Traffic Flow Prediction at Varied Time Scales via Ensemble Empirical Mode Decomposition and Artificial Neural Network

Xinqiang Chen 1,2,*, Jinquan Lu 3, Jiansen Zhao 3,*, Zhijian Qu 4, Yongsheng Yang 1 and Jiangfeng Xian 3

1 Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China; chenxinqiang@stu.shmtu.edu.cn (X.C.); yangys@shmtu.edu.cn (Y.Y.)
2 Institute of Atmospheric Sciences & Department of Atmospheric and Oceanic Sciences, Fudan University, Shanghai 200438, China
3 Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China; lujinquan@stu.shmtu.edu.cn (J.L.); 201530110005@stu.shmtu.edu.cn (J.X.)
4 Electrical and Automation Engineering College, East China Jiaotong University, Nanchang 330013, China; 08117324@bjtu.edu.cn
* Correspondence: jszhao@shmtu.edu.cn

Received: 15 April 2020; Accepted: 27 April 2020; Published: 2 May 2020

Abstract: Accurate traffic flow data is crucial for traffic control and management in an intelligent transportation system (ITS), and thus traffic flow prediction research attracts significant attention in the transportation community. Previous studies have suggested that raw traffic flow data may be contaminated by noises caused by unexpected reasons (e.g., loop detector damage, roadway maintenance, etc.), which may degrade traffic flow prediction accuracy. To address this issue, we proposed an ensemble framework via ensemble empirical mode decomposition (EEMD) and artificial neural network (ANN) to predict traffic flow under different time intervals ahead. More specifically, the proposed framework firstly employed the EEMD model to suppress the noises in the raw traffic data, which were then processed to predict traffic flow at time steps under different time scales (i.e., 1, 2, and 10 min). We verified our model performance on three loop detectors’ data, which were supported by the Department of Transportation, Minnesota. The research findings can help traffic participants collect more accurate traffic flow data and thus benefits transportation practitioners by helping them to make more reasonable traffic decisions.

Keywords: traffic flow data; prediction; denoising; varied time scales

1. Introduction

Rapid economic development has motivated a sharp increase in traffic demand and, thus, led to various traffic problems (e.g., traffic congestion, air pollution, and traffic accidents). Intelligent transportation systems (ITS) that support traffic control and management are considered as one of the efficient techniques for alleviating such problems. Traffic flow prediction provides critical traffic state information for the ITS system, which helps traffic participants make better traveling decisions and enhance traffic operation efficiency [1–4]. More specifically, we can develop better traffic control strategies (e.g., adaptive traffic signal control, dynamic speed-limit setting, etc.) and fine-tune more appropriate traffic parameters that consider roadway traffic condition fluctuation interference with the help of traffic flow data. In this manner, the successfullness of traffic control strategy is highly depended on the resolution of traffic flow prediction data [5]. Lane-level traffic flow data is
more sensitive to microscopic traffic state estimation accuracy (such as traffic speed, volume, and occupancy) and, thus, has become a hot topic in the traffic community [6–8].

Traffic flow prediction can be roughly divided into short-term and long-term levels in terms of the time span. More specifically, long-term traffic flow prediction aims to provide traffic flow data for several hours in advance, while short-term traffic flow prediction is implemented on the minute level (which is our research focus) [9,10]. Previous studies suggest that linear models, nonlinear models, and hybrid models are three types of typical traffic flow prediction techniques [11]. Linear models employ mathematical methods to conduct traffic flow prediction tasks. For instance, the Autoregressive Integrated Moving Average Model (ARIMA) showed a satisfactory performance on long-term traffic flow prediction tasks [12,13]. Nonlinear models introduce relevant machine learning methods to tackle the traffic flow prediction challenge. For instance, the relevant neural network models have shown success in many traffic flow prediction applications. Yasdi et al. employed an artificial neural network (ANN) to forecast traffic flow on a single roadway segment [14]. Yao et al. established a granular traffic flow forecasting model with an ANN model, and the experimental results indicated that the proposed model outperformed the other popular models (e.g., the Robertson model) [15].

It is noted that nonlinear models have enjoyed huge success in tackling many traffic flow prediction tasks. However, the uncertainty characteristic of traffic flow data may impose a negative effect on the nonlinear prediction models, which can degrade a model’s performance. The hybrid models (i.e., combining both linear and nonlinear models) are proposed to overcome the disadvantages. Hou et al. proposed a novel long-term traffic flow forecasting method, which was verified with real-time traffic data [16]. Jonathan et al. verified the hierarchical temporal memory (HTM) model’s performance when conducting a short-term traffic flow prediction task over real-world traffic data from Sydney. Although the long short-term memory (LSTM) showed better performance than that of the HTM model, Jonathan believed that the HTM model is a potentially efficient method for short-term traffic flow prediction tasks [17]. Zhao et al. proposed the temporal graph convolutional network (T-GCN) model for traffic forecasting based on the urban road network and found that the predictions outperformed state-of-the-art baselines on real-world traffic datasets [18]. Similar research can be found in [4,19–23].

