Article

Short-Term Forecast of OD Passenger Flow Based on Ensemble Empirical Mode Decomposition

Yi Cao, Xiaolei Hou * and Nan Chen

School of Transportation Engineering, Dalian Jiaotong University, Dalian 116028, China;
caoyi820619@aliyun.com (Y.C.); cn18510414695@163.com (N.C.)
* Correspondence: hxl18340877087@163.com; Tel.: +86-183-40877087

Abstract: The development of metro systems can be a good solution to many problems in urban transport and promote sustainable urban development. A metro system plays an important role in urban public transit, and the passenger-flow forecasting is fundamental to assisting operators in establishing an intelligent transport system (ITS). In order to accurately predict the passenger flow of urban metros in different periods and provide a scientific basis for schedule planning, a short-term metro passenger-flow prediction model is constructed by integrating ensemble empirical mode decomposition (EEMD) and long short-term memory neural network (LSTM) to solve the problem that the existing empirical mode decomposition (EMD) is prone to modal aliasing. According to the processed metro-card data, the time series of historical OD data of metro passenger flow is obtained. After EEMD modal decomposition, several intrinsic mode functions sub-items and residual items are obtained. Then, an LSTM network is constructed for prediction. The time step of the network is decided according to the partial autocorrelation functions. The prediction results of intrinsic mode function (IMF) and residual items are integrated to obtain prediction results. The station is classified according to the land types around the station, and the model is tested by using the metro automatic fare collection (AFC) data. In order to test the actual prediction, a different number of training set samples are selected to predict. The measured data of the next day is continuously added to the original training set to compare the prediction accuracy. The results show that the mean absolute percentage error (MAPE) and root mean square error (RMSE) of the EEMD-LSTM model are better than the EMD-LSTM in predicting the OD value of commercial–residential stations and scenic–residential stations. Compared with the EMD-LSTM model, the EEMD-LSTM model showed an average reduction by 3.112% in MAPE values and 15.889 in RMSE, indicating that the EEMD-LSTM has higher prediction accuracy, and EEMD-LSTM model has higher accuracy in short-term metro passenger-flow prediction. The average MAPE for the 35-to-42-day historical data sample decreased from 13.02% to 10.39% with a decreasing trend. It shows that the prediction accuracy keeps improving as the training set samples increase.

Keywords: traffic engineering; passenger-flow prediction; ensemble empirical mode decomposition; long short-term memory neural network; deep learning

1. Introduction

1.1. Background

In recent years, with the rapid development of urban society and economy, and the increasing travel demand of residents, urban traffic congestion has become increasingly prominent. The urban metro system boasts the advantages of a large capacity, fast speed, comfort and safety, and is not affected by ground traffic, which effectively alleviates traffic congestion. However, as an increasing number of passengers choose to travel by metro, there are rising concerns about the operation efficiency and safety of the metro system, such as the mismatch between the capacity and demand in peak hours, the delayed evacuation of passengers gathered on the platform, and the frequent occurrence of congestion and

even stampede accidents. Therefore, it is necessary to carry out in-depth research into high-precision prediction methods of short-term metro passenger flow, so as to provide a data basis for metro operation organization and management departments to formulate reasonable train operation plans. LSTM models are widely used in the field of traffic prediction. This is because the spatial and temporal distribution characteristics of metro passenger flow are becoming more and more complex. When the passenger flow fluctuates randomly, the prediction accuracy of the traditional model is often greatly affected. The prediction error of LSTM models can be very large, so combining with other methods will give better results. The EMD algorithm is the basic algorithm, and the EEMD algorithm is the improved algorithm. The EEMD algorithm is an effective method for analyzing and processing nonlinear non-stationary signals. The EEMD algorithm can effectively eliminate the interference noise in the signal and ensure the accuracy of prediction. In order to effectively process complex signals and reduce noise interference, a short-term metro passenger-flow prediction model is constructed by integrating ensemble empirical mode decomposition and a long short-term memory neural network to solve the problem that the existing empirical mode decomposition is prone to modal aliasing. The EEMD-LSTM model not only solves the mode mix problem of the EMD-LSTM model, but also improves the model prediction accuracy.

1.2. Purpose and Significance

In light of such a research background, it is necessary to achieve accurate OD point-to-point passenger-flow prediction between metro stations. With the development of intelligent transport devices such as AFC systems, real-time data such as the starting stations and destinations of passengers and their travel times can be quickly obtained and passenger-flow patterns can be quickly analyzed, making it possible to use intelligent technologies such as AFC data for OD demand forecasting [1]. According to existing research, the combined model of EEMD and LSTM can solve modal mixing and improve model prediction accuracy. Validated with the OD passenger-flow data of metro stations, the research shows that dynamic prediction, where the size of the training set continues to increase, has higher prediction accuracy than static prediction, where the size of the training set never changes. The research results have important practical significance for improving the prediction accuracy of OD passenger flow in short-term metro stations and reducing safety hazards such as metro station congestion. Accurate short-term forecast of metro passenger flow can reflect the travel needs of passengers on the metro network in the future, and passengers can obtain real-time and effective travel information, which helps them to choose premium travel routes and improve travel efficiency. It can provide effective information for relevant departments and operating organizations, and be used to make plans for a large number of passenger flows during peak hours. In case of emergency, it can serve as a basis for relevant departments to quickly formulate emergency plans, so as to efficiently evacuate passengers, which can not only ensure the safety of passengers, but also save operating costs and maintain the well-ordered operation of the metro.

2. Literature Review

At present, the methods of short-term passenger-flow prediction can be classified into two categories: the parametric model and non-parametric model. The former assumes that the passenger flow follows a certain probability distribution, and uses historical data to calibrate the model. The representative methods include the autoregressive differential average model [2,3], Kalman filter [4], and grey model [5]. Since the establishment of parametric models depends more on the a-priori knowledge of the research problem, it has a certain subjectivity. The latter makes no assumption on the probability distribution of passenger flow, and directly uses historical data to predict, which is more objective than the former, yet not conducive to being promoted. For example, Sun et al. [6] designed the wavelet-support vector machine (SVM) hybrid model, used the wavelet technology to decompose the passenger-flow data in high and low frequency, and then used the SVM

model to predict. Li et al. [7] proposed a multi-scale radial basis function network prediction method with sudden passenger flow as the research object. Zhang et al. [8] proposed a Conv-GCN deep-learning model for short-term passenger-flow prediction of urban rail transit by combining a graph convolution network and three-dimensional convolution neural network. In addition, the probability tree [9] and the decision tree [10] are also used in short-term metro passenger-flow forecasting.

