

## Research Article

# Short-Term Passenger Flow Forecasting for Rail Transit considering Chaos Theory and Improved EMD-PSO-LSTM-Combined Optimization

Lixin Zhao <sup>1</sup>, Hui Jin <sup>1</sup>, Xintong Zou <sup>1</sup>, and Xiao Liu<sup>2</sup>

<sup>1</sup>Liaoning University of Technology of China, Jinzhou, Liaoning 121000, China

<sup>2</sup>North China University of Technology, Beijing 100000, China

Correspondence should be addressed to Hui Jin; qcjinhui@lnut.edu.cn

Received 4 August 2022; Revised 20 January 2023; Accepted 5 April 2023; Published 3 May 2023

Academic Editor: Ren-Yong Guo

Copyright © 2023 Lixin Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a prediction method based on chaos theory and an improved empirical-modal-decomposition particle-swarm-optimization long short-term-memory (EMD-PSO-LSTM)-combined optimization process for passenger flow data with high nonlinearity and dynamic space-time dependence, using EMD to process the original passenger flow data and generate several eigenmodal functions (IMFs) and residuals with different characteristic scales. Based on the chaos theory, each component of the PSO algorithm was improved by introducing an inertia factor to facilitate the adjustment of its search capability to improve optimization. Each subsequence of the phase-space reconstruction was built into an improved PSO-LSTM prediction model, and the output of each prediction model was summed to determine the final output. Experimental studies were performed using data from the North Railway Station of Chengdu Rail Transit, and the results showed that the proposed model can generate better prediction results. The proposed model obtained root mean square error (RMSE) and mean absolute error (MAE) of 16.0908 and 11.3704, respectively. Compared with the LSTM, the improved PSO-LSTM, the improved EMD-PSO-LSTM, and the model proposed in this paper improved the RMSE values by 25.53%, 29.97%, and 58.76%, respectively, and the MAE values by 30.41%, 40.13%, and 63.08%, respectively, of the prediction results.

## 1. Introduction

At present, urban rail transport continues to develop well, as its speed, capacity, comfort, and safety help it become the main mode of transport for the urban public. Along with the gradually increasing intensity of passenger flow, short-time passenger flow prediction has become particularly important, as accurate traffic passenger flow prediction helps urban traffic managers better plan and manage their resources. Moreover, traffic short-time passenger flow prediction has become of strategic importance for the construction of urbanized intelligent transport systems to relieve traffic pressure, adjust operating times, and plan future construction. It is also the foundation of smart city ambitions and construction.

However, short-time passenger flow prediction can be a particularly challenging problem. Raw traffic-flow data are spatiotemporal data that simultaneously exhibit heterogeneity and correlation, as well as strong nonlinearity and chaos. In addition, most existing research captures relatively few traffic data attributes, resulting in unsatisfactory prediction results. Consequently, real-time accurate short-time passenger flow prediction is critical.

Short-term traffic forecasting models can be broadly classified into four categories—traditional statistical learning algorithms, machine learning models, deep learning algorithms, and combinatorial models. The tasks of statistical learning and machine learning algorithms are similar, which involve inferring model parameters and fitting and predicting data. However, the focus of the two differs—that is, statistical learning algorithms are more concerned with the

confidence of predictions, whereas machine learning algorithms are more concerned with the predictive effects of the model. Statistical learning algorithms include autoregressive-integrated moving average (ARIMA) models [1], seasonal ARIMA models [2], and Kalman filter models [3]. The advantage of these models is that they are simple to operate; however, owing to the complexity of changes in actual passenger flow data, there can be a certain subjective factor in their establishment, which can be easily influenced by a priori assumptions that can be difficult to satisfy in practice, limiting their predictive performance.