Initial traffic data collected from inductive loop detectors may involve unexpected outliers, and thus many research interests have been paid to enhance the data quality. More specifically, many data denoising models have been introduced to suppress noises before the implementation of traffic flow prediction tasks, and it has been found that the prediction performances of the model combined with a denoising algorithm are better than that of the model without a denoising process [24], such as the wavelet Kalman filter model [25–27], wavelet transform [28–30]. Empirical mode decomposition (EMD) was first proposed to remove nonlinear noises from the initial data series [31]. EMD extracts the intrinsic mode function (IMF) sets from the input data samples, which can be separated into high-frequency (HF) and low-frequency (LF) portions. The HF segments are considered as data details and noises embedded in the original traffic data, and the LF counterparts are the data’s contours. The EMD model is easily interfered with by the mode mixing challenge, which may severely degrade the EMD model’s performance when the IMF segments contain intermittence features. To address the issue, the ensemble empirical mode decomposition (EEMD) model was proposed by Wu et al. [32]. The EEMD model discards the noisy IMFs and selects noise-free IMF samples to reconstruct smoothed data.

This study aims to propose a simple but efficient traffic flow prediction framework based on EEMD and an artificial neural network (ANN). More specifically, we introduced the EEMD model to cleanse the initial traffic flow data collected from neighboring loop detectors, which was installed on the same roadway lane. The ANN model was then employed to forecast short-term traffic flow data at different time scales. We verified the proposed framework performance in both data cleansing and prediction procedures. The findings of the research can provide accurate traffic flow data in advance, which benefits traffic authorities by enabling them to take more reasonable traffic management and control measurements to reduce traffic congestion and enhance traffic safety. The remainder of the
paper is organized as follows: Section 2 illustrates the proposed traffic flow data denoising model and prediction model in detail. Section 3 describes the data source and specific experimental results. Section 4 briefly concludes the research.

2. Methodology

2.1. Schematic Overview

Loop-detector-generated traffic data is crucial for traffic flow prediction accuracy, which is supposed to be smooth and noise free. However, unwanted factors (e.g., detector damage, roadway maintenance, etc.) may deteriorate the original data quality, which can further affect the traffic flow data prediction accuracy. In this way, traffic flow data samples provided by loop deductive detectors are composed of smooth (i.e., ground truth traffic flow data) and noisy samples (anomalous data). We firstly introduced the EEMD method to correct the raw traffic data, and then the ANN model was employed to predict traffic flow. The flowchart for the proposed framework is shown in Figure 1.

![Flowchart for the proposed framework](image)
2.2. EEMD Model for Denoising Raw Traffic Flow Data

The EEMD model has shown great success in traffic flow data denoising tasks due to its features that adaptively decompose both stationary and nonstationary data. Due to such advantages, the EEMD model significantly outperforms other data denoising models (e.g., wavelet filters, the short-time Fourier transform, moving average, etc. [33,34]). The EEMD method exploits the IMFs from the input raw traffic data with the shifting procedure, which is described in detail as follows:

(a) initialize all the parameters. Add white noise into the raw traffic data;
(b) recognize the local maximum and minimum values for the input traffic flow data series;
(c) connect the maximum value points to obtain the upper envelop, U(t), and lower envelope, L(t), in a similar manner;
(d) calculate the average envelope, M(t), with the upper and lower envelopes using Equation (1):

\[ M(t) = \frac{L(t) + U(t)}{2} \]  

(1)

(e) obtain the data difference, H(t), between the raw traffic flow data, I(t), and the average envelope, M(t), using Equation (2):

\[ H(t) = I(t) - M(t) \]  

(2)

(f) obtain the IMF element when the following two conditions are satisfied: the number difference between the extrema points and the zero-crossing samples is less than 1; the average value for the two points from the upper and lower envelopes should be zero. More specifically, the H(t) data sample is considered as the IMF when the above two conditions are met. If not, the upper and lower envelops of H(t) are employed to conduct the kth round for the purpose of searching for the IMF with the following rule:

\[ H'_K(t) = H(t) - M'_K(t) \]  

(3)

where \( M'_K(t) \) is the kth (k = 1, 2, ..., s) round of mean envelope for the traffic flow data, and the parameter \( H'_K(t) \) is the potential IMF for the kth (k = 1, 2, ..., s) round iteration procedure.

The procedure stops when the updated \( H'_K(t) \) meets the previously mentioned two conditions. In other words, the obtained \( H'_K(t) \) is considered as the IMF (denoted as \( C_1(t) \)). Then, the difference between the initial traffic data I(t) and the \( C_1(t) \) is calculated (see Equation (4)), which are further processed to exploit the remaining IMFs. Note that the \( C_1(t) \) plays the role of dyadic filter bank, indicating that the residual traffic flow data, \( r_1(t) \), contained a longer traffic flow variation tendency.

\[ r_1(t) = I(t) - C_1(t) \]  

(4)

The EEMD shifting procedure is iteratively implemented unless the residual traffic flow data, \( r_1(t) \), is a monotonic data series, or only one extrema value is found. The relationship between raw traffic data I(t), decomposed IMFs \( C_i(t) \) (i = 1, 2, ..., m), and residual data is established as follows:

\[ I(t) = \sum_{i=1}^{m} C_i(t) - R(t) \]  

(5)

where m is the number for the IMF set and R(t) is the residual data.
When the EEMD model finished the data decomposition procedure, we could determine the final data denoising result by averaging the IMFs (see Equation (6)). It is noted that the added white noise was suppressed when calculating the mean value from IMF segments.

\[ I(t) = \frac{\sum_{j=1}^{E_n} C_i(t)}{E_n} + R(t) \]  

where parameter \( E_n \) is the ensemble number and \( R(t) \) is the residual traffic data.