With the development of big data, some scholars began to try to apply neural networks [11,12] to short-term passenger-flow forecasting, including long-term and short-term memory neural network [13,14], fully connected deep neural network [15], stack autoencoder [16], and deep belief network [17]. Among them, long short-term memory neural network (LSTM) prediction has attracted much attention due to its high accuracy. Li et al. [18] proposed an algorithm based on a back propagation long short-term memory algorithm, which can predict the speed of each vehicle on the road. Zhang et al. [19] used LSTM to predict short-term passenger flow through clustering to capture the characteristics and trends of passenger flow, and evaluated the reasonable time granularity interval. Li et al. [20] proposed a prediction model that combines inbound passenger flow characteristics with LSTM by extracting the characteristics affecting the accuracy of prediction. Some scholars improved the LSTM neural network. For example, Yang et al. [21] proposed an improved model based on the LSTM neural network, which improved the long-term correlation characteristics in passenger-flow data; Jia et al. [22] combined long short-term memory neural networks with stacked automatic encoders, and predicted the short-term passenger flow of metro stations. In order to effectively deal with complex signals and reduce noise interference, scholars combined empirical mode decomposition (EMD) with other models. Wei et al. [23] used an EMD and BP neural network combination model to predict short-term metro passenger flow, and the prediction accuracy was better than using a BP neural network alone; Wang et al. [24] used EMD and ARIMA to predict vehicle speed, and the prediction accuracy was better than ARIMA and ANN prediction methods in different scenarios.

Although domestic and foreign scholars have carried out a series of research on short-term metro passenger-flow prediction, the spatial and temporal distribution characteristics of metro passenger flow are becoming more and more complex. When the passenger flow fluctuates randomly, the prediction accuracy of the traditional model is often greatly affected.

According to the existing research, since the OD time series between subway stations is affected by the time and location of the entry or exit, it often exhibits nonlinearity and non-stationarity. If the basic data is directly input into the LSTM neural network, its own noise will interfere with the model to identify the spatial–temporal relationship between input and output passenger flow, resulting in reduced model prediction accuracy. Therefore, it is necessary to preprocess the time series before using neural network forecasting. Mixed modes refer to the inclusion of very different characteristic time scales in one IMF, or the distribution of similar characteristic time scales in different IMFs. This phenomenon is caused by the multiple jumps of the local extrema in a very short time interval in the process of empirical mode decomposition. Basically, the data is mixed with noise or discontinuous signals that we do not know about. For this problem, this paper proposes the ensemble empirical mode decomposition (EEMD) method to avoid mode mixing by stacking different Gaussian white noises with equal amplitude, and combines the advantages of long-term and short-term memory neural network (LSTM) in memory and forgetting. The EEMD-LSTM combination model is applied to the short-term passenger-flow prediction of a metro to ensure the accuracy of short-term prediction. Taking into account of the shortcomings of existing research, this paper makes two contributions as follows:

1.　In view of the fact that the EMD method is prone to modal aliasing, EEMD can avoid modal aliasing by stacking different Gaussian white noise with equal amplitude. With the advantages of the memory and forgetting of LSTM, the combination model

is applied to the field of metro short-term passenger-flow prediction to ensure the accuracy of short-term prediction.

2. The accuracy of EEMD-LSTM prediction in practical application is further explored. By changing the scale of the training set to achieve the effect of dynamic prediction, the feasibility of the model in practice is verified by comparing the static prediction results of dynamic prediction without changing the training set.

3. Methodology

3.1. Data Sources

Dalian Metro Line 1 and Line 2 basically covers the administrative center, the transportation hub, commercial area, universities and tourist attractions of Dalian. The AFC data of 42 stations of Dalian Metro Line 1 and Line 2 was selected for 91 days from 1 April to 30 June 2020. The missing value and abnormal value were processed to obtain the OD distribution among stations. The OD value among stations on 29 May is shown in Figure 1.
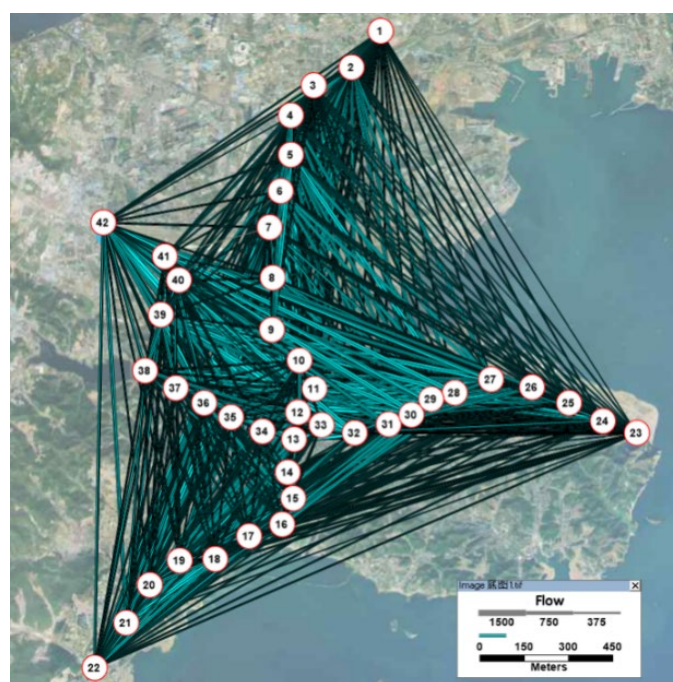

Figure 1. OD desire line of station.

The POI data of various metro stations and surrounding facilities are obtained from the Open Street map website. The K-Means clustering algorithm is easy to understand, operate and can be used in many fields. The algorithm is used as an unsupervised classification; it can effectively mine the internal features in a dataset [25]. According to the proportion of land use types within 500 m of the station, these stations are divided into residential stations, commercial stations, the hub station, scenic stations and the university stations by K-means cluster analysis in SPSS. The classification results are shown in Figure 2.

