mathematics

Article

Short-Term Prediction of Time-Varying Passenger Flow for Intercity High-Speed Railways: A Neural Network Model Based on Multi-Source Data

Huanyin Su *, Shanglin Mo and Shuting Peng

School of Railway Tracks and Transportation, Wuyi University, Jiangmen 529020, China
* Correspondence: suhuanyin@csu.edu.cn; Tel.: +86-0750-3296292

Abstract: The accurate prediction of passenger flow is crucial in improving the quality of the service of intercity high-speed railways. At present, there are a few studies on such predictions for railway origin–destination (O-D) pairs, and usually only a single factor is considered, yielding a low prediction accuracy. In this paper, we propose a neural network model based on multi-source data (NN-MSD) to predict the O-D passenger flow of intercity high-speed railways at different times in one day in the short term, considering the factors of time, space, and weather. Firstly, the factors that influence time-varying passenger flow are analyzed based on multi-source data. The cyclical characteristics, spatial and temporal fusion characteristics, and weather characteristics are extracted. Secondly, a neural network model including three modules is designed based on the characteristics. A fully connected network (FCN) model is used in the first module to process the classification data. A bi-directional Long Short-Term Memory (Bi-LSTM) model is used in the second module to process the time series data. The results of the first module and the second module are spliced and fused in the third module using an FCN model. Finally, an experimental analysis is performed for the Guangzhou–Zhuhai intercity high-speed railway in China, in which three groups of comparison experiments are designed. The results show that the proposed NN-MSD model can predict many O-D pairs with a high and stable accuracy, which outperforms the baseline models, and multi-source data are very helpful in improving the prediction accuracy.

Keywords: intercity high-speed railway; passenger flow prediction; multi-source data; neural network model; spatial–temporal fusion characteristics

MSC: 62F99; 65K10

check for updates

Citation: Su, H.; Mo, S.; Peng, S. Short-Term Prediction of Time-Varying Passenger Flow for Intercity High-Speed Railways: A Neural Network Model Based on Multi-Source Data. *Mathematics* **2023**, *11*, 3446. https://doi.org/10.3390/math11163446

Academic Editor: Nadir Farhi

Received: 12 July 2023
Revised: 1 August 2023
Accepted: 7 August 2023
Published: 8 August 2023

1. Introduction

In recent years, the urban agglomeration and metropolitan areas in China have been quickly developing, leading to a rapid rise in the passenger flow between cities. With the high speeds and high frequencies, intercity high-speed railways play an important role in the fast and efficient travel of passengers [1]. The high prediction accuracy of passenger flow is the base for optimizing transportation products of intercity high-speed railways, meeting the high-quality development of the urban agglomeration and metropolitan areas. Hence, it is necessary to study methods for predicting the passenger flow of intercity high-speed railways.

Traditional prediction methods for railway passenger flow are usually based on historical passenger flow time series data, to predict passenger flow at larger time granularities, such as one day, one month, or one year [2–4]. A single factor is considered in these methods. With the fast development of "one hour" metropolitan areas, passengers of intercity high-speed railways place a higher value on their time and the passenger travel demand is characterized by an obvious time fluctuation law. Different departure times in one day are preferred by passengers and such passenger demand is called a time-varying passenger flow [5]. Passenger travel demand shows obvious commuting characteristics, that is, the demands of workdays and weekends are obviously different and cyclical. The passenger demands of different origin–destination (O-D) pairs are also heterogeneous. These characteristics of passenger demand alter under different weather conditions, which has a great impact on passenger travel demand. Hence, for intercity high-speed railways, the prediction accuracy of O-D passenger flow at different times in one day can be improved by considering multiple factors, using multi-source data, and combining passenger demand characteristics.

Since access to the data has been made much easier, the methods for predicting the passenger flow that are based on multi-source data have been a hot research topic in recent years. Spatial and temporal factors are usually integrated and neural networks are designed to predict passenger flow. This research has focused on the prediction of the passenger flow in metro stations [6], railway stations [7] and on-demand ride services [8]. Few methods have solved the O-D passenger flow prediction problems of railways.

Considering the above research gaps, the objective of this paper is to propose a neural network model that considers time, space, and weather factors based on multi-source data, and to predict the time-varying passenger flow of intercity high-speed railways. The main novelty and contributions of this paper are as follows: (1) The distribution characteristics of time-varying passenger flow of O-D pairs are clustered using the *k*-means algorithm. The temporal characteristics and the spatial characteristics of passenger travel demand are fused in the clustering process, that is, the spatial–temporal fusion characteristics. (2) A modular neural network model is designed with the inputs of the cyclical characteristics, the spatial and temporal fusion characteristics, and the weather characteristics of passenger travel demand. Different neural network modules are used based on the input. (3) Multiple baseline methods are used for comparisons, to verify that the proposed neural network model can predict time-varying passenger flow of large sets of O-D pairs with high accuracy under multiple time steps for intercity high-speed railways and that the use of multi-source data yields a clear improvement in the accuracy of this prediction.

The remainder of the paper is arranged as follows. Section 2 summarizes the literature about passenger-flow forecasting. Section 3 analyzes the characteristics of passenger flow based on multi-source data. In Section 4, a neural network model including multiple modules is designed. In Section 5, experiments based on real-word data in China are presented. Finally, conclusions and future research are drawn in Section 6.

2. Related Literature Review

Passenger flow prediction methods have been studied in the past decades [9,10] and applied extensively, such as in the bus transportation system [11], in the metro system [12], in the railway system [13], and in the airline system [14], etc. Various types of models have been proposed to achieve a high prediction accuracy of passenger flow in recent years [15,16].

Initially, prediction models mainly included an exponential smoothing model [17], a regression model [18], a grey prediction model [19], and an autoregressive comprehensive moving average model [20]. These models are parametric forecasting models with simple structures and perform well for linear data. Williams et al. [21] proposed the theoretical basis for modeling univariate traffic condition data streams as seasonal autoregressive integrated moving average (SARIMA) processes, and used this model to predict the passenger flow of a freeway in the short term. Ghosh et al. [22] used an improved SARIMA model to predict the short-term passenger flow of streets. Okutani et al. [23] proposed a prediction method by employing the theory of Kalman filters.

Traffic data are often random, non-stationary, and non-linear, and are hard to predict very well using parametric models. Hence, some researchers use a support vector machine [24], a random forest model [25], and a neural network model [26] to predict passenger flow. These models are non-parametric forecasting models. Neural networks have been the hot topics recently and are presented frequently in the literature as they are more suitable for passenger flow prediction with non-linear and more complex data and perform well [27]. Liu et al. [28] modeled a metro system as graphs with various topologies and proposed a unified Physical–Virtual Collaboration Graph Network to predict the short-term passenger flow of metro stations and O-D pairs. Jing et al. [29] used a BP neural network to predict the short-term passenger flow of intercity high-speed railway stations. Han et al. [30] proposed a novel deep learning-based approach to predict the short-term passenger flow of metro stations. Yu et al. [31] used an artificial neural network model to predict the short-term passenger flow of buses. Wang et al. [32] proposed an improved BP neural network model, and predicted the short-term passenger flow of intercity high-speed railway stations. Liu et al. [33] proposed a deep learning architecture to handle the problem of large-scale fine-grained traffic state prediction. Açıkbaş and Söylemez [34] used an artificial neural network to estimate energy consumption and travel time for mass rail transit lines.