2.3. Traffic Flow Prediction with ANN Model

The artificial neural network model aims to learn intrinsic nonlinear relationships between the input and output traffic flow data, which follows an information perception rule by human being. The ANN model presents a high performance on incomplete associative memory, pattern recognition, and other similar tasks. Moreover, it demonstrates the efficacy in dealing with the nondeterministic polynomial-time hardness problems, for which it is difficult to build a model. Considering the traffic flow data features, we employed the back-propagation (BP) neural network (a type of feed-forward ANN model) to fulfill the traffic data prediction task.

The input layer of the BP network receives the raw traffic flow data sequence, which are further processed in hidden layers. Note that the input traffic data points are mapped into hidden nodes with different weights. More specifically, the input connection weights and thresholds (which is defined as bias) serve as the BP network input. After that, the hidden layers learn the intrinsic traffic flow data patterns in an iterative manner. By successfully exploring the intrinsic features, the hidden node outputs can be obtained by applying a transfer function (the sigmoidal function was used in our study). During the network training procedure, the BP network quantifies the difference (i.e., error) between the prediction outputs and ground truth traffic flow data. The BP network fine-tunes the network (both the structure and parameter setting) according the back-propagated errors, and such a procedure stops when the default stopping criteria is met. The BP network structure used in our study is shown in Figure 2.

![Figure 2. Overview of the back-propagation (BP) network structure in our proposed network.](image)

Suppose the neuron number in the BP network input layer, hidden layer, and the output layer are \( g, p, \) and \( q \). The parameter \( V^z_r (r = 1, 2, \ldots, g; z = 1, 2, \ldots, p) \) is the weight between neurons connecting the input and hidden layers, and \( W^z_r (r = 1, 2, \ldots, g; z = 1, 2, \ldots, q) \) is the weight linking the hidden
and output layers. The thresholds for outputting the network-learned traffic flow data pattern are $\theta_r$ \((r = 1, 2, \ldots, p)\) and $\mu_r$ \((r = 1, 2, \ldots, q)\), respectively. In our study, the ANN employed the BP network as the default model when no further specifications were provided.

2.4. Prediction Goodness Measurements

To verify the proposed framework performance, we compared the predicted traffic flow data with ground truth data by statistical measurements. Three goodness measurement indicators were introduced to quantify the model performance, which were the mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RSME). For predicting any given traffic flow data series, we obtained the above three indicators with Equations (7)–(9). Note that a smaller MAE, MAPE, or RMSE indicated higher prediction accuracy, while larger values indicated that the model prediction performance may not have been satisfied.

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^{m} |p_i - y_i|$$  \hspace{1cm} (7)

$$\text{MAPE} = \frac{100}{m} \sum_{i=1}^{m} \frac{|y_i - p_i|}{y_i}$$  \hspace{1cm} (8)

$$\text{RSME} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (p_i - y_i)^2}$$  \hspace{1cm} (9)

where $p_i$ predicted the traffic flow data samples and $y_i$ is the ground truth data.

3. Experiment

Traffic flow prediction is crucial for traffic control and management, and accurate traffic flow prediction can significantly benefit roadway safety and traffic efficiency. To this aim, we quantitatively evaluated our proposed model (i.e., a combination of EEMD and an ANN model, abbreviated as EEMD+ANN) performance at different time spans ahead. More specifically, we predicted traffic flow at varied steps ahead based on traffic flow data at different time scales. For instance, traffic flow prediction on the 3-step ahead, based on 2 min data (the time scale was 2 min), was to predict traffic flow at 6 min time ahead based on the 2 min data. The rule was applicable to 1-step, 6-step, and 10-step traffic flow prediction under other time scales (1 and 10 min). For the purpose of model performance comparison, we implemented the combination of EMD and an ANN model (abbreviated as EMD+ANN) and a conventional ANN model to predict the traffic flow. Note that 70% of traffic flow data was employed to train the artificial neural network, and 10% of data samples were used for the purpose of network validation and parameter fine-tuning. The remaining 20% of data were used for evaluating the performance of the traffic flow prediction model. Both the input and feedback delays were selected from 1 to 2, and the hidden layer size was 10.

3.1. Data

We collected the traffic flow data from loop detectors installed in freeways in Minnesota State, USA, and data access was supported by the Minnesota Department of Transportation and the Transportation Data Research Laboratory at the University of Minnesota Duluth [35]. The original publicly accessible data included traffic volume, speed, occupancy, and we only recorded the traffic flow data from three neighboring detectors installed on the same freeway lane. Besides, we collected the traffic flow data during the time period from 1 January 2016 to 15 January 2016. Note that the time resolution for the raw data was 30 s, and we have aggregated the data into 1, 2, and 10 min for the traffic flow prediction task. More details about the experiment are shown in Table 1.
Table 1. Experimental data—detailed information.

| Detector Location | Minnesota State |
|-------------------|-----------------|
| time duration     | 15 days         |
| time resolution   | 30 s            |
| data sample for 1 min (1 step) | 21,600          |
| data sample for 2 min (1 step)  | 10,800          |
| data sample for 10 min (1 step) | 2160            |

3.2. Traffic Flow Prediction Results Analysis

3.2.1. Parameter Settings

Added white noise, $A_n$, and the ensemble number, $E_n$, are closely related to the EEMD denoising performance, and thus we firstly describe the two parameters’ determination procedures in detail. The relationship between the ensemble number, $E_n$, and added white noise, $A_n$, is shown in Equation (10), where $\varepsilon_n$ demonstrates the EEMD denoising performance on the input traffic flow data. Note that a larger $\varepsilon_n$ indicates superior data cleansing performance and vice versa. However, an obvious weakness is that more computation cost (e.g., computation resources, longer computation time, etc.) is required to obtain better denoising results (i.e., a larger $\varepsilon_n$). Previous studies have suggested that a default $A_n$ and $E_n$ of 0.2 and 1000 can obtain satisfactory performance in many denoising applications [31].