During the period of research, the passenger flow between commercial–residential stations is relatively high, while that between scenic–residential stations is relatively low, which is suitable for testing the prediction accuracy of the model. According to the classification results of Figure 2, the historical OD values between the typical commercial station Xi'an Road Station, the scenic station Xinghai Square Station and the residential station Malan Square Station were selected as research data, as shown in Figure 3.
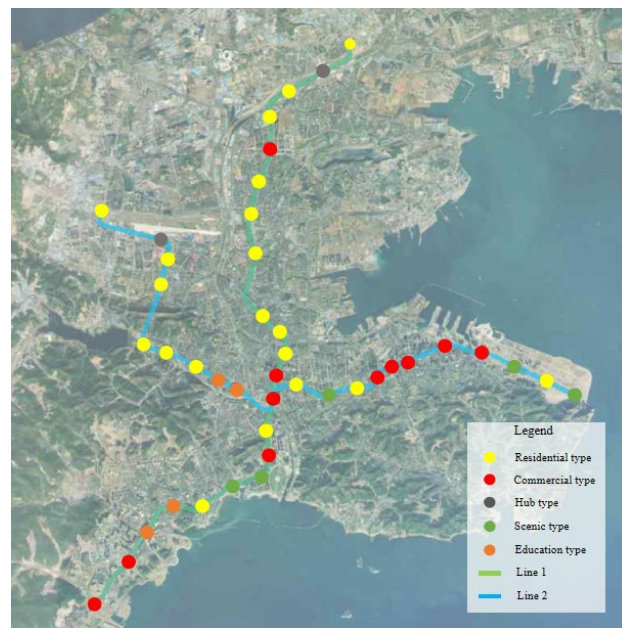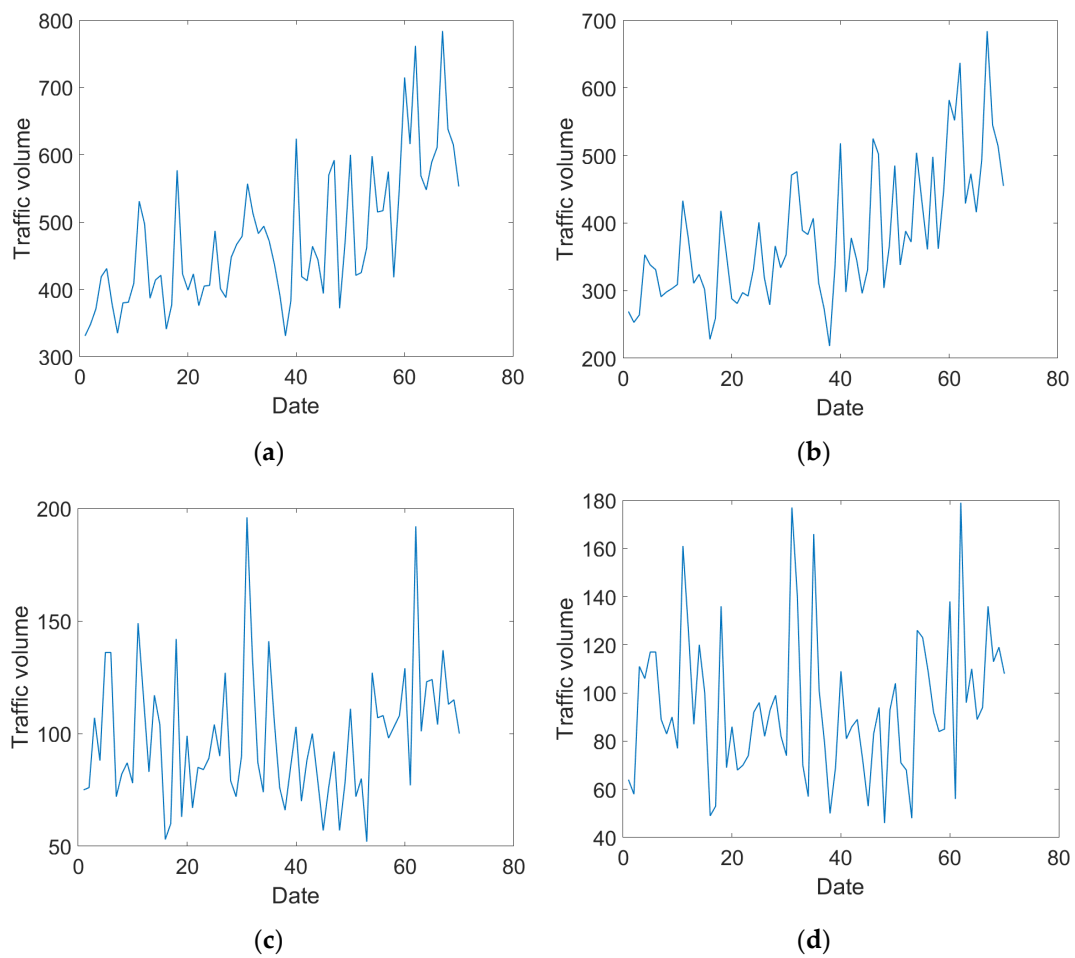
Figure 2. Classification results of station.

Figure 3. Historical OD values. (**a**) Historical OD of commercial–residential station; (**b**) historical OD of residential–commercial station; (**c**) historical OD of scenic–residential station; (**d**) historical OD of residential–scenery station.

3.2. Ensemble Empirical Mode Decomposition (EEMD)

EMD is a signal-processing method for non-stationary time series, and it believes that any signals can be decomposed into several simple signals with different periods and a residual signal. However, EMD method is easy to give rise to modal aliasing, that is, the signals of the same scale or frequency are divided into multiple eigenfunctions. Once the IMF experiences aliasing, the IMF is not a single frequency component, and the accurate instantaneous frequency cannot be obtained. EMMD is a noise-aided data-analysis method to make up for the shortcomings of EMD method. In EEMD, random Gaussian white noise sequence is added into the input signal, and the noised signal is decomposed by EMD. When the number of decompositions reaches the overall average number, the decomposition stops. When calculating the average number of IMF in the end, the minimum number of code components in each IMF decomposed by M times of experiments should be selected. The larger the overall average number of times M is, the closer the overall average of the random Gaussian white noise added is to zero [26]. In M times of experiments, the average value of each IMF is shown in Formula (1).

$$\bar{d}_i = \frac{\sum\limits_{m=1}^{M} d_{i,m}}{M} \tag{1}$$

where, d_i is the ith IMF obtained by EEMD decomposition; m is the number of times of experiments; and M is the overall average number of times.