Machine learning models include support vector machines (SVMs) [4], artificial neural networks [5, 6], and Bayesian networks [7], which can capture the nonlinear data features of short-time passenger flow using their own learning abilities. Leng et al. [8] established an improved neural network prediction model that was optimized using a genetic optimization algorithm that not only improved the convergence of its search capability but also its prediction accuracy. However, traditional machine learning models cannot effectively process high-dimensional data, and the complex variability of nonlinearities in time series traffic data can be difficult to capture. Moreover, their predictive performance depends on expert experience, and their generalization ability is weak. Consequently, many scholars have researched deep learning models to handle high-dimensional spatiotemporal traffic data.

Based on deep learning algorithms such as recurrent neural networks (RNNs) [9, 10] and long short-term memory (LSTM) neural networks [11], Huang et al. [12] processed information from traffic sequence data using long- and short-time neural networks and gated recurrent units in RNNs, and performed noise reduction of raw passenger flow data using wavelet transforms. Zhang et al. [13] used multigraph convolutional neural networks to explore the spatial features of traffic data. However, deep learning models are prone to overfitting or underfitting [14].

There is also a class of combinatorial models. Zhai et al. [15] proposed a hybrid traffic flow prediction method by combining the  $k$ -nearest neighbor and LSTM algorithms based on the spatiotemporal features of transportation data. Their experimental results showed that their proposed model improved by 12.59% on average compared with the comparison model. Gao et al. [16] proposed a new hierarchical hybrid model to forecast short-term passenger flows, with an average absolute error of approximately 10% in the forecasting results. Moreover, experiments showed that the prediction results of this combined model exhibited greater accuracy.

However, the collected traffic flow data can be disturbed by noise factors, reducing the predictive performance of the models. To minimize the impact of external factors on forecast accuracy, a prediction model that employs chaos theory can directly analyze the intrinsic regularity of traffic flow data through a priori cognition without establishing a subjective model. Consequently, considering the characteristics of nonlinearity and nonsmoothness in urban rail traffic time series data, the empirical mode decomposition (EMD) algorithm—which can be applied to nonlinear and

nonsmooth signal processing methods—was examined for rail traffic short-time passenger flow predictions. The improved complete ensemble EMD with adaptive noise (CEEMDAN) method was proposed and used for the decomposition of highway data by decomposing the time series into different features, which could dramatically reduce the prediction error of the mainstream model. The improved CEEMDAN-fuzzy entropy (FE)-temporal convolutional network model was shown to exhibit high predictive accuracy and strong robustness when using the US101-S highway in California as the research object. Wang Xiao Quan et al. [17] proposed an SVM model for short-time traffic flow prediction, incorporating the principles of chaos theory to map traffic flows into a hyper-dimensional structure by performing phase-space reconstruction based on its nonlinear characteristics, and to calculate the embedding dimension parameter by using the maximum conditional entropy method with a time delay parameter obtained using mutual information techniques. Finally, the reconstructed subseries were used as inputs and predicted using genetic-algorithm optimized support vector regression. Numerical experiments showed that the proposed method exhibited excellent predictive accuracy. Lingling Wu et al. [18] proposed an empirical model decomposition and differential evolution algorithm to optimize the back propagation neural network for a short-time traffic flow prediction model. They used the EMD algorithm to decompose different modal components in the traffic timing data step-by-step, generating a series of eigenmodal functions (IMFs) and residuals at different scales to remove certain noise effects, thereby improving the accuracy of results.

The main contributions of this paper are as follows:

- (1) Given the nonlinear characteristics of traffic flow and the fact that particle swarm optimization (PSO) algorithms usually fall into local optimality, we combined the chaos theory and an improved EMD-PSO-LSTM model to design a short-time passenger flow forecasting method for urban rail transport, and applied it to the field of rail transit passenger flow forecasting for the first time.
- (2) We used the EMD algorithm to decompose the original time series data and perform phase-space reconstruction using chaos theory principles to reconstruct the useful aspects of the EMD to further explore the internal characteristics of traffic flow and improve prediction accuracy. An improved PSO-LSTM prediction model was then developed for each reconstructed subsequence. Owing to the tendency of the original model to fall into a local optimum, we improved it to increase its PSO-seeking search capability, and the predictions of each component were summed to produce the final output.
- (3) We conducted the experiments using a dataset comprising data from the North Railway Station of the Chengdu Municipal Railway to validate the effectiveness of the proposed model. The results showed that the proposed model performed better than existing methods.