$$\varepsilon_n = \frac{A_n}{\sqrt{E_n}}$$ (10)

In our study, we applied the EEMD model to smooth the 1 min traffic flow data for determining the optimal parameter settings in the EEMD model. More specifically, we obtained the spectrum of the EEMD smoothed data by applying different settings to the two parameters, which were further analyzed for the purpose of parameter determination. The specific parameter settings were set as follows: (1) $A_n = 0.1$ and $E_n = 500$, (2) $A_n = 0.2$ and $E_n = 1000$, (3) $A_n = 0.3$ and $E_n = 1500$, and (4) $A_n = 0.4$ and $E_n = 2000$. For the purpose of simplicity, we only present the first four IMF spectrums, which clearly show a difference in the comparison with the remaining IMF counterparts. The spectrogram distributions of 1 min traffic data for the first four IMFs (denoted as IMF1, IMF2, IMF3, and IMF4) for detector ID # 5802 are shown in Figure 3, while Figures 4 and 5 are IMF spectrogram distributions for detector ID # 5805 and # 5808, respectively.

As shown in Figure 3, the spectrum features were very similar under four groups of different parameter settings in the EEMD model. Note that the larger intensity (i.e., close to the red area) in the color bar in Figure 3 (which is applicable to Figures 4 and 5) indicated high-frequency components and vice versa. Note that the four IMFs were the details of the traffic flow data, and thus the IMF spectrums that contain more high-frequency components were considered to have better smoothing results. IMF4 and IMF3 in Figure 3a,b, respectively, were quite close to low-frequency area, which indicated that the EEMD models may over-suppress noises in the raw traffic flow data. Although the spectrograms of IMF1, IMF2, and IMF4 were quite similar, we observed that the IMF3 spectrogram in Figure 3d contained more high-frequency components (the color of the spectrogram is closer to the red area) compared to the IMF3 spectrogram in Figure 3c. Note that the above spectrogram distribution analyses were applicable to Figures 4 and 5 (i.e., detector ID #5805 and #5808). Based on the above analysis, we set the default values of $A_n$ and $E_n$ as 0.4 and 2000, respectively.
Figure 3. Spectrogram distributions of 1 min traffic data for the first four intrinsic mode functions (IMFs) at detector ID #5802. Descriptions for (a–d) are given in figure.

Figure 4. Spectrogram distributions of 1 min traffic data for the first four IMFs at detector ID #5805. Descriptions for (a–d) are given in figure.
Figure 5. Spectrogram distributions of 1 min traffic data for the first four IMFs at detector ID #5808. Descriptions for (a–d) are given in figure.

3.2.2. Traffic Flow Smoothing Results

Traffic volume is supposed to be smooth considering that the vehicles passing through a detector area gradually increase (or decrease) without significant vehicle speed (or traffic volume) variation at a given time span. In that manner, we believe that traffic volume data is composed of continuous and smooth components (i.e., noise-free traffic flow data) and noises. Moreover, the EEMD model can efficiently exploit the stochastic data outliers that exist in the traffic volume (which are identified as noise-IMFs). Based on the above analysis, we employed the EEMD model to minimize the data noise. We presented the EEMD smoothing results on the traffic flow data collected at loop detector ID #5802, and then verified the denoising performance on the traffic flow data from the other two detectors (i.e., detector ID #5805 and #5808, respectively).

Although the 1 min traffic flow dataset contained a larger number of samples, we could still observe that the EEMD method suppressed the obvious outliers in the raw traffic flow data, as burr samples were significantly reduced (see Figure 6a). More specifically, the raw 1 min traffic data obtained from detector ID #5802 (see the black curve in Figure 6a) showed sudden fluctuations in peak and trough areas. More specifically, the spikes, dips, and choppy samples could be periodically found at different time spans. Note that a few traffic flow volume samples were 30% larger than the neighboring data (with an extreme data sample reaching two-fold higher than its neighbors). Such data errors imposed a significant threat to the traffic flow prediction performance. It is noted that the traffic volume in larger time scales (e.g., 2 and 10 min) obtained less outliers compared to those of the 1 min time scale (see the Figure 6b,c, respectively). The main reason was that traffic volume outliers in the larger time scales were significantly reduced by the data aggregation procedure.
The raw 2 and 10 min traffic volume data and the corresponding smoothed data for detector ID #5802 are shown in Figure 6b,c. The EEMD smoothing results on the 2 and 10 min traffic data could be better observed (compared to the 1 min counterparts) due to the data samples being evenly dispersed in the Cartesian coordinate system. We noticed that the 10 min traffic data showed quite smooth results, which successfully suppressed the spikes, dips, and choppy samples. More specifically, the obvious abnormal oscillations and burr data (caused by the detector damages, etc.) in the 10 min...
traffic volume data series were corrected into more reasonable values. The EEMD smoothing results on the data from loop detector ID #5805 and #5808 (see Figures 7 and 8, respectively) showed very similar results to those of detector ID #5802. From the perspective of quantitative analysis, we considered that the EEMD model successfully suppressed outliers in the raw traffic flow data.