3.3. Long Short-Term Memory Neural Network (LSTM)

LSTM is a memory cell designed on the basis of RNN and can select to memory important information and filter out the noise information, thereby reducing the memory burden and maintaining the long-term memory of the neural network, so that the model can also be well used for long-term sequences. For RNN, since the network layer updates information without restriction, it will make the information chaotic and easy to disappear and change, so the problem of gradient disappearance may happen. LSTM network adds forgetting unit and memory unit in the hidden layer. When there is new information input, LSTM network will filter out information which will be kept or discarded, and store important information into long-term memory. LSTM network has a relatively complex internal structure, conducts information transmission in a selective manner through unique gating unit and has a circular network structure with more complex neurons. The LSTM internal structure diagram is shown in Figure 4 [27].
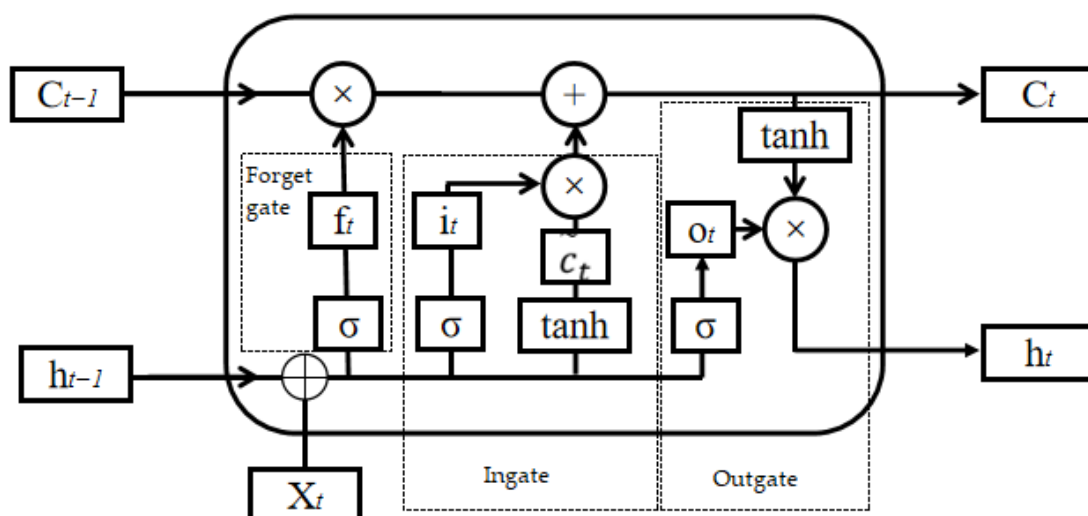

Figure 4. LSTM internal structure diagram.

3.4. EEMD-LSTM Model

The OD time series s(t) between metro stations is affected by the time and locations of entering or leaving the station, and often exhibits nonlinearity and non-stationarity. If s(t) is directly input to the LSTM neural network as the basic data, its own noise will interfere with the spatial–temporal relationship between the input and output passenger flow of the model identification, resulting in a decline in the prediction accuracy of the model. Therefore, it is necessary to preprocess s(t) before conducting neural network prediction. Through the original AFC data processing, the OD time series s(t) between stations is obtained. In view of the nonlinear and non-stationary characteristics of s(t) and to avoid modal superposition, this paper adopts the ensemble empirical mode decomposition (EEMD) to decompose s(t), and changes the signal extremum by adding different white noises with the same amplitude each time. After many rounds of decomposition, a series of intrinsic mode function sub-items (IMFs) are obtained and their overall average value is calculated to offset the added white noise, so as to suppress the occurrence of modal aliasing. For each IMF sub-item and residual, the corresponding LSTM models are trained, respectively. The Adam algorithm was used for model training, and it was determined through trial-and-error experiments that the number of iterations with the lowest average absolute percentage error was 200, the initial learning rate was 0.01, and the learning rate degradation factor was 0.2. After the LSTM network of each component is trained, the prediction results of each component are integrated and superimposed. Given that the result at this time is the one obtained by normalizing the data and is distributed in [0, 1], the final predicted results are obtained by conducting inverse normalization to the integrated results. Finally, the prediction results of each model are integrated to obtain the prediction results of passenger flow.

The specific steps of prediction are as follows:

1. The AFC data is pre-processed to obtain the OD time series s(t) between stations.
2. EEMD method is used to decompose s(t) to obtain n components of intrinsic mode function (IMF) and residuals.
3. The partial autocorrelation function (PACF) is used as an index to calculate the autocorrelation between the components of each eigenfunction and the residual, and the corresponding LSTM time step is determined.
4. The function components and residuals of the intrinsic model are divided into training set and test set to predict IMFs and residuals, respectively.
5. The predicted daily OD passenger-flow data is obtained by integrating the predicted intrinsic mode function components and residuals.

3.5. Model Building

The OD value from 1 April to 9 June 2020 was selected as the training sample, accounting for about 3/4 of the total, and the OD value from 10 June to 30 June 2020 was selected as the verification sample. MATLAB was employed to construct the historical OD time series from 1 April to 9 June 2020, and then EEMD decomposition was carried out. The standard deviation of white noise was set to 0.1, the number of adding noise was set to 10, and the number of integrations was set to 100. For example, in commercial–residential OD, five IMF subsequences and one residual item were obtained after decomposition. The decomposition results are shown in Figure 6. The flow chart of the prediction model is shown in Figure 5.

The lags of IMF and residuals are synchronized with the corresponding LSTM time step and usually determined by PACF, whose confidence interval was set to 95%. The decomposition results are shown in Figure 7. It can be seen from Figure 7 that the time steps of component 1–5 and residual are 3, 5, 7, 7, 7, 6, in turn, and the component 3–5 with low frequency is related to the OD value of the previous 7 days. This is because the passenger-flow regularity is mostly presented with one week as a cycle.
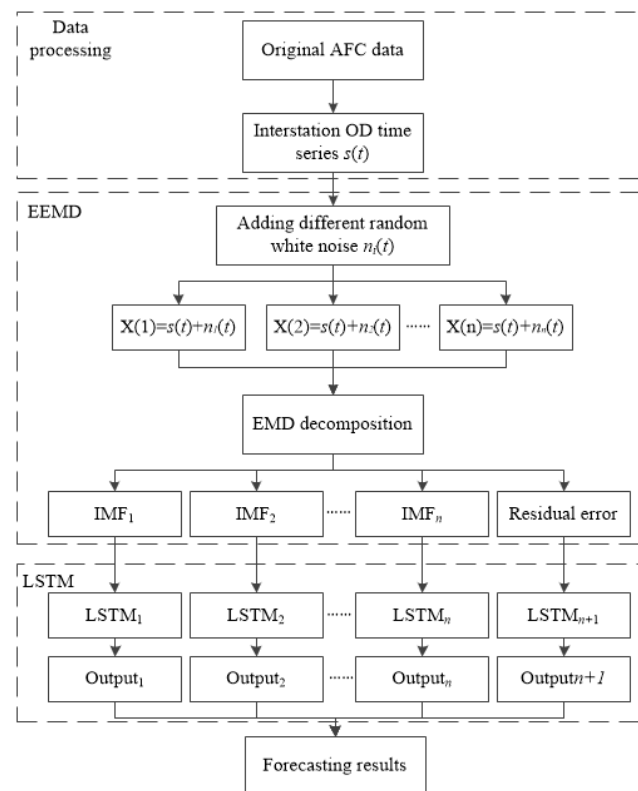
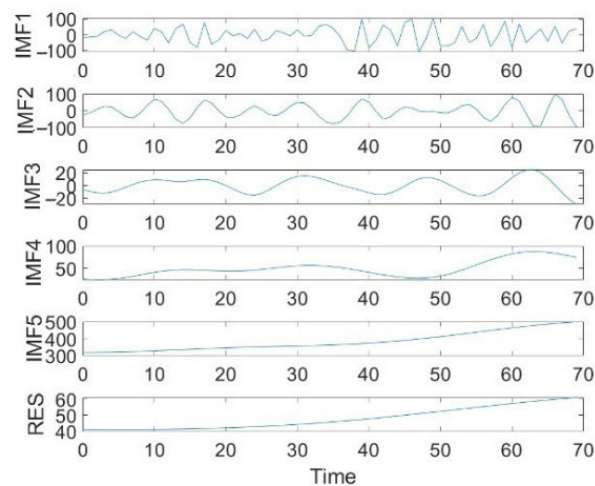**Figure 5.** EMMD-LSTM passenger-flow prediction model.