Considering that single prediction methods always have some defects, some hybrid methods have been proposed [35]. A hybrid model can integrate the advantages of multiple models and improve the accuracy of passenger flow forecasting. Wen et al. [36] proposed a short-term prediction method of high-speed railway station passenger flow during holidays. This method used the SARIMA model to predict linear time series, and the non-linear time series were acquired and transformed as feature-label samples with a feature selection to transfer learning. Chen et al. [37] proposed an Empirical Mode Decomposition (EMD)-based Long Short-Term Memory (LSTM) neural network model for predicting short-term metro station passenger flow. Jiang et al. [38] developed a hybrid short-term demand forecasting approach by combining the ensemble empirical mode decomposition and grey support vector machine models to predict the short-term passenger flow of high-speed railway O-D pairs. Zhao et al. [39] proposed a novel hybrid model for short-term high-speed railway passenger demand forecasting that explicitly considered the relevance of neighbor time data. This model was the SSA–WPDCNN–SVR model. Su et al. [5] designed three neural network-based hybrid forecasting models to predict the short-term time-varying passenger flow of high-speed railway O-D pairs, namely, the Variational Mode Decomposition-Multilayer Perceptron, the Variational Mode Decomposition-Gated Recurrent Unit Neural Network, and the Variational Mode Decomposition-Bidirectional Long Short-Term Memory Neural Network. Olayode et al. [40] used a hybrid artificial neural network optimized by particle swarm optimization to predict the traffic flow of long and short trucks.

Since access to data has become much easier in recent years, some prediction methods based on multi-source data have been proposed [41,42]. The prediction accuracy is improved when more information from the multi-source data is used. Li et al. [43] used the SARIMA and support vector machines to establish a metro station passenger flow prediction model. The model was built with intelligent data provided by a large-scale urban traffic flow warning system, such as accurate passenger flow data and weather data. Bei et al. [44] proposed a weather factor model, which was plugged into a macroscopic traffic prediction model, to predict the short-term passenger flow of a freeway. Fu et al. [45] proposed a neural network model for the short-term prediction of metro passenger flow with multi-source data, including smart card data, mobile phone data, and metro network data. Zhang et al. [46] proposed a novel method based on a multi-layer LSTM, which integrated multi-source traffic data and multi-techniques (including feature selection based on Spearman correlation and time feature clustering) to predict the short-term passenger flow of high-speed railway stations. Du et al. [47] proposed a deep irregular convolutional residual LSTM network model called DST-ICRL for urban traffic passenger flow prediction, and they fused other external factors to facilitate a real-time prediction.

In conclusion, the features and deficiencies of passenger flow prediction methods in the field of transportation are shown as follows.

(1) Prediction models have become more complicated, such as hybrid models, deep learning models, and neural network architectures. These models are more suitable for passenger

flow prediction with non-linear and more complex data. The potential information features in the data are extracted and used to improve the prediction accuracy.

(2) The prediction methods mainly focus on predicting the passenger flow of stations and few predict the passenger flow of O-D pairs. Single data are often used in the methods and multi-source data need more attention.

3. Characteristic Analysis of Passenger Flow Based on Multi-Source Data

The travel of passengers on intercity high-speed railways is impacted by many factors, for example, time, space, and the environment. To improve the passenger flow prediction accuracy, the main factors and characteristics of passengers need to be analyzed and extracted. The base data of relevant factors are collected and used to analyze the characteristics of O-D passenger flow based on the Guangzhou–Zhuhai intercity high-speed railway in China, which is located in the Guangdong–Hong Kong–Macao Greater Bay Area with a length of 143 km, 20 stations, and a design speed of 250 km/h, as shown in Figure 1.

Figure 1. Guangzhou–Zhuhai intercity high-speed railway.

3.1. Impact Factor Analysis

(1) Temporal factors

Passenger flow along the intercity high-speed railway is characterized by short distance and commuting. The passenger demand distributions for different departure times in a day and different days in a week show fluctuation rules. To extract the rules, the O-D passenger demand is analyzed statistically in dimensions of the departure time and the day of the week based on the historical ticket reservation data. The O-D passenger demand per hour in each day of one week is calculated, and the passenger demand distributions of Guangzhounan–Zhongshanbei in two successive weeks are shown in Figure 2.

(2) Spatial factors

The O-D passenger flow distribution of the intercity high-speed railway is greatly impacted by spatial factors. The stations of each O-D pair are different in geographical location and technical facilities, resulting in a disparity for the passenger flow of different O-D pairs. Railway stations are often classified into several categories [48], hence the 20 stations of the Guangzhou–Zhuhai intercity high-speed railway being divided in this paper into three grades, as shown in Table 1. The daily arrival and departure passenger flow of each station is distinctly different, as shown in Figure 3. The passenger flow distributions of O-D pairs with different station grades are also different, as shown in Figure 4.

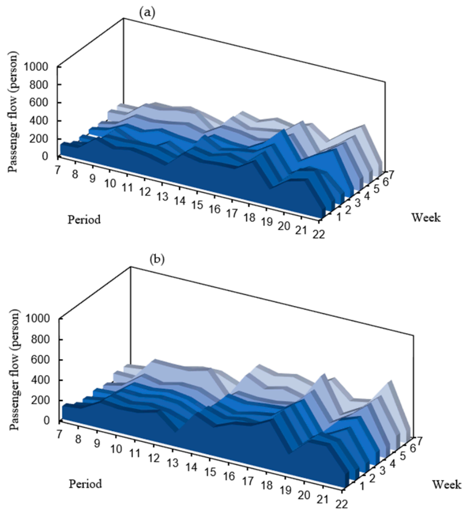


Figure 2. Passenger flow distribution. (**a**) Passenger flow distribution of the first week, (**b**) Passenger flow distribution of the second week.

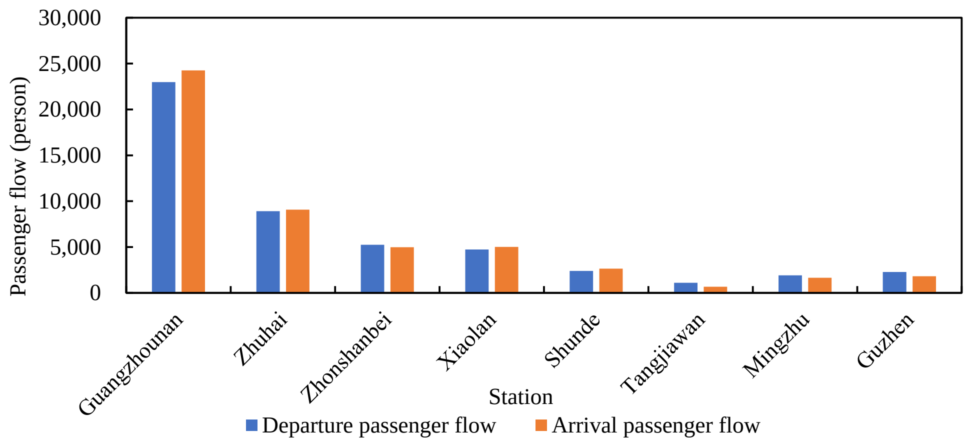


Figure 3. The daily arrival and departure passenger flow of stations.

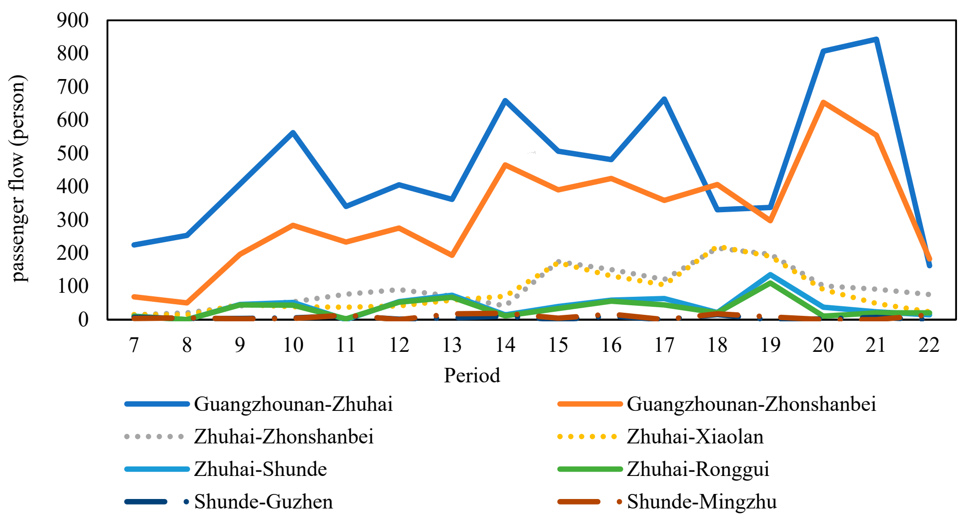


Figure 4. The passenger flow distributions of O-D pairs at different departure times in a day.