We employed three evaluation metrics (i.e., median absolute deviation (MAD), mean square error (MSE), and Pearson correlation coefficient (PCC)) to qualitatively analyze the traffic flow data denoising performance, and the corresponding results are shown in Table 2 [31]. The MAD values for the three detectors under a 1 min time scale were 6.942 (for detector ID #5802), 6.921 (for detector ID #5805), and 6.554 (for detector ID #5808). The MAD distributions under 2 and 10 min time scales showed similar variation tendency to that of the 1 min time scale. In this manner, the MAD distributions for different detectors at the same time scale were quite close. Moreover, the distributions of the MSE and PCC indicators’ distributions confirmed that the EEMD smoothing results for different detectors under the same time scale were very similar.

Figure 7. EEMD smoothing results on the traffic data for detector ID #5805. Descriptions for (a–c) are given in figure.
Table 2 also shows that the aggregated traffic flow data at a larger time scale may lead to smoothing performance loss. More specifically, the MAD and MSE indicators for the 2 min time scale for the same detector were nearly two-fold compared to the counterparts of the 1 min time scale (e.g., the MAD and MSE for detector ID #5802 at the 2 min time scale were 13.665 and 14.540, respectively, while the counterparts at 1 min were 6.942 and 7.576). Note that the MAD and MSE at 10 min were approximately three-fold higher than those of the 2 min data, which were ten-fold larger in comparison with the
1 min data. Additionally, the PCC indicator variation at different time scales did not show an obvious difference and was bigger than 0.940 for each detector at each time scale. Based on the above qualitative and quantitative analyses, we considered that EEMD obtained a satisfactory smoothing performance at different time scales, and the smoothing accuracy was higher at smaller time scales (i.e., smoothing accuracy at 1 min > smoothing accuracy at 2 min > smoothing accuracy at 10 min).

Table 2. EEMD smoothing performance evaluation for the three detectors.

| Detector ID | 1 min | 2 min | 10 min |
|-------------|-------|-------|--------|
|              | MAD   | MSE   | PCC    | MAD   | MSE   | PCC    | MAD   | MSE   | PCC    |
| #5802       | 6.942 | 7.576 | 0.949  | 13.665| 14.540| 0.974  | 67.639| 61.827| 0.995  |
| #5805       | 6.921 | 7.724 | 0.947  | 13.759| 14.607| 0.974  | 69.240| 69.081| 0.995  |
| #5808       | 6.554 | 7.273 | 0.943  | 12.930| 13.715| 0.971  | 64.003| 56.378| 0.995  |

3.2.3. Traffic Flow Prediction Analysis

The 1-step traffic flow prediction results are shown in Table 3, and the 3-step, 6-step, and 10-step prediction results are presented from Tables 4–6. We will describe the 1-step traffic flow prediction results in detail, considering page limitation. From the perspective of 1 min traffic flow data, the MAE values for the EEMD+ANN framework (detector IDs were #5802, #5805, and #5808) were 0.145, 0.148, and 0.139, which were approximately the same. The MAEs of EMD+ANN and traditional ANN models were both larger than those of the EEMD+ANN counterparts (from the same detector). For instance, the MAE values for EMD+ANN and ANN at 1 min for detector ID #5802 were 0.206 and 2.765, which were at least 50% higher than the EEMD+ANN MAE. The traffic flow prediction performance at 2 and 10 min were similar to the 1 min results.

Table 3. Prediction performance on the 1-step traffic flow data for the three detectors.

|                  | EEMD+ANN | EMD+ANN | ANN  |
|------------------|----------|---------|------|
|                  | #5802    | #5805   | #5808|
| MAE 1 min        | 0.145    | 0.148   | 0.139|
| MAPE 2 min       | 1.233    | 1.281   | 1.363|
| MAPE 10 min      | 3.650    | 4.220   | 3.859|
| RMSE 2 min       | 0.296    | 0.303   | 0.268|
| RMSE 10 min      | 4.758    | 4.716   | 4.435|

|                  | EEMD+ANN | EMD+ANN | ANN  |
|------------------|----------|---------|------|
|                  | #5802    | #5805   | #5808|
| MAE 2 min        | 1.173    | 1.020   | 1.096|
| MAPE 10 min      | 4.902    | 5.290   | 4.340|
| RMSE 2 min       | 1.915    | 1.979   | 1.967|
| RMSE 10 min      | 6.565    | 7.115   | 7.481|

Table 4. Prediction performance on the 3-step traffic flow data for the three detectors.

|                  | EEMD+ANN | EMD+ANN | ANN  |
|------------------|----------|---------|------|
|                  | #5802    | #5805   | #5808|
| MAE 1 min        | 0.155    | 0.157   | 0.159|
| MAPE 2 min       | 1.173    | 1.020   | 1.096|
| MAPE 10 min      | 4.902    | 5.290   | 4.340|
| RMSE 2 min       | 1.915    | 1.979   | 1.967|
| RMSE 10 min      | 6.565    | 7.115   | 7.481|

|                  | EEMD+ANN | EMD+ANN | ANN  |
|------------------|----------|---------|------|
|                  | #5802    | #5805   | #5808|
| MAE 2 min        | 1.173    | 1.020   | 1.096|
| MAPE 10 min      | 4.902    | 5.290   | 4.340|
| RMSE 2 min       | 1.915    | 1.979   | 1.967|
| RMSE 10 min      | 6.565    | 7.115   | 7.481|
Table 5. Prediction performance on the 6-step traffic flow data for the three detectors.