Figure 6. EEMD decomposition results of OD time series of commercial–residential stations.

When constructing LSTM network, the training data was normalized to have zero mean and unit variance. The IMF components obtained by EEMD decomposition were predicted, respectively. In this paper, 20 neurons were set, and the number of neurons in the output layer is 1. Since only the passenger-flow characteristics in the time dimension are considered, the input length at each time step is 1. In order to improve the learning efficiency, Adam optimization algorithm was introduced, and 200 rounds of training were set. The initial learning rate was 0.01. After 100 rounds of training, the learning rate was reduced by multiplying factor of 0.2, so as to ensure that the optimal solution was quickly approached and there was no significant fluctuation.
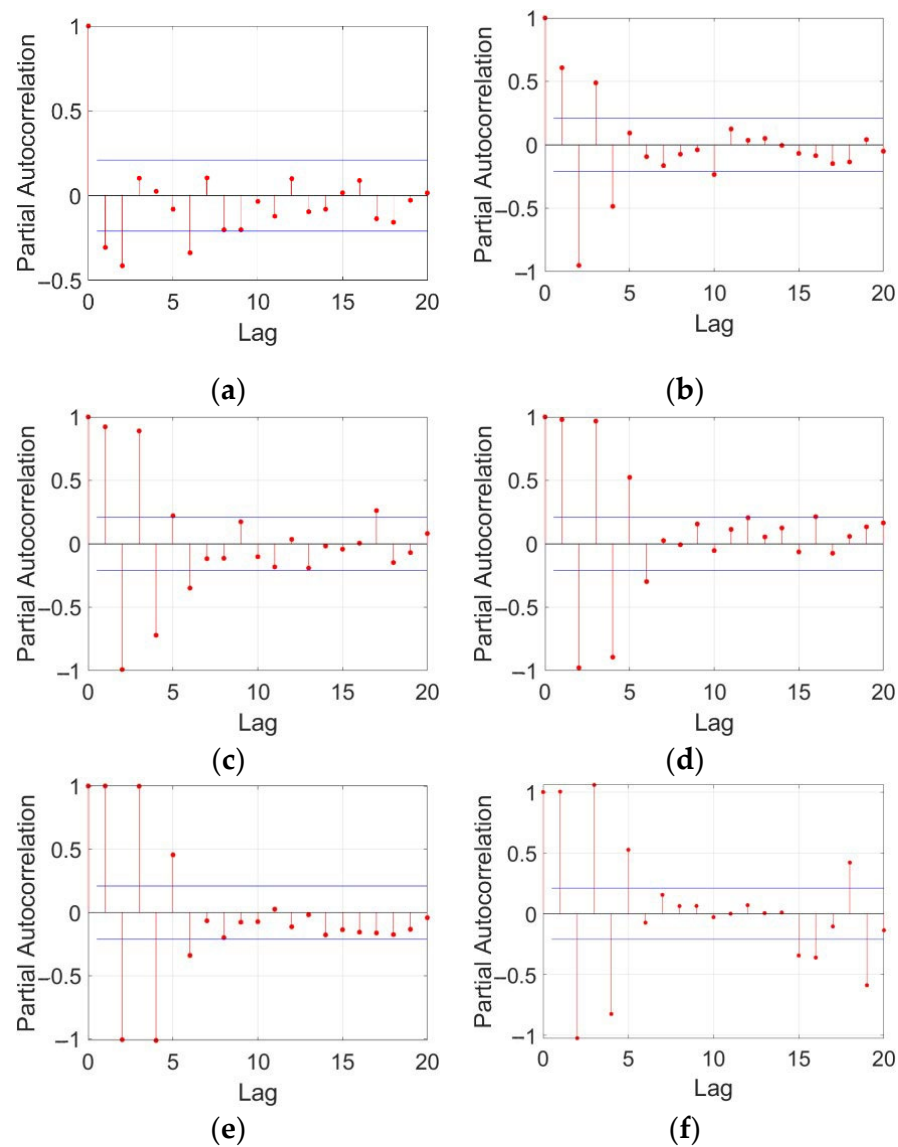
Figure 7. Partial autocorrelation function of intrinsic mode function components and residual terms. (**a**) IMF1; (**b**) IMF2; (**c**) IMF3; (**d**) IMF4; (**e**)IMF5; (**f**) RES.