## 2. Methods

**2.1. EMD Algorithm.** EMD can be used for analyzing both linear and smooth, and nonlinear and nonsmooth signals. The core of this approach is to gradually smoothen the signal and decompose the vibration modes in the signal into a finite number of components that tend to be smoothed based on different characteristic scales or trends. Moreover, in short-period passenger flow forecasting for rail transport, the transformation of nonlinear and nonsmooth passenger flow signals into linear and smooth signals better reflects their intrinsic physical meaning [18]. Compared with other signal-processing methods, EMD methods are more indirect, intuitive, and adaptive.

EMD divides the raw traffic flow signal into several empirical mode components (IMFs) based on its adaptive timing analysis, each of which contains local features of different feature scales in the previous traffic flow signal, and residuals (RES), which represent the mean or trend in the original traffic flow signal. Each IMF must satisfy two conditions at the same time—that is, the difference between the number of extremes and zeros in the domain of definition should not exceed 1, and the mean value of the upper and lower envelope functions should be 0 [18]. The process is as follows:

**Step 1:** In the original traffic signal, all extreme points are calculated, and the upper and lower envelopes are fitted  $m_1(t)m_2(t)$  with the cubic spline interpolation function to calculate the mean value  $m(t)$  of the upper and lower envelopes. The mean of the original signal envelope is calculated as follows:

$$m(t) = \frac{m_1(t) + m_2(t)}{2}. \quad (1)$$

**Step 2:** Subtracting the original sequence  $m(t)$  from  $X(t)$  to obtain a new sequence, as follows:

$$h_1(t) = X(t) - m(t), \quad (2)$$

where  $h_1(t)$  satisfies the IMF condition, so that the first  $f_1(t)$  IMF component is obtained. If  $h_1(t)$  is still unstable, the abovementioned process is repeated once with  $h_1(t)$  instead of  $X(t)$  until the resulting average envelope tends to 0, defining the  $f_1$  component as  $h_{1k}$ .

$$f_i = h_{ik}. \quad (3)$$

**Step 3:** The original sequence  $X(t)$  is subtracted from the first IMF component to obtain the first  $f_1(t)$  difference sequence with the high-frequency component and removed  $RES_1(t)$ . The above processing of  $RES_1(t)$  is used to obtain a second empirical modal component until it is no longer possible to disaggregate it, with the last one obtained being a residual  $RES_n(t)$ . After decomposition, which represents the actual average trend of the primary series  $X(t)$ , the original sequence  $X(t)$  can be expressed as follows:

$$X(t) = \sum_{i=1}^{n-1} f_i(t) + RES_n(t). \quad (4)$$

**2.2. Chaos Theory.** The study of the chaos theory began in 1980 with the phase-space reconstruction theory proposed by Packard et al. The theory states that the evolution of each component in a chaotic system is jointly dictated by the other individual components of these interactions, and that the variable contains information about the long-term evolution of all the variables in the system. The basic principle of phase-space reconstruction is the delayed embedding theorem proposed by Takens [19]. For chaotic time series, chaotic models can be built and predicted in the so-called phase space, wherein phase-space reconstruction based on the chaos theory is an essential component in the processing of a chaotic time series. There are two key index values in the phase-space reconstruction algorithm—the embedding dimension ( $d$ ) and time delay ( $\tau$ )—and in [19], the parameters of both the embedding dimension and time delay are only proved via theoretical studies, and no specific formula is given. In practical applications, the time delay and embedding dimension parameters should be calculated considering the actual situation because raw traffic flow time-series data are influenced by external variables.

