M. Paridab, and V. K. Katiyar, "Short term traffic flow prediction for a non urban highway using artificial neural network," *Procedia-Social and Behavioral Sciences*, vol. 104, pp. 755–764, 2013.

[17] B. Park, C. J. Messer, and T. Urbanik II, "Short-term freeway traffic volume forecasting using radial basis function neural network," *Transportation Research Record*, no. 1651, pp. 39–47, 1998.

[18] Y. Gao and M. J. Er, "NARMAX time series model prediction: feedforward and recurrent fuzzy neural network approaches," *Fuzzy Sets and Systems*, vol. 150, no. 2, pp. 331–350, 2005.

[19] M.-C. Tan, S. C. Wong, J.-M. Xu, Z.-R. Guan, and P. Zhang, "An aggregation approach to short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 1, pp. 60–69, 2009.

[20] K. Y. Chan, T. S. Dillon, J. Singh, and E. Chang, "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and levenberg-marquardt algorithm," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 644–654, 2012.

[21] G. A. Davis and N. L. Nihan, "Nonparametric regression and short-term freeway traffic forecasting," *Journal of Transportation Engineering*, vol. 117, no. 2, pp. 178–188, 1991.

[22] A. L. Ding, X. M. Zhao, and L. C. Jiao, "Traffic flow time series prediction based on statistics learning theory," in *Proceedings of the 5th IEEE International Conference on Intelligent Transportation Systems (ITSC '02)*, pp. 727–730, September 2002.

[23] A.-R. Mohamed, T. N. Sainath, G. Dahl, B. Ramabhadran, G. E. Hinton, and M. A. Picheny, "Deep belief networks using discriminative features for phone recognition," in *Proceedings of the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '11)*, pp. 5060–5063, May 2011.

[24] V. Nair and G. E. Hinton, "3D object recognition with deep belief nets," in *Proceedings of the 23rd Annual Conference on Neural Information Processing Systems (NIPS '09)*, pp. 1339–1347, December 2009.

[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 26th Annual Conference on Neural Information Processing Systems (NIPS '12)*, pp. 1097–1105, Lake Tahoe, Nev, USA, December 2012.

[26] H. Lee, L. Yan, P. Pham, and A. Y. Ng, "Unsupervised feature learning for audio classification using convolutional deep belief networks," in *Proceedings of the 23rd Annual Conference on Neural Information Processing Systems (NIPS '09)*, vol. 9, pp. 1096–1104, December 2009.

[27] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2191–2201, 2014.

[28] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.

[29] P. Lingras, S. Sharma, and M. Zhong, "Prediction of recreational travel using genetically designed regression and time-delay neural network models," *Transportation Research Record*, no. 1805, pp. 16–24, 2002.

[30] J. W. C. Van Lint, S. P. Hooqendoorn, and H. J. Van Zuvlen, "Freeway travel time prediction with state-space neural networks modeling state-space dynamics with recurrent neural networks," *Transportation Research Record*, no. 1811, pp. 30–39, 2002.

[31] S. Ishak, P. Kotha, and C. Alecsandru, "Optimization of Dynamic Neural Network Performance for Short-Term Traffic Prediction," *Transportation Research Record*, no. 1836, pp. 45–56, 2003.

[32] J. W. C. van Lint, S. P. Hoogendoorn, and H. J. van Zuylen, "Accurate freeway travel time prediction with state-space neural networks under missing data," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 5-6, pp. 347–369, 2005.

[33] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[34] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.

[35] Z. Zhao, W. Chen, X. Wu, P. C. Chen, and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68–75, 2017.

[36] K. C. Dey, A. Mishra, and M. Chowdhury, "Potential of intelligent transportation systems in mitigating adverse weather impacts on road mobility: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1107–1119, 2015.

[37] S. Dunne and B. Ghosh, "Weather adaptive traffic prediction using neurowavelet models," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 370–379, 2013.

[38] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[39] D. Ackley, G. Hinton, and T. Sejnowski, "A learning algorithm for Boltzmann machines," *Cognitive Science*, vol. 9, no. 1, pp. 147–169, 1985.

[40] G. E. Hinton, "Learning multiple layers of representation," *Trends in Cognitive Sciences*, vol. 11, no. 10, pp. 428–434, 2007.

[41] J. C. Tanner, "Effect of weather on traffic flow," *Nature*, vol. 169, no. 4290, p. 107, 1952.

[42] A. T. Ibrahim and F. L. Hall, "Effect of adverse weather conditions on speed-flow-occupancy relationships," *Transportation Research Record*, no. 1457, pp. 184–191, 1994.

[43] L. B. Smith, G. K. Byrne, B. R. Copperman et al., "An investigation into the impact of rainfall on freeway traffic flow," in *Proceedings of the 83rd Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 2004.

[44] *Highway Capacity Manual 2000*, Transportation Research Board, Washington, DC, USA, 2000.

[45] T. H. Maze, M. Agarwal, and G. Burchett, "Whether weather matters to traffic demand, traffic safety, and traffic operations and flow," *Transportation Research Record*, no. 1948, pp. 170–176, 2006.

[46] E. Chung, O. Ohtani, H. Warita et al., "Does weather affect highway capacity," in *Proceedings of the 5th International Symposium on Highway Capacity and Quality of Service*, Yakoma, Congo, 2006.

[47] M. Cools, E. Moons, and G. Wets, "Assessing the impact of weather on traffic intensity," *Weather, Climate, and Society*, vol. 2, no. 1, pp. 60–68, 2010.

[48] T. Hou, H. Mahmassani, R. Alfelor, J. Kim, and M. Saberi, "Calibration of traffic flow models under adverse weather and application in mesoscopic network simulation procedures," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2, no. 2391, pp. 92–104, 2013.

[49] W. H. K. Lam, M. L. Tam, X. Cao, and X. Li, "Modeling the effects of rainfall intensity on traffic speed, flow, and density relationships for urban roads," *Journal of Transportation Engineering*, vol. 139, no. 7, pp. 758–770, 2013.

[50] Y. Jia, J. Wu, Y. Du et al., "Impacts of rainfall weather on urban traffic in beijing: analysis and modeling," in *Proceedings of the 94th Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 2015.