4. Model Validation

Results and Analysis of Precision

Mean absolute percentage error (MAPE) reflects the relative deviation between the observed value and the true value. It can directly measure the quality of the prediction results and is often used to evaluate the pros and cons of the prediction model, but it cannot directly reflect the difference between the observed value and the true value. Root mean square error (RMSE) can directly reflect the absolute difference between the observed value and the true value, and is very sensitive to large or small errors. It is an effective supplement to MAPE when comparing the accuracy of the model. Therefore, in this paper, MAPE and RMSE were used as evaluation indexes of the prediction model. The EMD-LSTM and EEMD-LSTM were used to predict the intrinsic mode function (IMF) components and residuals, and the prediction results were integrated to obtain the prediction results—the OD passenger flow from 10 June to 30 June. The prediction results were compared with the actual value, as shown in Figure 8. MAPE and RMSE were used to compare the accuracy of the two models, as shown in Table 1.
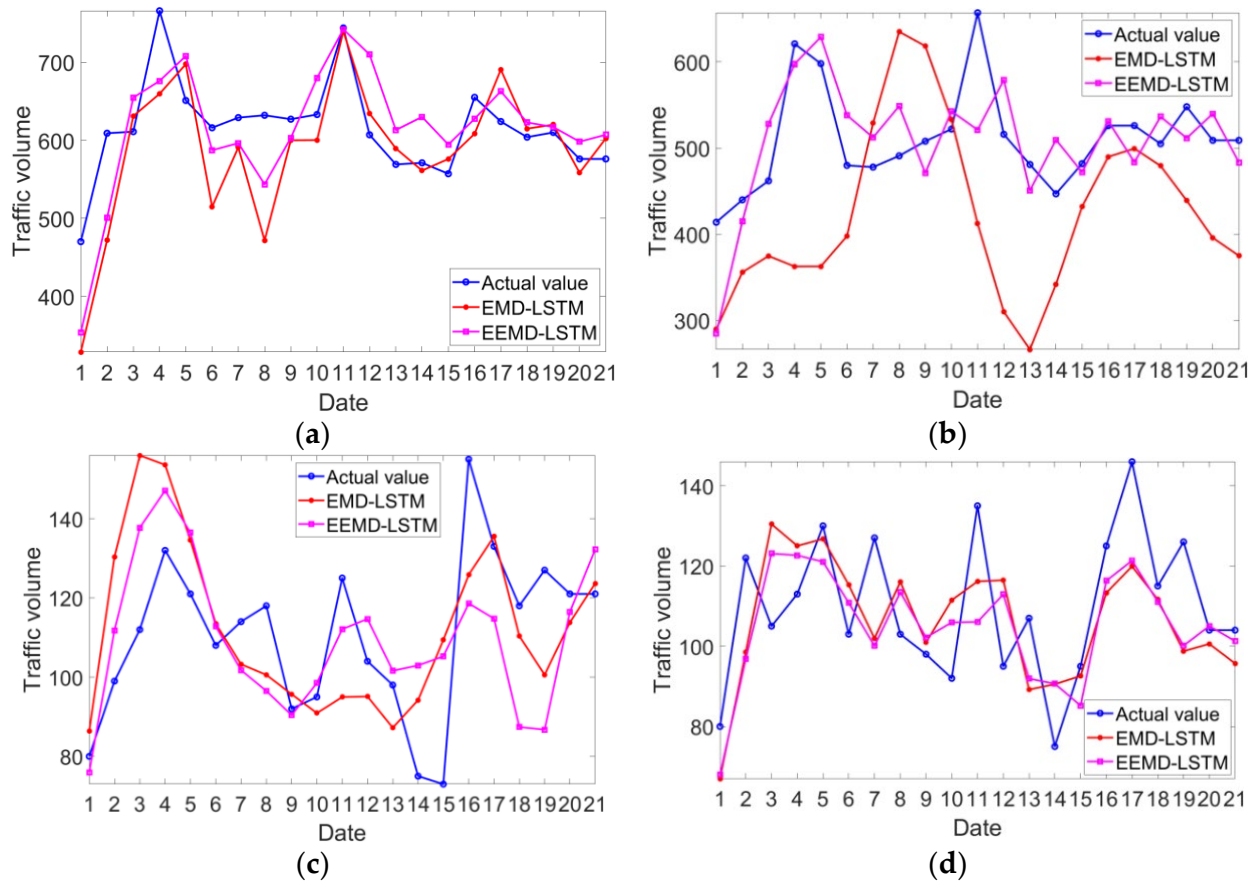
Figure 8. OD passenger-flow prediction results. (**a**) OD prediction of commercial–residential station; (**b**) OD prediction of residential–commercial station; (**c**) OD prediction of scenic–residential station; (**d**) OD prediction of residential–scenery station.

Table 1. Model prediction error comparison.

	Commercial–Residential Station		Residential–Commercial Station		Scenic–Residential Station		Residential–Scenery Station	
	EMD-LSTM	**EEMD-LSTM**	**EMD-LSTM**	**EEMD-LSTM**	**EMD-LSTM**	**EEMD-LSTM**	**EMD-LSTM**	**EEMD-LSTM**
MAPE	8.434	8.158	22.418	9.082	15.195	15.124	13.375	12.556
RSME	69.557	58.842	138.272	56.070	20.174	20.039	16.749	16.171

Note: mean absolute percentage error (MAPE); root mean square error (RMSE); ensemble empirical mode decomposition (EEMD); long short-term memory neural network (LSTM).

According to Table 1, compared with the EMD-LSTM model, the EEMD-LSTM model showed an average reduction of 3.112% in MAPE values and 15.889 in RMSE, indicating that the EEMD-LSTM has higher prediction accuracy. The MAPE and RMSE values of EEMD-LSTM are lower than EMD-LSTM in predicting the OD traffic between commercial–residential and scenic–residential types, indicating that EEMD-LSTM has higher accuracy. Modal aliasing occurs in the EMD-LSTM model prediction of residential–commercial type, resulting in high MAPE. The MAPE value of the OD prediction results of the commercial–residential type are lower than that of the scenic-residential type, indicating that this method is more suitable for the prediction of large traffic volume. The RMSE between the commercial area and residential area is higher than that between the scenic spot and residential area because the passenger flow of the commercial area and residential area were higher in the target day, which makes the error between the predicted value and the real value more sensitive.

In order to further verify the accuracy of the model in practical application, the 35-day historical data of commercial–residential and scenic–residential stations were used as the training set to predict the following 7-day amount, and then the actual data of the latter day were added to the training set. Then, 36-day historical data was used to predict the following 7-day amount, and so on, until the 42-day historical data was used to predict the following 7-day amount, when the prediction accuracy was compared. The prediction results are shown in Figure 9.
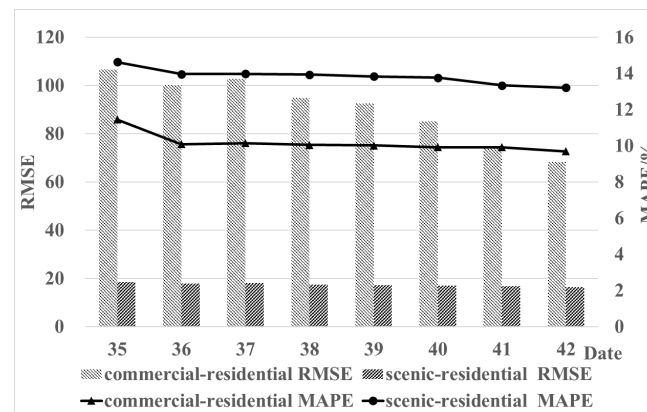

Figure 9. Comparison of prediction accuracy of different historical data.