In a chaotic system, a set of observations that vary with time can be obtained by examining them—noted as chaotic time series  $Y_i, i = 1, 2, \dots, n$ —and a set of  $m$ -dimensional vectors can be constructed using the observations, as follows:

$$Y_i = (y_i, y(i + \tau), \dots, y(i + (m - 1)\tau)), \quad (5)$$

where  $\tau = k\Delta t, k = 1, 2, \dots, \Delta t$  denotes the time interval,  $Y_i$  denotes a sample point in the constructed phase space with  $m$  components, and  $m = n - (m - 1)k$ . If the parameters of  $m$  are chosen appropriately, then  $Y_i$  can represent the state in the original system and dynamic characteristics of primary passenger flow data in the multidimensional phase space.

**2.3. LSTM-Based Short-Time Passenger Flow Forecasting for Rail Transit.** LSTM is particularly suitable for processing traffic data sequences with certain time intervals. In the LSTM neural network structure, each neuron comprises three gating units as a solution to the drawback of disappearing gradients owing to a long time series. The LSTM model comprises four main parts—memory cells, forgetting gate, input gate, and output gate [20]—as shown in Figure 1. The LSTM network structure is shown in Figure 2.

Here, the memory cell is the core component of the entire LSTM model used for storing the cell states of past information, and the output of the memory cell at moment  $t$  can be expressed as follows:

$$\begin{aligned} \tilde{C}_t &= \tan h(W_c \cdot [h_{t-1}, Z_t] + b_c), \\ C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \end{aligned} \quad (6)$$

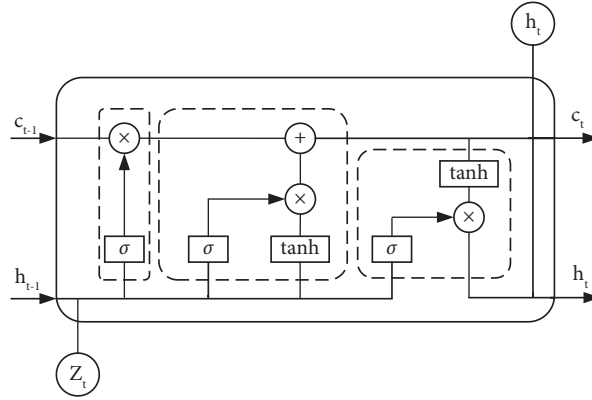


FIGURE 1: LSTM structure.

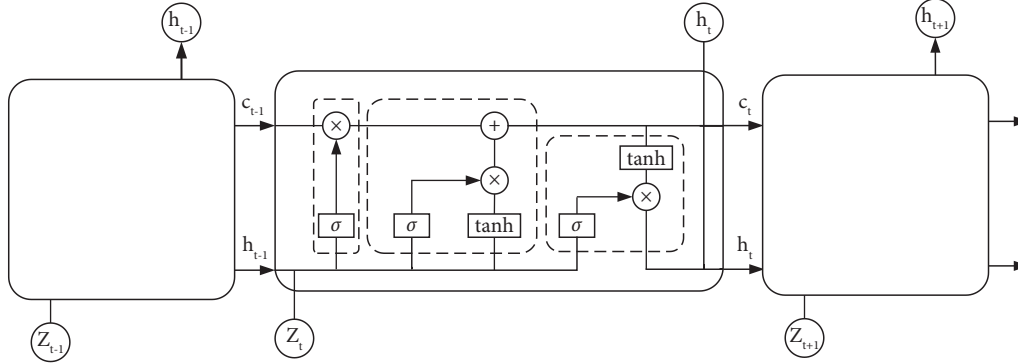


FIGURE 2: LSTM network structure.

where  $Z_t$  denotes the input at the present time,  $i_t$  denotes the input gate, which refreshes the stored information in the cell state, and  $C_{t-1}$  is the update to  $C_t$ .

The forgetting gate is used to determine which part of the information in the cell state needs to be removed by fusing information from the preceding point in time and that on the time at hand. The output of the forgetting gate at moment  $t$  is obtained as follows:

$$f_t = \sigma(W_f \cdot [h_t - 1, Z_t] + b_f), \quad (7)$$

where  $f_t$  denotes the output value of the forgetting gate,  $\sigma$  denotes the activation function,  $W_f$  denotes the weighting matrix representing the forgetfulness gate, and  $b_f$  denotes the bias term for the forgotten door.