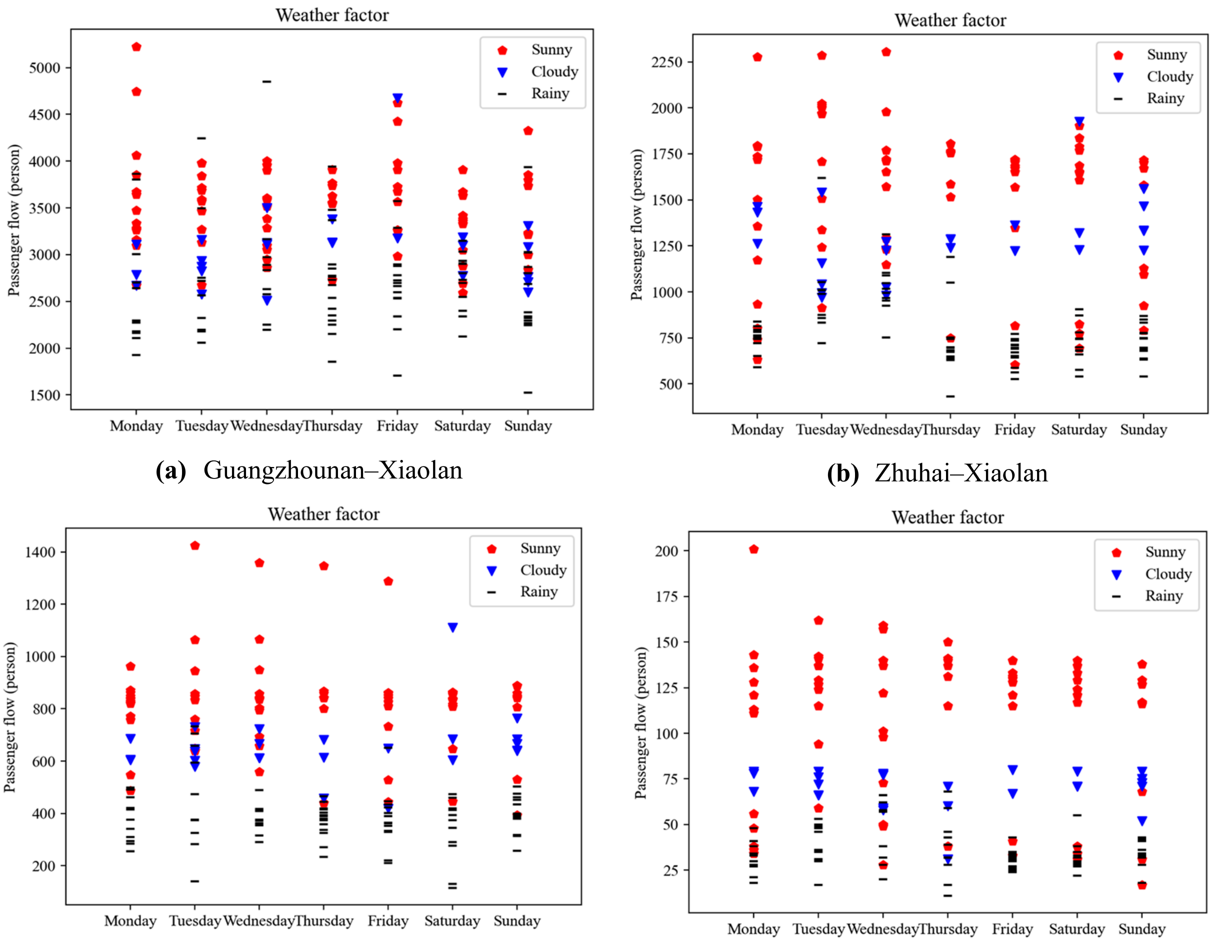
Table 1. The station grades of Guangzhou–Zhuhai intercity high-speed railway.

Level	Station
1	Guangzhounan
2	Zhuhai, Zhonshanbei, Xiaolan
3	Shunde, Ronggui, Mingzhu, Guzhen, Jiangmendon, Xinghui, Bijiang, Beijiao, Shundexueyuan, Nantou, Donsheng, Nanlang, Zhuhaibei, Tangjiawan, Qianshan, Zhonshan

According to Table 1 and Figure 4, the fluctuation rules of the passenger flow distributions at different departure times are similar for O-D pairs with the same arrival and departure station grades, otherwise, they are different. For the ease of use of this characteristic in the subsequent predicting model, it can be extracted based on the arrival and departure station grades and the passenger flow at different departure times.

(3) Weather factors

Besides the spatial and temporal factors, the short-term passenger flow of the intercity high-speed railway is also affected by the environment (such as weather and emergency). The weather of the departure stations of the passengers is mainly considered and analyzed in this paper [49]. The relevant weather and passenger flow data of Guangzhounan–Xiaolan, Zhuhai–Xiaolan, Zhuhai–Shunde, and Shunde–Guzhen from March to August 2015 are analyzed, as shown in Figure 5.



(a) Guangzhounan–Xiaolan **(b)** Zhuhai–Xiaolan **(c)** Zhuhai–Shunde **(d)** Shunde–Guzhen

Figure 5. The passenger flow distributions of O-D pairs under different weather conditions.

According to Figure 5, the passenger flow of O-D pairs on each day of a week is greatly impacted by the weather conditions. The passenger flows on sunny days clearly increase, while those on rainy days clearly decrease, and those on cloudy days stay in the middle. The differences are significant. Hence, the weather conditions are classified into three categories: sunny, cloudy, and rainy.

3.2. Characteristic Expression

For the sake of description, notations are designed as follows. The set of stations is denoted by V. The O-D pair is represented by (r, s), where r is the origin station and s is the destination station. Let RS be the set of all O-D pairs. The grade of r is denoted by C_r and the grade of s is denoted by C_s, with $C_r, C_s \in \{1, 2, 3\}$. The passenger demand before 7:00 and after 23:00 is small, so only the passenger flow between 7:00 and 23:00 is studied. The service time horizon [7:00, 23:00] is divided into 16 periods by hours; for example, [7:00, 7:59] is called period 1, [8:00, 8:59] is called period 2, and so on. For an O-D pair (r, s), the actual passenger flow in the mth period of day n is recorded as $x^m_{n,r,s}$, and the predicted passenger flow is recorded as $\hat{x}^m_{n,r,s}, n = 1, 2 \ldots\ldots N, m = 1, 2, 3, \ldots, 16$, and the average actual passenger flow in the mth period is denoted by $\overline{x}^m_{r,s}$.

The number of influential factors is large, but only a few are suitable for being used in the prediction model. Hence, the key characteristics of passenger flow distribution are extracted by fusing the multi-source data.

(1) Cyclical characteristics

The O-D passenger flow distribution is cyclical with a period of one week. Hence, for an O-D pair, the passenger flows in the same period of the first 14 days are used to predict the passenger flow in the same period of the 15th day, which means that $x^m_{n-14,r,s}, x^m_{n-13,r,s}, \ldots\ldots x^m_{n-1,r,s}$ are used to predict $x^m_{n,r,s}$.