According to Figure 9, the average MAPE for the 35- to 42-day historical data sample decreased from 13.02% to 10.39% with a decreasing trend. The MAPE and RMSE values of the 42-day sample are the lowest, indicating that its prediction accuracy is the best. That is, with the increase in the number of training-set samples, the prediction error gradually decreases. The reason behind this is that, the more samples of statistical passenger flow there are, the more obvious the passenger-flow law will be, and the predictability of passenger flow will also increase. In practical application, the measured OD value of the next day can be added to the original historical data to predict the future OD value, which can effectively improve the prediction accuracy.

5. Discussion

The traditional passenger-flow prediction model achieves an effect similar to the passenger-flow characteristics by adjusting the neural network parameters [20,21]. EEMD is an improved algorithm for EMD that is prone to modal aliasing and can avoid modal aliasing by adding Gaussian white noise [26]; due to its unique memory forgetting function, LSTM has an advantage over RNN and ARIMA models in dealing with long text data [20]. Although EMD is a flexible and adaptive time-frequency data analysis method and performs good analysis and interpretation effects on nonlinear or non-stationary noise, it also has some defects: EMD does not consider the noise in the original signal that will interfere in actual conditions, so to adopt EMD to decompose signals with noise will give rise to modal aliasing, that is, signals of the same scale or frequency are divided into multiple eigenfunctions. Once the IMFs are aliased, the IMFs are not a single frequency component and cannot accurately obtain the instantaneous frequency, which is mainly related to the frequency characteristics of the original signal and the algorithm of the EMD itself [28]. When the signal is stochastic, the modal aliasing will lead to a large fluctuation in the time–frequency distribution of the EMD decomposition results. The EMD decomposition results are poor when the signal is random. EEMD is an analysis method for the phenomenon of modal aliasing. The main idea for improvement is to insert a random Gaussian white noise sequence into the input signal, so that the original signal is continuous on different characteristic time scales, thereby eliminating the sawtooth lines appearing in the time–frequency distribution.

RNN can integrate historical information and current information well. When it comes to long text data, RNN will experience gradient dispersion and gradient explosion due to the fact that historical data transmission becomes too large or too small. When gradient dispersion occurs, the weights in the RNN will not be updated, which will eventually lead to training failure; when the gradient explosion occurs, the parameters in the RNN will change greatly and the optimal parameters cannot be obtained. In addition, the RNN is in the long text data. There will also be a phenomenon of long-distance dependence, that is, the input at the beginning has less and less influence on the subsequent moments. Compared with RNN, a forgetting unit and memory unit are added to the hidden layer of the LSTM network, which enables the model to memorize or forget information. In the model operation, the network will determine the information that needs to be retained or the information that needs to be discarded, and continue to transmit the information that needs to be retained to the next neuron.

Combining the advantages of EEMD and LSTM, the combined model for short-term passenger-flow prediction of a metro is proposed. The prediction accuracy is compared with the existing research on EMD-LSTM [29] by using AFC data of the Dalian metro. Moreover, most existing studies employ static prediction, that is, using the same size of the training set to predict [18–20]. By changing the size of the training set to achieve the effect of dynamic prediction, EEMD-LSTM is used to compare dynamic and static prediction accuracy. The prediction error of the EMD-LSTM model is 14.855%, and the error of EEMD-LSTM is 11.230%. In contrast, the error of the EEMD-LSTM model is reduced by 3.625%. Using the EEMD-LSTM model and the dynamic prediction method, with the continuous conversion of new data into training sets to participate in training, the error is further reduced from 12.57% to 10.06%, which is a great improvement compared to the original. The results show that EEMD-LSTM has higher prediction accuracy than EMD-LSTM in commercial–residential and scenic–residential OD prediction. On the other hand, by comparing the static prediction without changing the scale of the training set and the dynamic prediction with the gradually increasing scale of the training set, the prediction accuracy of the dynamic prediction is higher.

This study has important theoretical value and practical significance for improving the accuracy of short-term passenger-flow prediction, alleviating traffic congestion and making passengers feel more comfortable, and for the planning and operation of urban rail transit. Accurate passenger-flow prediction can better synergize the metro and other transportation modes, provide theoretical support for the development of overall urban public transportation systems, and realize the sustainable development of cities.

6. Conclusions

1. On the basis of the existing LSTM neural network prediction of short-term passenger flow, EEMD is used to decompose the local characteristic signal of the passenger-flow sequence at the entry and exit stations at different time, so as to weaken the interference of sample noise in the accuracy of the prediction model. With the AFC data of the Dalian metro Line 1 and Line 2 used for testing, the prediction error of the EEMD-LSTM model was reduced by 3.625% on average, compared with that of the EMD-LSTM model, indicating that EEMD-LSTM has higher prediction accuracy.

2. Starting from the 35-day historical data, the OD value of the next 7 days was predicted. The actual amount of the next day was added to historical data, and then the OD value of next 7 days was predicted again. By analogy, until 42 days is taken as historical data, the prediction accuracy of training samples with different historical data was compared. The results show that the average prediction error of historical samples from the 35-day one to the 42-day one decreases from 12.57% to 10.06%, and shows a trend of further decreases, indicating that the dynamic prediction has higher accuracy than the static prediction method by continuously increasing the scale of the training set.

Due to limited research resources and conditions, the long-term OD volume was not selected to verify the accuracy of the model in the case analysis, and the influence of weather, season, the epidemic situation and other factors on passenger flow were not considered. It is one-sided to only consider the comparison with the EMD-LSTM model, but despite this, in terms of error improvement, data preprocessing, etc., the prediction model and general rules constructed in this study can still be used as reference for similar studies. The problem of the small volume of case object can be gradually overcome in subsequent research.