Unlike the forgetting gate, the input gate decides which information can enter the unit state based on its threshold, and the candidate value vector is created in the tanh layer to generate candidate memories, wherein new information passing through the screening is added to the unit state to replenish the lost attribute information. In addition, the input gate updates the information stored in the cell state, and its output at time  $t$  is calculated as follows:

$$i_t = \sigma(W_i \cdot [h_t - 1, Z_t] + b_i). \quad (8)$$

The output gate result can be determined via three main components—the previous moment's input information, the information stored after the cell state is updated, and the output information at the last moment. Thus, the output at moment  $t$  can be expressed as follows:

$$\begin{aligned} O_t &= \sigma(W_O \cdot [h_{t-1}, Z_t] + b_O), \\ h_t &= O_t \cdot \tan h(C_t), \end{aligned} \quad (9)$$

where  $W$  denotes the weight value,  $b$  denotes the bias term, and  $\tan h$  denotes the hyperbolic slice employment factor.

**2.4. Improved PSO-LSTM Algorithm for Rail Transit Short-Time Passenger Flow Prediction.** PSO is an exploratory method and a classical swarm intelligence algorithm used for solving the optimal search problem. The principle underpinning this optimization method originated from the search learning of particle foraging behavioral approach, wherein each bird is abstractly viewed as a particle and used to represent a feasible solution [21]. By evaluating the degrees of superiority and inferiority of each particle through a fitness value, a series of random searches are performed, and the current optimal solution search is dynamically tracked by exchanging information with other particles, discovering information, and adaptively changing the