In addition, the demands of workdays and weekends are clearly different, so day n is labeled by WD_n. If it is a workday, then $WD_n = 1$, otherwise $WD_n = 0$.

(2) Spatial and temporal fusion characteristics

The distribution characteristics of passenger flow $x^m_{n,r,s}$ are clustered by the k-means algorithm [50] taking into consideration the departure station's grade C_r, the arrival station's grade C_s and the average passenger flow $\overline{x}^m_{r,s}$. The class of the spatial and temporal fusion characteristics for passenger flow $x^m_{n,r,s}$ is denoted by $C_{r,s,m}$.

(3) Weather characteristics

The weather condition of O-D pair (r, s) in the nth day is labeled by $WEA_{n,r}$, $WEA_{n,r} = 1, 2, 3$, representing, respectively, sunny, cloudy, and rainy.

4. Methodology

In this paper, a neural network model based on multi-source data (NN-MSD) is proposed, taking the cyclical characteristics, spatial and temporal fusion characteristics, and weather characteristics of passenger flow as the input. The input data include two parts. The first part consists of WD_n, $C_{r,s,m}$ and $WEA_{n,r}$, which are classification data. The second part consists of historical time series data of passenger flow, that is, $x^m_{n-14,r,s}, x^m_{n-13,r,s}, \ldots\ldots x^m_{n-1,r,s}$. Selecting a suitable neural network based on the kind of input data can increase the prediction accuracy. The bi-directional Long Short-Term Memory (Bi-LSTM) model performs well when processing time series data and extracting historical characteristics [51]. The fully connected network (FCN) is good at extracting potential relationships between characteristics [52]. Hence, a neural network architecture including 3 modules has been designed, as shown in Figure 6. An FCN model is used in the first module to process the classification data. A Bi-LSTM model is used in the second module to process the time series data. The results of the first module and the second module are spliced and fused in the third module using an FCN model. The details are as follows.

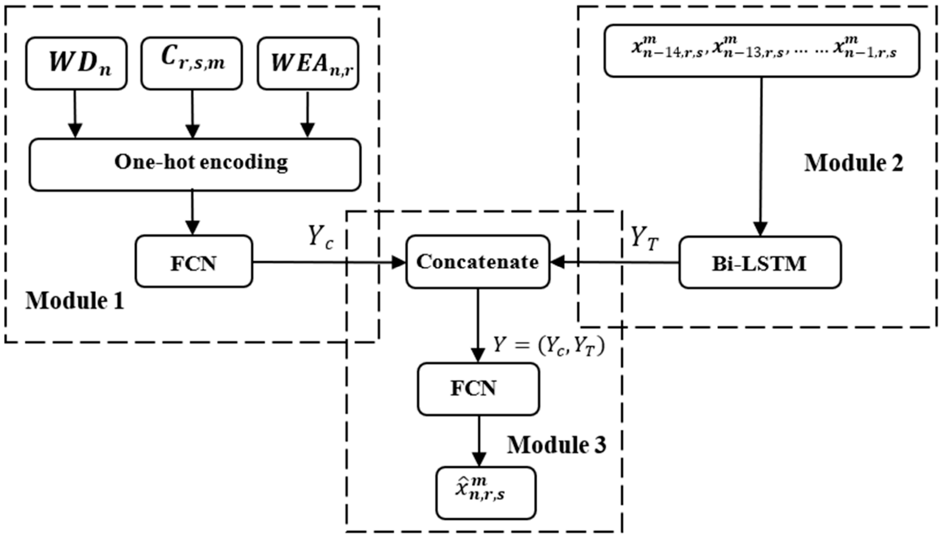


Figure 6. The structure of NN-MSD Model.

The first module: The input data WD_n, $C_{r,s,m}$ and $WEA_{n,r}$ are processed by one-hot encoding [53] and the results are spliced and input into the FCN model. The output is denoted by Y_c.

The second module: time series data $x^m_{n-14,r,s}, x^m_{n-13,r,s}, \ldots\ldots x^m_{n-1,r,s}$ are input into the Bi-LSTM model and the output is denoted by Y_T.

The third module: Concatenate Y_c and Y_T and let $Y = (Y_c, Y_T)$. Input Y into the FCN model and then the output is the predicted passenger flow $\hat{x}^m_{n,r,s}$.

4.1. Bi-LSTM Model

The Bi-LSTM is a model design based on the LSTM model. This model can obtain the data feature information in both directions of the hidden layer in the calculation process, which helps to improve the prediction accuracy. The structure of the Bi-LSTM model is shown in Figure 7, which contains six weight matrices, W_1–W_6. The forward layer performs forward calculation from time 1 to time t to obtain and save data at each time, while the backward layer reverses the calculation from time t to time 1 to obtain and save the data at each time, with h_t as the final output.

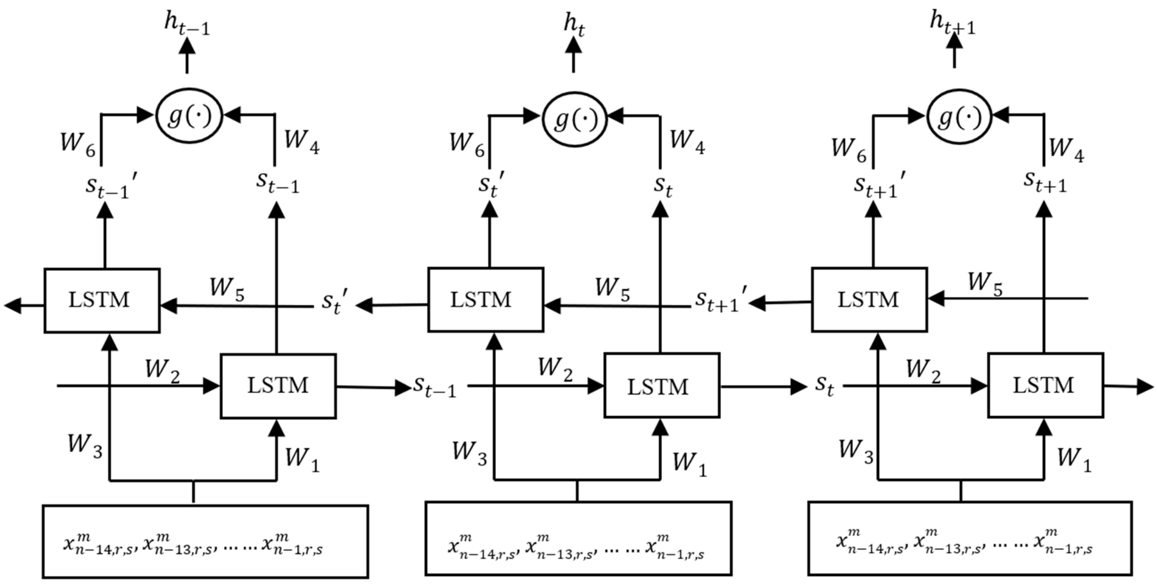


Figure 7. Bi-LSTM structure.

The operation formulas of the Bi-LSTM model are as follows:

$$s_t = f\left(W_1\left[x^m_{n-14,r,s}, x^m_{n-13,r,s}, \ldots\ldots x^m_{n-1,r,s}\right] + W_2 s_{t-1}\right) \quad (1)$$

$$s_{t\prime} = f\left(W_3\left[x^m_{n-14,r,s}, x^m_{n-13,r,s}, \ldots\ldots x^m_{n-1,r,s}\right] + W_5 s_{t+1\prime}\right) \quad (2)$$

$$h_t = g(W_4 s_t + W_6 s_{t\prime}) \quad (3)$$

4.2. FCN Model

The FCN model is a multi-layer perceptron structure. Every node is fully connected with the other nodes in the neighbor layers. An FCN model includes the input layer, several hidden layers, and the output layer. Figure 8 shows the structure of the FCN model with a single hidden layer.