| EEMD+ANN | EMD+ANN | ANN   |
|----------|---------|-------|
|          | #5802   | #5805 | #5808 | #5802   | #5805 | #5808 | #5802   | #5805 | #5808 |
| MAE      |         |       |       |         |       |       |         |       |       |
| 1 min    | 0.778   | 0.774 | 0.777 | 1.042   | 0.980 | 0.917 | 2.888   | 2.964 | 2.859 |
| 2 min    | 1.210   | 1.275 | 1.067 | 1.468   | 1.675 | 1.621 | 4.436   | 4.799 | 4.474 |
| 10 min   | 6.976   | 6.201 | 6.753 | 7.121   | 7.261 | 7.690 | 25.850  | 25.535 | 24.055 |
| MAPE     |         |       |       |         |       |       |         |       |       |
| 1 min    | 9.188   | 10.212| 12.253| 14.574  | 15.747| 19.822| 38.098  | 39.108 | 38.326 |
| 2 min    | 7.860   | 7.938 | 7.668 | 13.095  | 11.938| 9.257 | 31.175  | 40.335 | 34.201 |
| 10 min   | 7.040   | 7.217 | 10.659| 9.893   | 9.365 | 12.816| 27.051  | 29.169 | 25.727 |
| RMSE     |         |       |       |         |       |       |         |       |       |
| 1 min    | 1.104   | 1.005 | 1.029 | 1.435   | 1.343 | 1.243 | 3.814   | 4.010 | 3.693 |
| 2 min    | 1.592   | 1.663 | 1.473 | 1.999   | 2.281 | 2.268 | 5.824   | 6.361 | 6.192 |
| 10 min   | 9.055   | 8.549 | 8.253 | 9.206   | 9.061 | 10.771| 34.029  | 35.688 | 33.995 |

The MAPE indicated the loss performance for the traffic flow prediction models, while the RMSE measured the deviation between the predicted and the ground truth data. The MAPE indicators for detector ID #5802 at the 1 min time scale for the three models (i.e., EEMD+ANN, EMD+ANN, and ANN) were 1.780, 2.903, and 34.450, which indicated that the ANN-obtained prediction error was significantly larger than the other two models. Moreover, the EEMD+ANN prediction error (in terms of MAPE) was approximately half that of EMD+ANN. The MAPE distributions at the 10 min scale showed similar variation to those of the 1 and 2 min scales (see Table 3). To sum up, the MAPE distributions indicated that the EEMD+ANN model obtained minimal prediction loss (i.e., maximal prediction accuracy). Besides, the RMSE indicator variation in Table 3 shows a similar performance to the MAE and MAPE statistics. Based on the above analysis, we considered that the EEMD+ANN model obtained a satisfactory performance on the 1-step traffic flow prediction task.

The 3-step traffic flow prediction results, which confirmed that the prediction errors (in terms of the MAE, MAPE, and RMSE) for the ANN model were larger than that of the EEMD+ANN model, are shown in Table 4. Similarly, the 6-step and 10-step traffic flow prediction results shown in Tables 5 and 6 demonstrated that the EEMD+ANN model prediction accuracy at different time scales (for the same detector data) were higher than those of the counterparts. Moreover, it is observed that the prediction accuracy showed a decreasing tendency along with the increase of the prediction step. For instance, the MAPEs for the EEMD+ANN model at the time scale of 1 min under 1-step, 3-step, 6-step, and 10-step were 1.780, 1.915, 9.188, and 10.046 (see the sixth row in Tables 3–6). We can infer that traffic flow prediction at longer steps may be interfered with by unexpected factors (e.g., asymmetric volume distributions).

For the purpose of visualizing the prediction performance, we calculated the average MAE, MAPE, and RMSE values for the three detectors, which are shown in Figure 9. From the perspective of 1 min average traffic flow prediction accuracy, the hybrid models (i.e., EEMD+ANN and EMD+ANN) obtained better performance than that of the ANN model. More specifically, the average MAE for the...
EEMD+ANN hybrid model at 1-step, 3-step, 6-step, and 10-step were all smaller than 1 (see the left uppermost subplot in Figure 9), which were slightly lower than the counterparts of the EMD+ANN model. Moreover, the EEMD+ANN model outperformed the EMD+ANN model on the prediction task at different time scales. Such a phenomenon can be observed in the 2 and 10 min traffic flow prediction accuracy (see Figure 9). It was noted that the ANN-obtained average MAEs were three times higher than those of the hybrid models. The main reason was that the denoising procedure eliminated noisy traffic flow data, which were substituted with reasonable data. In this manner, interference coming from abnormal oscillations in the raw data was suppressed during the traffic flow prediction procedure.

![Figure 9. Average prediction error distributions at different time scales. Descriptions for (a–c) are given in figure.](image)

Figure 9 also indicated that predicting a longer time period may result in larger errors (i.e., a larger MAE, MAPE, and RMSE). Particularly, the MAE, MAPE, and RMSE showed an obvious increasing tendency when the prediction step became larger. More specifically, the statistical indicators’ variation on the prediction task at a larger time step demonstrated a prediction performance decrease. One of the potential reasons is that training samples under the longer time interval is insufficient, and thus the prediction models may fail to fully exploit the intrinsic traffic flow pattern. To sum up, the traffic flow prediction task on the 1 min data obtained better performance (in terms of aggregated MAE, MAPE, and RMSE) than that of the 2 and 10 min data.