References

1. Xu, X.; Xie, L.; Li, H.; Qin, L. Learning the Route Choice Behavior of Subway Passengers from AFC Data. *Expert Syst. Appl.* **2018**, *95*, 324–332. [CrossRef]
2. Moreira-Matias, L.; Gama, J.; Ferreira, M.; Moreira, J.; Damas, L. Predicting taxi–passenger demand using streaming data. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 1393–1402. [CrossRef]
3. Li, W.; Sui, L.; Zhou, M.; Dong, H. Short-term passenger flow forecast for urban rail transit based on multi-source data. *J. Wirel. Commun. Netw.* **2021**, *9*, 2021. [CrossRef]
4. Wang, Y.; Papageorgiou, M.; Messmer, A. Real-time freeway traffic state estimation based on extended Kalman filter: A case study. *Transp. Sci.* **2007**, *41*, 167–181. [CrossRef]
5. Song, H.; Tao, T.; Li, C.; Ding, Y. Long Time Forecasting of Rail Transit Passenger Volume. In Proceedings of the 2nd International Symposium on Rail Transit Comprehensive Development (ISRTCD) 2014, Beijing, China, 11–12 January 2013; pp. 387–394.
6. Sun, Y.; Leng, B.; Guan, W. A Novel Wavelet-SVM Short-time Passenger Flow Prediction in Beijing Metro System. *Neurocomputing* **2015**, *166*, 109–121. [CrossRef]
7. Li, Y.; Wang, X.D.; Sun, S.; Ma, X.; Lu, G. Forecasting Short-term Metro Passenger Flow Under Special Events Scenarios Using Multiscale Radial Basis Function Networks. *Transp. Res. Part C Emerg. Technol.* **2017**, *77*, 306–328. [CrossRef]
8. Zhang, J.; Chen, F.; Guo, Y.; Li, X. Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit. *IET Intell. Transp. Syst.* **2020**, *14*, 1210–1217. [CrossRef]
9. Leng, B.; Zeng, J.B.; Xiong, Z.; Lv, W.; Wan, Y. Probability Tree Based Passenger Flow Prediction and its Application to the Beijing Metro System. *Front. Comput. Sci.* **2013**, *7*, 195–203. [CrossRef]
10. Ding, C.; Wang, D.G.; Ma, X.L.; Li, H. Predicting Short-term Metro Ridership and Prioritizing its Influential Factors Using Gradient Boosting Decision tree. *Sustainability* **2016**, *8*, 1100. [CrossRef]
11. Vlahogianni, E.I.; Karlaftis, M.G.; Golias, 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]
12. Sun, B.; Cheng, W.; Goswami, P.; Bai, G. Short-term traffic forecasting using self-adjusting k-nearest neighbors. *IET Intell. Transp. Syst.* **2018**, *12*, 41–48. [CrossRef]
13. Zhao, Z.; Chen, W.; Wu, X.; Chen, P.C.Y.; Liu, J. LSTM network: A Deep Learning Approach for Short-term Traffic Forecast. IET Intell. *Transp. Syst.* **2017**, *11*, 68–75. [CrossRef]

14. Ma, X.; Tao, Z.; Wang, Y.; Yu, H.; Wang, Y. Long Short-term Memory Neural Network for Traffic Speed Prediction Using Remote Microwave Sensor Data. *Transp. Res. Part C Emerg. Technol.* **2015**, *54*, 187–197. [CrossRef]
15. Polson, N.G.; Sokolov, V.O. Deep Learning for Short-term Traffic Flow Prediction. *Transp. Res. Part C Emerg. Technol.* **2017**, *79*, 1–17. [CrossRef]
16. Lv, Y.; Duan, Y.; Kang, W.; Li, Z.; Wang, F.-Y. Traffic Flow Prediction with Big Data: A Deep Learning Approach. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 865–873. [CrossRef]
17. Huang, W.; Song, G.; Hong, H.; Xie, K. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks with Multitask Learning. IEEE Trans. Intell. *Transp. Syst.* **2014**, *15*, 2191–2201.
18. Li, Y.; Chen, M.; Zhao, W. Investigating long-term vehicle speed prediction based on BP-LSTM algorithms. *IET Intell. Transp. Syst.* **2019**, *13*, 1281–1290.
19. Zhang, J.; Chen, F.; Shen, Q. Cluster-Based LSTM Network for Short-Term Passenger Flow Forecasting in Urban Rail Transit. IEEE Trans. Intell. *Transp. Syst.* **2019**, *7*, 147653–147671.
20. Li, Y.; Yin, M.; Zhu, K. Short Term Passenger Flow Forecast of Metro Based on Inbound Passenger Plow and Deep Learning. In *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*; IEEE: Piscataway, NJ, USA, 2021; pp. 777–780.
21. Yang, D.; Chen, K.; Yang, M.; Zhao, X. Urban Rail Transit Passenger Flow Forecast Based on LSTM with Enhanced Long-term Features. *IET Intell. Transp. Syst.* **2019**, *13*, 1475–1482. [CrossRef]
22. Jia, F.; Li, H.; Jiang, X.; Xu, X. Deep learning-based hybrid model for short-term metro passenger flow prediction using automatic fare collection data. *IET Intell. Transp. Syst.* **2019**, *13*, 1708–1716. [CrossRef]
23. Wei, Y.; Chen, M. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. *Transp. Res. Part C Emerg. Technol.* **2012**, *21*, 148–162. [CrossRef]
24. Wang, H.; Liu, L.; Dong, S.; Qian, S.; Wei, H. A novel work zone short-term vehicle-type specific traffic speed prediction model through the hybrid EMD–ARIMA framework. *Transp. Metr. B Transp. Dyn.* **2016**, *4*, 1735–1780. [CrossRef]
25. Zhu, C.; Sun, X.; Li, Y. Urban public bicycle traffic demand prediction under station classification. *J. Jilin Univ.* **2021**, *51*, 531–540.
26. Chen, M.; Chen, L.; Wei, Y. Apply ensemble empirical mode decomposition to discover time variants of metro station passenger flow. In Proceedings of the 2017 4th International Conference on Industrial Engineering and Applications (ICIEA), Nagoya, Japan, 21–23 April 2017; pp. 239–243.
27. Cheng, Z.; Zhang, X.; Liang, Y. Railway freight volume prediction based on LSTM networks. *J. Railw.* **2020**, *42*, 15–21.
28. Yang, Y.; Wu, Y. *Application of Empirical Modal Decomposition in Vibration Analysis*; National Defense Industry Press: Beijing, China, 2013.
29. Chen, J.; Cao, L.; Yu, K. EMD-LSTM based deep learning inbound and outbound passenger flow prediction. In Proceedings of the 2021 2nd International Conference on Artificial Intelligence and Information Systems, Chongqing, China, 28–30 May 2021; Volume 174, pp. 1–6.