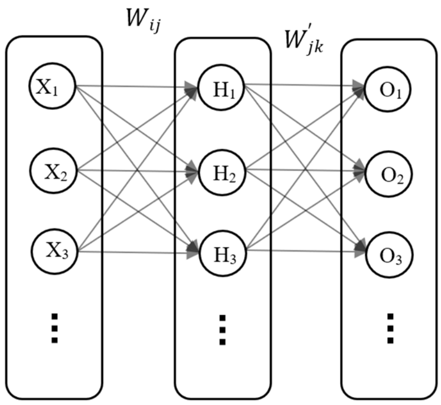


Figure 8. FCN structure.

The operation formulas of the FCN model are as follows:

$$tanh(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}} \quad (4)$$

$$H_j = tanh\left(X_1 \cdot W_{1j} + X_2 \cdot W_{2j} + \cdots + X_U \cdot W_{Uj} + b_j\right) \quad (5)$$

$$O_k = tahn\left(H_1 W'_{1k} + H_2 W'_{2k} + \cdots + H_Q W'_{Qk} + b_k\right) \quad (6)$$

$tanh(\cdot)$ is the activation function. X_i is the ith input in the input layer, a total of U. H_j is the jth neuron in the hidden layer, a total of Q. O_k is the kth neuron in the output layer, a total of P. W_{ij} and W'_{jk} are weights and b is the bias.

5. Experiments and Results

5.1. Sample Setting

The relevant data of the Guangzhou–Zhuhai intercity high-speed railway in China from March to August 2015 are used in the experiments. Data during national legal holidays are deleted and a total of 184 days of the data remain. The first 80% (147 days) of the data set is selected as the training set, and the last 20% (37 days) is selected as the test set. Fifty-three O-D pairs are involved because passenger demand of other O-D pairs is low and insufficient for study. The input time step is 14. The output time steps are 1, 2, and 3. In the first module, the FCN model is set with a single hidden layer consisting of 64 neurons and the output dimension is 3. In the second module, the Bi-LSTM model is set with a single hidden layer, in which the selection range of the number of neurons is (8, 64), and the output dimension is 14. The learning rate is set to 0.01, the loss function is RMSE, the optimizer is Adam, and the range of iterations is (10, 50). The input dimension of the third module is 17, which is processed using an FCN model with a single hidden layer consisting of 32 neurons.

Three commonly used measures of the prediction errors, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), are shown in Formulas (7)–(9), where z is the number of samples in the dataset.

$$RMSE = \frac{1}{53}\sum\nolimits_{(r,s)\in RS}\sum\nolimits_{m=1}^{16}\sqrt{\frac{\sum_{n=1}^{z}\left(x_{n,r,s}^{m}-\hat{x}_{n,r,s}^{m}\right)^{2}}{z}} \tag{7}$$

$$MAE = \frac{1}{z}\sum\nolimits_{n=1}^{z}\frac{1}{53}\sum\nolimits_{(r,s)\in RS}\sum\nolimits_{m=1}^{16}\left|\left(x_{n,r,s}^{m}-\hat{x}_{n,r,s}^{m}\right)\right| \tag{8}$$

$$MAPE = \frac{1}{z}\sum\nolimits_{n=1}^{z}\frac{1}{53}\sum\nolimits_{(r,s)\in RS}\sum\nolimits_{m=1}^{16}\left|\frac{\left(x_{n,r,s}^{m}-\hat{x}_{n,r,s}^{m}\right)}{x_{n}^{m}}\right| \tag{9}$$

The historical ticket reservation data of high-speed trains are added in the dimensions of the departure time and the day of the week and then the time series data $x_{n-14,r,s}^{m}, x_{n-13,r,s}^{m}, \ldots\ldots x_{n-1,r,s}^{m}$ are obtained. WD_n is determined based on the date of day n. The weather condition $WEA_{n,r}$ of an O-D pair (r,s) in the nth day is determined based on the data from https://weather.cma.cn/.

The class of the spatial and temporal fusion characteristics $C_{r,s,m}$ is calculated using the k-means algorithm (the departure station's grade C_r and the arrival station's grade C_s are presented in Section 3.1 and the average passenger flow $\overline{x}_{r,s}^{m}$ can be calculated based on the time series data), where the best clustering number is determined by the Silhouette Coefficient method [54], as shown in Figure 9. When the clustering number k = 10, the contour coefficient reaches its maximum, hence the best clustering number is 10. The clustering results of the spatial and temporal fusion characteristics for the passenger flow of the O-D pairs Guangzhounan–Mingzhu and Xiaolan–Guangzhounan are shown in Table 2.

Table 2. The partial clustering results from the k-means algorithm.

r Is Guangzhounan Station and s Is Mingzhu Station ($C_r = 1$, $C_s = 3$)			**r Is Xiaolan Station and s Is Guangzhounan Station ($C_r = 2$, $C_s = 1$)**		
m	$\overline{x}_{r,s}^{m}$	$C_{r,s,m}$	m	$\overline{x}_{r,s}^{m}$	$C_{r,s,m}$
1	48	7	1	300	6
2	73	7	2	407	4
3	0	3	3	233	6
4	0	3	4	257	6
5	147	5	5	267	6
6	0	3	6	230	6
7	143	5	7	344	9
8	169	10	8	226	10
9	153	10	9	196	10
10	165	10	10	220	10
11	49	7	11	194	10
12	286	6	12	161	10
13	0	3	13	98	5
14	99	5	14	41	7
15	134	5	15	14	3
16	68	7	16	12	3

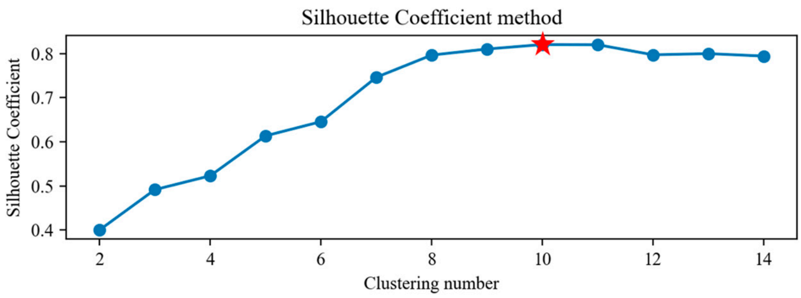


Figure 9. The clustering number determined using contour coefficient method.

5.2. Experimental Results

Based on the above setting, the NN-MSD model is used to predict the time-varying passenger flow of the Guangzhou–Zhuhai intercity high-speed railway. The comparison of the predicted passenger flow from Guangzhounan station to Zhongshanbei station at different times each day in one week with the actual one under different predicting time steps is shown in Figure 10. The prediction errors under different predicting time steps are shown in Table 3.