### 4. Conclusions

It is not easy to predict accurate traffic flow data via historical information due to their nonlinear and unstable features. We proposed an ensemble framework with EEMD and an ANN model to predict traffic flow data at different yet typical time scales (i.e., 1-step, 2-step, 6-step, and 10-step). The proposed EEMD model decomposed the raw traffic flow data into different IMFs, and the noisy IMFs were suppressed while the other IMFs were aggregated into the noise-free traffic flow data. After this, the ANN model was introduced to predict the traffic flow at 1-step, 3-step, 6-step, and 10-step
ahead under different time scales (1, 2, and 10 min). The experimental results showed that hybrid models (i.e., EEMD+ANN and EMD+ANN) significantly outperform the conventional ANN model on the traffic flow prediction task. More specifically, the MAEs, MAPEs, and RMSEs obtained by the EEMD+ANN model at different time scales were significantly smaller than the counterparts of EMD+ANN and ANN. The proposed framework can be easily transferred to support other traffic data prediction tasks (speed, density, etc.) due to the generative features of both EEMD and the ANN model.

We can further expand our research in the following aspects: First, although we tested the performance of the proposed prediction model at different time scales, we can obtain a more holistic performance by comparing it with other traffic flow prediction models, such as ARIMA. Second, we only tested our model’s performance on a short-term traffic flow prediction task. It deserves further attention to test the model’s performance on a long-term traffic prediction task. Third, we can implement relevant deep learning methods to implement the short-term traffic flow prediction task, which may provide us with additional interesting results and findings. Last but not least, we can obtain more realistic and real-time traffic state information with accurate traffic flow prediction results, which can benefit transportation efficiency improvement on a more refined level.

**Author Contributions:** Conceptualization, X.C., J.L.; methodology, X.C., J.L., J.Z., and Y.Y.; formal analysis, X.C., J.Z., Y.Y., and Z.Q.; investigation, J.L.; writing—review and editing, X.C., Y.Y., and J.X.; funding acquisition, J.Z., Z.Q., and Y.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China: No. 51709167, 51708094, and 51867009; Shanghai Committee of Science and Technology, China: No. 18040501700, 18295801100, and 17595810300; the Foundation Plan for Distinguished Young Scholars in Jiangxi Province: No. 20162BCB23045; and the Applied and Cultivation Program of Science and Technology Department of Jiangxi Province: No. 20181BB58010.