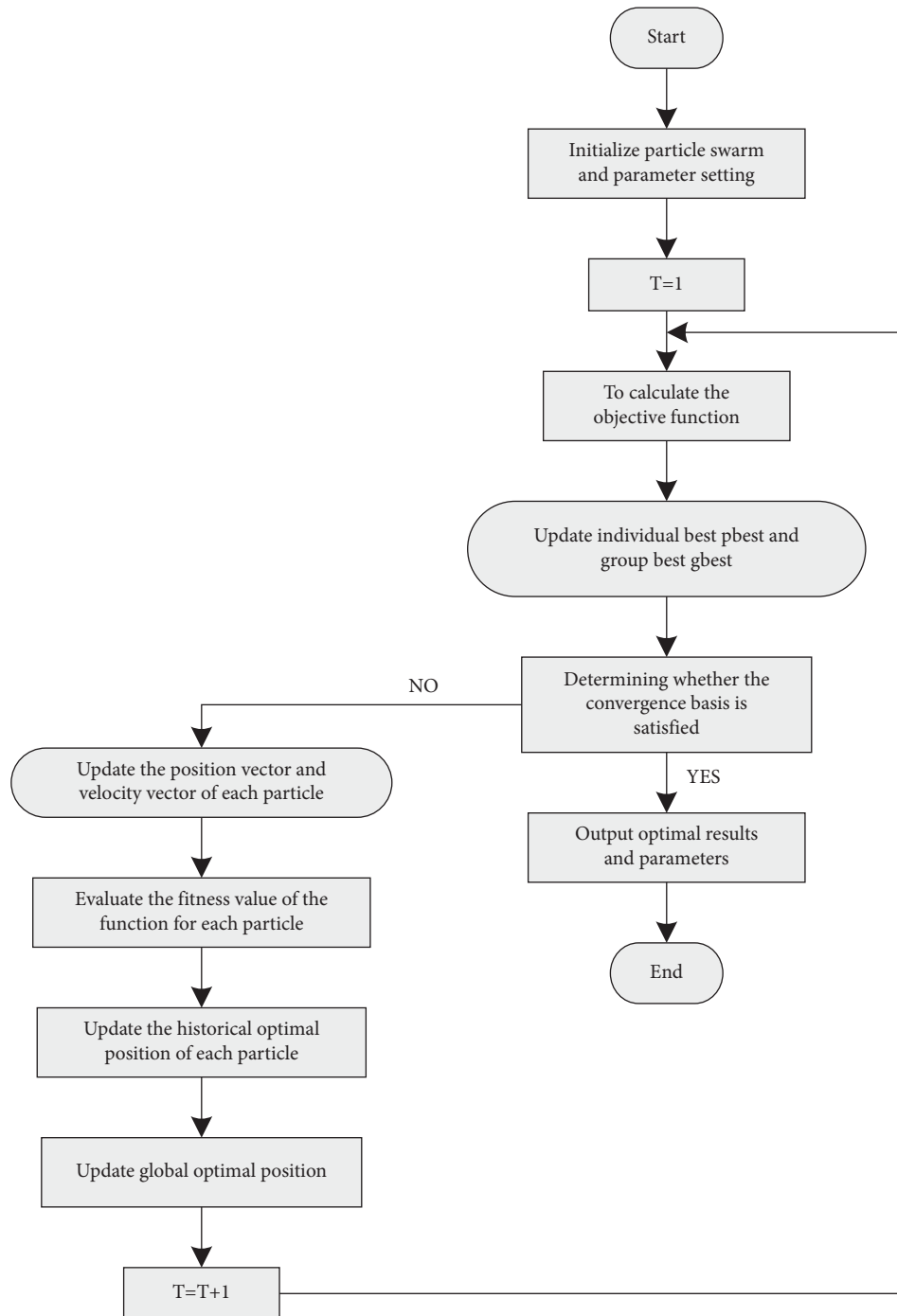


FIGURE 3: Improved PSO solution process.

direction of the next search by collective information sharing such that the group can determine the optimal destination location.

To overcome the shortcomings of traditional PSO algorithms, which can easily fall into optimal solutions, the particle swarm algorithm was improved by introducing an inertia factor, which can reasonably and effectively regulate the global search and partial search capabilities of the algorithm, such that it can change during the PSO search process based on the search

function of the PSO algorithm [22]. When the particles update their velocity and position vectors with each repeated motion, the best value of these two results can be obtained by tracking the positions through which the particles and swarm pass. The specific method can be expressed as follows:

$$\begin{aligned} v_i &= v_i + c_1(pbest_i - x_i) + c_2(gbest_i - x_i), \\ x_i &= x_i + v_i, \end{aligned} \quad (10)$$

where  $i = (1, 2, \dots, N)$ ,  $N$  denotes the total number of particles in the swarm,  $x_i$  denotes the current position of the particle,  $v_i$  denotes the velocity of the particle,  $c_1$  and  $c_2$  denote the learning factors,  $pbest_i$  denotes the most optimum point for a mass to pass through, and  $gbest_i$  denotes the optimal position experienced by the swarm as a whole [20].

$$\begin{aligned} v_i &= w \times v_i + c_1 (pbest_i - x_i) + c_2 (gbest_i - x_i), \\ w &= w_e + (w_s - w_e) \cdot \left(1 - \frac{t_1}{t}\right), \end{aligned} \quad (11)$$

where  $w_s$  denotes the weight at the beginning of the inertia factor,  $w_e$  denotes the weight at the end of the inertia factor,  $t$  denotes the maximum number of iterations after all iterations are completed, and  $t_1$  denotes the number of iterations at the current moment [23].

The improved PSO solution can be obtained by constructing the LSTM short-time passenger flow prediction model and using a modified PSO optimization algorithm to discover the optimal parameters for the LSTM prediction model [24], as shown in Figure 3. The pseudocode of the improved PSO algorithm is shown in Figure 4.

The specific operational steps are as follows:

**Step 1:** Initialize the particle swarm and set the relevant parameters, including the population size, random position, and velocity.

**Step 2:** Determine the fitness value of each particle, as well as the optimal positions for the particle and particle population to pass.

**Step 3:** Determine whether the particles satisfy the convergence condition, and if they do, output the result. If they do not satisfy the convergence condition, continue with the following steps.