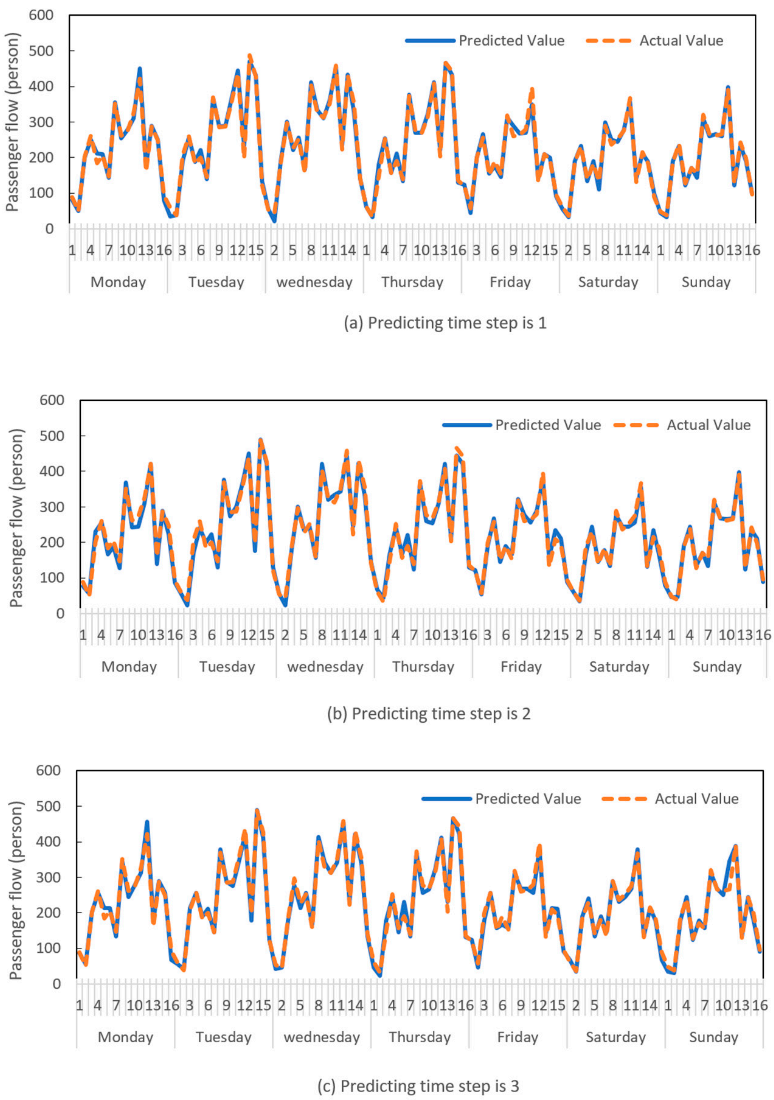


Figure 10. Comparison of the predicted passenger flow at different times every day in one week with the actual one.

Table 3. Prediction errors under different predicting time steps.

Error	Predicting Time Step		
	1	2	3
RMSE	47.4	47.9	45.2
MAE	24.5	27.9	24.1
MAPE	0.06	0.08	0.07

In Figure 10, the distribution characteristics of the predicted passenger flow at different times are consistent with the actual one overall under different predicting time steps, which shows that the NN-MSD model can better fit the time-varying characteristics of passenger flow and are suitable for predicting the time-varying passenger flow of an intercity high-speed railway. In Table 3, the MAPE errors under different predicting time steps are between 6% and 8%, which are quite low compared with existing models in the literature, as shown in Table 4. These verify that the prediction accuracy of the NN-MSD model is high and stable.

Table 4. Prediction accuracy comparison with existing models in the literature.

Literature	Field of Passenger Flow Forecasting	Including Multiple O-D Pairs (Stations) in One Experiment	Multi-Source Data	Method	MAPE
This study	Railway O-D pair	Yes	Yes	NN-MSD	0.06~0.08
Milenkovic et al. [3]	Railway station	No	No	SARIMA	0.04~0.05
Wang et al. [6]	Metro station	No	No	TGACN	0.08
Lin et al. [13]	Railway station	No	No	MLP	0.28~0.36
Liu et al. [28]	Metro O-D pair	Yes	No	PVCGN	0.10~0.19
Han et al. [30]	Metro station	No	No	STGCNNmetro	0.24
Zhao et al. [35]	Metro station	No	No	STL-HW-LSTM	0.06~0.15
Wen et al. [36]	Railway station	No	No	Tra-Decom	0.02~0.05
Jiang et al. [38]	Railway O-D pair	No	No	EEMD-GSVM	0.05~0.07
Zhao et al. [39]	Railway O-D pair	No	No	SSA-WPDCNN-SVR	0.03~0.13
Li et al. [41]	Metro station	Yes	Yes	STLSTM	0.25~0.35
Li et al. [43]	Metro station	No	Yes	SARIMA-SVM	0.09~0.16
Fu et al. [45]	Metro station	No	Yes	NN	0.22~0.30

To illustrate the outstanding performance of the NN-MSD model, three groups of comparison experiments including ten baseline models were performed. The results under prediction time step 1 of the baseline models are shown in Table 5. The structures of the baseline models are as follows.

(1) Baseline models in the first group: Replace the Bi-LSTM model in the second module of the NN-MSD model with the LSTM model [55], the GRU model [56], and the MLP model [57], respectively, and then obtain the baseline models, named, respectively, the NN-MSD-LSTM model, the NN-MSD-GRU model, and the NN-MSD-MLP model. The other parts of the baseline models in the first group are the same as the NN-MSD model.

(2) Baseline models in the second group: Remove the input data WD_n, $C_{r,s,m}$, and $WEA_{n,r}$ from the NN-MSD model separately, and then obtain the baseline models, named, respectively, the NN-MSD-1 model, the NN-MSD-2 model, and the NN-MSD-3 model. The other parts of the baseline models in the second group are the same as the NN-MSD model.

(3) Baseline models in the third group: Take the LSTM model, the GRU model, the MLP model, and the ARIMA model as the baseline models in the third group. Only time series data $x^m_{n-14,r,s}, x^m_{n-13,r,s}, \ldots\ldots x^m_{n-1,r,s}$ are input into the models and the related parameters are the same as the NN-MSD model.

Table 5. Prediction accuracy comparison under prediction time step 1.

Experiment	Input Data	Model	RMSE	MAE	MAPE
The first group	Multi-source data	NN-MSD	47.4	24.5	0.06
		NN-MSD-LSTM	49.1	23.4	0.07
		NN-MSD-GRU	47.1	25.4	0.07
		NN-MSD-MLP	43.4	23.6	0.07
The second group	Remove WD_n from the multi-source data	NN-MSD-1	50.0	29.5	0.08
	Remove $C_{r,s,m}$ from the multi-source data	NN-MSD-2	49.1	29.1	0.08
	Remove $WEA_{n,r}$ from the multi-source data	NN-MSD-3	50.1	24.1	0.07
The third group	Only $\{x^m_{n-14,r,s}, x^m_{n-13,r,s}, \ldots\ldots x^m_{n-1,r,s}\}$	LSTM	61.5	29.2	0.09
		GRU	71.6	32.2	0.10
		MLP	52.0	28.6	0.09
		ARIMA	93.4	69.8	0.35

The following findings are obtained based on the above comparison results in Table 4.

(1) The proposed NN-MSD model outperforms the baseline models. The MAPE error is 6%, obviously lower than the others, and the RMSE and MAE errors are also quite low, which show that the proposed NN-MSD model achieves a high prediction accuracy with multi-source data.

(2) In the second group, the prediction errors increase markedly, and are higher than those of the first group. These verify that the characteristics WD_n, $C_{r,s,m}$, and $WEA_{n,r}$ greatly influence the prediction accuracy. What is more, removing characteristics WD_n and $C_{r,s,m}$ results in higher errors than those when removing characteristics $WEA_{n,r}$, which shows that the cyclical characteristics and the spatial and temporal fusion characteristics have a greater influence on the prediction accuracy.

(3) The prediction errors in the third group are higher than those in the first and second groups, which verifies that using only a single data source yields a low prediction accuracy.