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

**References**

1. Tian, Y.; Zhang, K.; Li, J.; Lin, X.; Yang, B. LSTM-based traffic flow prediction with missing data. *Neurocomputing* **2018**, *318*, 297–305. [CrossRef]
2. Chen, X.; Li, Z.; Wang, Y.; Cui, Z.; Shi, C.; Wu, H. Evaluating the impacts of grades on vehicular speeds on interstate highways. *PLoS ONE* **2017**, *12*, e0184142. [CrossRef]
3. Liu, J.; Guan, W. A summary of traffic flow forecasting methods. *J. Highw. Transp. Res. Dev.* **2004**, *3*, 82–85.
4. Chen, X.; Xu, X.; Yang, Y.; Wu, H.; Tang, J.; Zhao, J. Augmented Ship Tracking Under Occlusion Conditions From Maritime Surveillance Videos. *IEEE Access* **2020**, *8*, 42884–42897. [CrossRef]
5. Emami, A.; Sarvi, M.; Bagloee, S.A. Short-term traffic flow prediction based on faded memory Kalman Filter fusing data from connected vehicles and Bluetooth sensors. *Simul. Model. Pract. Theory* **2019**, *102025*, in press. [CrossRef]
6. Gu, Y.; Lu, W.; Qin, L.; Li, M.; Shao, Z. Short-term prediction of lane-level traffic speeds: A fusion deep learning model. *Transp. Res. Part C Emerg. Technol.* **2019**, *106*, 1–16. [CrossRef]
7. Vlahogianni, E.I.; Karlaftis, M.G.; Golas, J.C. Short-term traffic forecasting: Where we are and where we’re going. *Transp. Res. Part C Emerg. Technol.* **2014**, *43*, 3–19. [CrossRef]
8. Yu, S.; Fu, R.; Guo, Y.; Xin, Q.; Shi, Z. Consensus and optimal speed advisory model for mixed traffic at an isolated signalized intersection. *Phys. A Stat. Mech. Appl.* **2019**, *531*, 121789. [CrossRef]
9. Zhang, Z.; Wang, Y.; Chen, P.; He, Z.; Yu, G. Probe data-driven travel time forecasting for urban expressways by matching similar spatiotemporal traffic patterns. *Transp. Res. Part C Emerg. Technol.* **2017**, *85*, 476–493. [CrossRef]
10. Chen, X.; Qi, L.; Yang, Y.; Luo, Q.; Postolache, O.; Tang, J.; Wu, H. Video-Based Detection Infrastructure Enhancement for Automated Ship Recognition and Behavior Analysis. *J. Adv. Transp.* **2020**, *2020*, 1–12. [CrossRef]
11. Wu, Y.; Tan, H.; Qin, L.; Ran, B.; Jiang, Z. A hybrid deep learning based traffic flow prediction method and its understanding. *Transp. Res. Part C Emerg. Technol.* **2018**, *90*, 166–180. [CrossRef]
12. Wang, Y.; Li, L.; Xu, X. A piecewise hybrid of ARIMA and SVMs for short-term traffic flow prediction. In *Proceedings of the International Conference on Neural Information Processing*, Guangzhou, China, 14–18 November 2017.
13. Williams, B.M.; Hoel, L.A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. *J. Transp. Eng.* 2003, 129, 664–672. [CrossRef]
14. Yasdi, R. Prediction of road traffic using a neural network approach. *Neural Comput. Appl.* 1999, 8, 135–142. [CrossRef]
15. Zhihong, Y.; Yangsheng, J.; Peng, H.; Xiaoling, L.; Tao, X. Traffic Flow Prediction Model Based on Neural Network in Small Time Granularity. *J. Transp. Syst. Eng. Inf. Technol.* 2017, 17, 67–73.
16. Hou, Z.; Li, X. Repeatability and Similarity of Freeway Traffic Flow and Long-Term Prediction Under Big Data. *IEEE Trans. Intell. Transp. Syst.* 2016, 17, 1786–1796. [CrossRef]
17. Mackenzie, J.; Roddick, J.F.; Zito, R. An Evaluation of HTM and LSTM for Short-Term Arterial Traffic Flow Prediction. *IEEE Trans. Intell. Transp. Syst.* 2019, 20, 1847–1857. [CrossRef]
18. Zhao, L.; Song, Y.; Zhang, C.; Liu, Y.; Wang, P.; Lin, T.; Deng, M.; Li, H. T-gcn: A temporal graph convolutional network for traffic prediction. *IEEE Trans. Intell. Transp. Syst.* 2019, 1–11, Early Access. [CrossRef]
19. Babu, C.N.; Reddy, B.E. A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. *Appl. Soft Comput.* 2014, 23, 27–38. [CrossRef]
20. Do, L.N.N.; Vu, H.L.; Vo, B.Q.; Liu, Z.; Phung, D. An effective spatial-temporal attention based neural network for traffic flow prediction. *Transp. Res. Part C Emerg. Technol.* 2019, 108, 12–28. [CrossRef]
21. Bogaerts, T.; Masegosa, A.D.; Angarita-Zapata, J.S.; Onieva, E.; Hellinckx, P. A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data. *Transp. Res. Part C Emerg. Technol.* 2020, 112, 62–77. [CrossRef]
22. Chen, X.; Yang, Y.; Wang, S.; Wu, H.; Tang, J.; Wang, Z. Ship Type Recognition via a Coarse-to-Fine Cascaded Convolution Neural Network. *J. Navig.* 2020. [CrossRef]
23. Chen, X.; Wang, S.; Shi, C.; Wu, H.; Zhao, J.; Fu, J. Robust Ship Tracking via Multi-view Learning and Sparse Representation. *J. Navig.* 2019, 72, 176–192. [CrossRef]
24. Chen, X.; He, Z.; Sun, L. A Bayesian tensor decomposition approach for spatiotemporal traffic data imputation. *Transp. Res. Part C Emerg. Technol.* 2019, 98, 73–84. [CrossRef]
25. Xie, Y.; Zhang, Y.; Ye, Z. Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition. *Comput. Aided Civ. Infrastruct. Eng.* 2007, 22, 326–334. [CrossRef]
26. Dunne, S.; Ghosh, B. Weather adaptive traffic prediction using neurowavelet models. *IEEE Trans. Intell. Transp. Syst.* 2013, 14, 370–379. [CrossRef]
27. Xiao, H.; Sun, H.; Ran, B.; Oh, Y. Fuzzy-neural network traffic prediction framework with wavelet decomposition. *Transp. Res. Rec.* 2003, 1836, 16–20. [CrossRef]
28. Jiang, X.; Adeli, H. Wavelet packet-autocorrelation function method for traffic flow pattern analysis. *Comput. Aided Civ. Infrastruct. Eng.* 2004, 19, 324–337. [CrossRef]
29. Tan, M.; Li, Y.; Xu, J. A Hybrid ARIMA and SVM Model for Traffic Flow Prediction Based on Wavelet Denoising. *J. Highw. Transp. Res. Dev.* 2009, 26, 127–132, 138.
30. Chen, X.; Chen, H.; Wu, H.; Huang, Y.; Yang, Y.; Zhang, W.; Xiong, P. Robust Visual Ship Tracking with an Ensemble Framework via Multi-View Learning and Wavelet Filter. *Sensors* 2020, 20, 932. [CrossRef]
31. Chen, X.; Li, Z.; Wang, Y.; Tang, J.; Zhu, W.; Shi, C.; Wu, H. Anomaly Detection and Cleaning of Highway Elevation Data from Google Earth Using Ensemble Empirical Mode Decomposition. *J. Transp. Eng. Part A Syst.* 2018, 144, 04018015. [CrossRef]
32. Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* 2009, 1, 1–41. [CrossRef]
33. Srivastava, M.; Anderson, C.L.; Freed, J.H. A new wavelet denoising method for selecting decomposition levels and noise thresholds. *IEEE Access* 2016, 4, 3862–3877. [CrossRef] [PubMed]
34. Luo, X.; Bhakta, T. Estimating observation error covariance matrix of seismic data from a perspective of image denoising. *Comput. Geosci.* 2017, 21, 205–222. [CrossRef]
35. Tang, J.; Chen, X.; Hu, Z.; Zong, F.; Han, C.; Li, L. Traffic flow prediction based on combination of support vector machine and data denoising schemes. *Phys. A Stat. Mech. Appl.* 2019, 534, 1–19. [CrossRef]