6. Conclusions

In this study, we propose a neural network model, called an NN-MSD model, to predict the time-varying O-D passenger flow of intercity high-speed railways, considering historical ticket reservation data, high-speed railway network data, and weather condition data. The cyclical characteristics, spatial and temporal fusion characteristics, and weather characteristics of passenger travel demand are extracted from the multi-source data. The NN-MSD model includes three modules with these characteristics as inputs. An FCN model is used to process the classification data. A Bi-LSTM model is used to process the time series data. The processed results are spliced and fused using an FCN model. Finally, an experimental analysis is performed regarding the Guangzhou–Zhuhai intercity high-speed railway in China, in which three groups of comparison experiments are designed. The results show the following:

(1) The proposed NN-MSD model can predict the time-varying passenger flow of many O-D pairs for intercity high-speed railways with a high and stable accuracy. The MAPE errors under multiple prediction time steps are between 6% and 8%.

(2) Compared with the baseline models, the prediction accuracy of the NN-MSD model is higher and is influenced greatly by the cyclical characteristics, spatial-temporal fusion characteristics, and weather characteristics. If one of these characteristics is removed from the model, the MAPE error increases markedly.

(3) The MAPE errors of the baseline models with a single data source are between 9% and 10%, significantly higher than that of the NN-MSD model, which shows that

the proposed model achieves a high prediction accuracy due to its use of multi-source data.

In further research, we will study the prediction methods of metro O-D passenger flows considering multi-source data. A metro system has a complex network structure and high travel frequencies of passenger demand. How to extract the spatial–temporal fusion features and improve the prediction accuracy of metro O-D passenger flow deserves more attention.

Author Contributions: Conceptualization, H.S.; data curation, H.S., S.M. and S.P.; formal analysis, H.S., S.M. and S.P.; funding acquisition, H.S.; investigation, H.S., S.M. and S.P.; methodology, H.S., S.M. and S.P.; resources, H.S.; validation, H.S. and S.P.; writing—original draft, H.S.; writing—review and editing, H.S., S.M. and S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the High-Level Personnel Research Start Foundation of Wuyi University (2017RC51) and the Science and Technology Planning Project of Jiangmen (2022JC01004).

Data Availability Statement: Not applicable.

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

References

1. National Development and Reform Commission. Expert Interpretation of the "14th Five Year Plan for the Implementation of New Urbanization" Series III: Optimizing the Spatial Layout and Form of Urbanization to Promote the High Quality Development of New Urbanization in the 14th Five Year Plan [EB/OL]. Available online: https://www.ndrc.gov.cn/xxgk/jd/jd/202207/t20220728_1332352.html?code=&state=123 (accessed on 28 July 2022).
2. Cao, W.X.; Sun, S.L.; Li, H.T. A new forecasting system for high-speed railway passenger demand based on residual component disposing. *Measurement* **2021**, *183*, 109762. [CrossRef]
3. Milenkovic, M.; Svadlenka, L.; Melichar, V.; Bojovic, N.; Avramovic, Z. SARIMA modelling approach for railway passenger flow forecasting. *Transport* **2018**, *33*, 1113–1120. [CrossRef]
4. Wang, H.; Wang, Y.H.; Wu, D.D. A new seasonal cycle GM (1,1) model and its application in railway passenger volume forecasting. *Grey Syst. Theory Appl.* **2022**, *12*, 293–317. [CrossRef]
5. Su, H.Y.; Peng, S.T.; Mo, S.L.; Wu, K.X. Neural Network-Based Hybrid Forecasting Models for Time-Varying Passenger Flow of Intercity High-Speed Railways. *Mathematics* **2022**, *10*, 4554. [CrossRef]
6. Wang, Y.G.; Qin, Y.; Guo, J.Y.; Cao, Z.W.; Jia, L.M. Multi-point short-term prediction of station passenger flow based on temporal multi-graph convolutional network. *Physica A* **2022**, *604*, 127959. [CrossRef]
7. Tsai, T.H.; Lee, C.K.; Wei, C.H. Neural network based temporal feature models for short-term railway passenger demand forecasting. *Expert Syst. Appl.* **2009**, *36*, 3728–3736. [CrossRef]
8. Ke, J.T.; Zheng, H.Y.; Yang, H.; Chen, X.Q. Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. *Transp. Res. Part C* **2017**, *85*, 591–608. [CrossRef]
9. Ahmed, M.S.; Cook, A.R. Analysis of Freeway Traffic Time-Series Data by Using Box-Jenkins Techniques. *Transp. Res. Rec.* **1979**, *722*, 361–1981.
10. Williams, B.M.; Durvasula, P.K.; Brown, D.E. Urban Freeway Traffic Flow Prediction: Application of Seasonal Autoregressive Integrated Moving Average and Exponential Smoothing Models. *Transp. Res. Rec.* **1998**, *1644*, 132–141. [CrossRef]
11. Nagaraj, N.; Gururaj, H.L.; Swathi, B.H.; Hu, Y.C. Passenger flow prediction in bus transportation system using deep learning. *Multimed. Tools Appl.* **2022**, *81*, 12519–12542. [CrossRef]
12. Tang, L.; Yang, Z.; Cabrera, J. Forecasting Short-Term Passenger Flow: An Empirical Study on Shenzhen Metro. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 3613–3622. [CrossRef]
13. Lin, L.Z.; Gao, Y.Z.; Cao, B.X.; Wang, Z.F.; Jia, C. Passenger Flow Scale Prediction of Urban Rail Transit Stations Based on Multilayer Perceptron (MLP). *Complexity* **2023**, *2023*, 1430449. [CrossRef]
14. Yu, J.N. Short-term Airline Passenger Flow Prediction Based on the Attention Mechanism and Gated Recurrent Unit Model. *Cogn. Comput.* **2022**, *14*, 693–701. [CrossRef]
15. Jiang, W.W.; Luo, J.Y. Graph neural network for traffic forecasting: A survey. *Expert Syst. Appl.* **2022**, *207*, 117921. [CrossRef]
16. Yong, G.; Lee, G. Trends, Topics, Leaders, Influential Studies, and Future Challenges of Machine Learning Studies in the Rail Industry. *J. Infrastruct. Syst.* **2022**, *28*, 03122001. [CrossRef]
17. Gould, P.G.; Koehler, A.B.; Ord, J.K.; Snyder, R.D.; Hyndman, R.J.; Vahid-Araghi, F. Forecasting time series with multiple seasonal patterns. *Eur. J. Oper. Res.* **2008**, *191*, 207–222. [CrossRef]
18. Hawas, Y.E. Simulation-Based Regression Models to Estimate Bus Routes and Network Travel Times. *J. Public Transp.* **2013**, *16*, 107–130. [CrossRef]

19. Xiao, X.P.; Lu, Y.Y. Grey linear regression model and its application. *Kybernetes* **2012**, *41*, 622–632. [CrossRef]
20. Du Preez, J.; Witt, S.F. Univariate versus multivariate time series forecasting: An application to international tourism demand. *Int. J. Forecast.* **2003**, *19*, 435–451. [CrossRef]
21. 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]
22. Ghosh, B.; Basu, B.; O'Mahony, M. Bayesian time-series model for short-term traffic flow forecasting. *J. Transp. Eng.* **2007**, *133*, 180–189. [CrossRef]
23. Okutani, I.; Stephanedes, Y.J. Dynamic prediction of traffic volume through Kalman filtering theory. *Transp. Res. Part B* **1984**, *11*, 1. [CrossRef]
24. Hu, Y.R.; Wu, C.; Liu, H.J. Prediction of passenger flow on the highway based on the least square support vector machine. *Transport* **2011**, *26*, 197–203. [CrossRef]
25. Liu, L.J.; Chen, R.C.; Zhao, Q.F.; Zhu, S.Z. Applying a multistage of input feature combination to random forest for improving MRT passenger flow prediction. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 4515–4532. [CrossRef]
26. Jiang, W.H.; Ma, Z.L.; Koutsopoulos, H.N. Deep learning for short-term origin-destination passenger flow prediction under partial observability in urban railway systems. *Neural Comput. Appl.* **2022**, *34*, 4813–4830. [CrossRef]
27. Huang, W.X.; Cao, B.K.; Li, X.M.; Zhou, M.L. Passenger Flow Prediction for Public Transportation Stations Based on Spatio-Temporal Graph Convolutional Network with Periodic Components. *J. Circuits Syst. Comput.* **2022**, *31*, 2250134. [CrossRef]
28. Liu, L.B.; Chen, J.W.; Wu, H.F.; Zhen, J.J.; Li, G.B.; Lin, L. Physical-Virtual Collaboration Modeling for Intra- and Inter-Station Metro Ridership Prediction. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 3377–3391. [CrossRef]
29. Jing, Z.C.; Yin, X.L. Neural Network-Based Prediction Model for Passenger Flow in a Large Passenger Station: An Exploratory Study. *IEEE Access* **2020**, *8*, 36876–36884. [CrossRef]
30. Han, Y.; Wang, S.K.; Ren, Y.B.; Wang, C.; Gao, P.; Chen, G. Predicting Station-Level Short-Term Passenger Flow in a Citywide Metro Network Using Spatiotemporal Graph Convolutional Neural Networks. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 243. [CrossRef]
31. Yu, S.Q.; Shang, C.Y.; Yu, Y.; Zhang, S.Y.; Yu, W.L. Prediction of bus passenger trip flow based on artificial neural network. *Adv. Mech. Eng.* **2016**, *8*, 1–7. [CrossRef]
32. Wang, Y.; Cheng, H.; Li, S. Passenger flow forecast model for intercity high speed railway—A neural network-based analysis. *J. Interdiscip. Math.* **2018**, *33*, 1113–1120. [CrossRef]
33. Liu, Y.; Lyu, C.; Zhang, Y.; Liu, Z.Y.; Yu, W.W.; Qu, X.B. DeepTSP: Deep traffic state prediction model based on large-scale empirical data. *Commun. Transp. Res.* **2021**, *1*, 100012. [CrossRef]
34. Açıkbaş, S.; Söylemez, M.T. Coasting point optimisation for mass rail transit lines using artificial neural networks and genetic algorithms. *IET Electr. Power Appl.* **2008**, *2*, 172–182. [CrossRef]
35. Zhao, Y.Y.; Ma, Z.L.; Yang, Y.; Jiang, W.H.; Jiang, X.G. Short-Term Passenger Flow Prediction With Decomposition in Urban Railway Systems. *IEEE Access* **2020**, *8*, 107876–107886. [CrossRef]
36. Wen, K.Y.; Zhao, G.T.; He, B.S. A decomposition-based forecasting mothod with transfer learning for railway short-term passenger flow in holidays. *Expert Syst. Appl.* **2022**, *189*, 116102. [CrossRef]
37. Chen, Q.C.; Wen, D.; Li, X.Q.; Chen, D.J.; Lv, H.X.; Zhang, J.; Gao, P. Empirical mode decomposition based long short-term memory neural network forecasting model for the short-term metro passenger flow. *PLoS ONE* **2019**, *14*, e0222365. [CrossRef] [PubMed]
38. Jiang, X.S.; Zhang, L.; Chen, X.Q. Short-term forecasting of high-speed rail demand: A hybrid approach combining ensemble empirical mode decomposition and gray support vector machine with real-world applications in China. *Transp. Res. Part C* **2014**, *44*, 110–127. [CrossRef]
39. Zhao, S.; Mi, X.W. A Novel Hybrid Model for Short-Term High-Speed Railway Passenger Demand Forecasting. *IEEE Access* **2019**, *7*, 175681–175692. [CrossRef]
40. Olayode, O.I.; Du, B.; Tartibu, K.L.; Alex, J.F. Traffic flow modelling of long and short trucks using a hybrid artificial neural network optimized by particle swarm optimization. *Int. J. Transp. Sci. Technol.* **2023**, *4*, 4.
41. Li, D.W.; Cao, J.M.; Li, R.Y.; Wu, L.F. A spatio-temporal structured LSTM model for short-term prediction of origin-destination matrix in rail transit with multisource data. *IEEE Access* **2020**, *8*, 84000–84019. [CrossRef]
42. Fang, S.; Prinet, V.; Chang, J.L.; Werman, M.; Zhang, C.X.; Xiang, S.M.; Pan, C.H. MS-Net: Multi-source spatio-temporal network for traffic flow prediction. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 7142–7155. [CrossRef]
43. Li, W.; Sui, L.Y.; Zhou, M.; Dong, H.R. Short-term passenger flow forecast for urban rail transit based on multi-source data. *EURASIP J. Wirel. Commun. Netw.* **2021**, *2021*, 9. [CrossRef]
44. Bie, Y.W.; Qiu, T.Z.; Zhang, C.; Zhang, C.B. Introducing Weather Factor Modelling into Macro Traffic State Prediction. *J. Adv. Transp.* **2017**, *4879170*, 15. [CrossRef]
45. Fu, X.; Zuo, Y.F.; Wu, J.J.; Yuan, Y.; Wang, S. Short-term prediction of metro passenger flow with multi-source data: A neural network model fusing spatial and temporal features. *Tunn. Undergr. Space Technol.* **2022**, *124*, 104486. [CrossRef]
46. Zhang, Z.; Cheng, W.; Gao, Y. Passenger Flow Forecast of Rail Station Based on Multi-Source Data and Long Short Term Memory Network. *IEEE Access* **2020**, *8*, 28475–28483. [CrossRef]
47. Du, B.W.; Peng, H.; Wang, S.Z.; Bhuiyan, M.Z.A.; Wang, L.H.; Gong, Q.R.; Liu, L.; Li, J. Deep Irregular Convolutional Residual LSTM for Urban Traffic Passenger Flows Prediction. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 972–985. [CrossRef]

48. Huiling, F.; Lei, N.; Lingyun, M.; Benjamin, R.; Sperry, Z.H. A hierarchical line planning approach for a large-scale high speed rail network: The China case. *Transp. Res. Part A* **2015**, *75*, 61–83.
49. Saneinejad, S.; Roorda, M.J.; Kennedy, C. Modelling the impact of weather conditions on active transportation travel behavior. *Transp. Res. Part D—Transp. Environ.* **2012**, *17*, 129–137. [CrossRef]
50. Sinaga, K.P.; Yang, M.S. Unsupervised K-Means Clustering Algorithm. *IEEE Access* **2020**, *8*, 80716–80727. [CrossRef]
51. Gao, J.C.; Wang, H.Y.; Shen, H.Y. Task Failure Prediction in Cloud Data Centers Using Deep Learning. *IEEE Trans. Serv. Comput.* **2022**, *15*, 1411–1422. [CrossRef]
52. Zhao, J.C.; Deng, F.; Cai, Y.Y.; Chen, J. Long short-term memory—Fully connected (LSTM-FC) neural network for PM2.5 concentration prediction. *Chemosphere* **2018**, *12*, 128. [CrossRef] [PubMed]
53. Cerda, P.; Varoquaux, G. Encoding High-Cardinality String Categorical Variables. *IEEE Comput. Soc.* **2022**, *34*, 1164–1176. [CrossRef]
54. He, H.; Zhao, Z.H.; Luo, W.W.; Zhang, J.H. Community Detection in Aviation Network Based on K-means and Complex Network. *Comput. Syst. Sci. Eng.* **2021**, *39*, 251–264. [CrossRef]
55. Ma, X.L.; Tao, Z.M.; Wang, Y.H.; Yu, H.Y.; Wang, Y.P. 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]
56. Che, Z.P.; Purushotham, S.; Cho, K.; Sontag, D.; Liu, Y. Recurrent Neural Networks for Multivariate Time Series with Missing Values. *Sci. Rep.* **2018**, *8*, 6085. [CrossRef]
57. Bui, D.T.; Tuan, T.A.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378.