**Step 4:** The velocity vector is updated with the optimum positions passed by the particles and particle swarm, and the position vector of the particles is updated with the updated velocity vector, after which all optimal particles are updated.

**Step 5:** Return to Step 3 until the convergence condition is met, before outputting the optimal result and number of iterations.

### 3. Design of Short-Time Passenger Flow Forecasting Algorithm for Rail Transit Based on Chaos Theory and Improved EMD-PSO-LSTM

The EMD algorithm can decompose the time series of rail transit short-time passenger flow into IMF components of different frequencies based on their intrinsic characteristics, which describe the local characteristics of the original series more clearly. An improved PSO-LSTM prediction model can then be separately built for each subsequence, before adding the predicted values of each subsequence to obtain the final output [25]. The model construction process is shown in Figure 5. The pseudocode of the improved EMD-

PSO-LSTM passenger flow forecasting algorithm based on the chaos theory is shown in Figure 6.

**Step 1:** By performing EMD on the rail traffic short-time passenger flow data, several IMFs and RES terms can be obtained.

**Step 2:** The decomposed components are screened, which are then reconstructed in the phase space.

**Step 3:** An improved PSO-LSTM prediction model is developed for each component after phase-space reconstruction, and the improved PSO algorithm is then used to find the optimal parameters and train the LSTM model.

**Step 4:** The predicted values of each component are superimposed and fitted to obtain the final prediction results of the model.

**Step 5:** The final prediction results are output [26].

## 4. Results: Case Studies

**4.1. AFC Data and Processing.** The operational data of the North Railway Station of Chengdu Metro Line 1—that is, its incoming passenger flow data from January 4–13, 2020—were selected for this study. In the Chengdu rail transit system, the AFC platform recorded the entrance and exit information of each passenger using smart card data from the automatic ticketing system at each metro station. Inbound traffic data at the metro stations were obtained between 5:00 and 00:55 the following day, with a data collection interval of 5 min. The data contained a total of 2,260 time series, each of which included the start time, end time, input flow, and output flow. The Chengdu Metro Rail Transit North Station road network map is shown in Figure 7.

A total of 2260 time series of data were input in the experiments of this paper. The input data are inbound passenger flow in person/5 min. The output data are the inbound passenger flow in person/5 min. The simulation environment used to test the predictive performance of the model in this study was MATLAB 2019a. Generally, the larger the training set, the more accurate the prediction results. Therefore, to take full advantage of the data, the first 90% of the original traffic data (that of eight days from January 4–11) were used in the training set, and the remaining 10% (January 12) were used in the test set.

First, anomalies and missing data from the original data were processed, wherein the anomalous data were considered as missing data. Lagrangian interpolation methods were used to process the missing data, wherein four neighboring data before and after the missing datum are selected for interpolation to ensure the reliability of the interpolated data.

The data was then normalized using the min-max method as follows:

$$y^* = \frac{y - y_{\min}}{y_{\max} - y_{\min}}, \quad (12)$$

---

Input: Population size  $M$ , random position  $P$ , velocity  $V$

Output: Outputs optimal results and related parameters

---

- 1 Initialize the particle swarm
- 2 Set population size  $M$ , random position  $P$ , velocity  $V$
- 3 Set the parameter  $T \leftarrow 1$
- 4 Computes the objective function value
- 5 Update individual optimal Update (pbest) , population optimal Update (gbest)
- 6 If (convergence condition satisfied)
- 7 Output optimal results and related parameters
- 8 quit ( )
- 9 Else (does not meet the convergence condition)
- 10 Updates the particle's position vector  $P$  and velocity vector  $V$
- 11 Calculate the adaptation value  $W$  for each particle
- 12 Update the historical optimal position of each particle Update (pbest)
- 13 Update the global optimal location of the population Update (gbest)
- 14 Set the parameter  $T \leftarrow T+1$
- 15 Recalculate the objective function value
- 16 End if

---

FIGURE 4: Pseudocode of the improved PSO algorithm.

where  $y_{\min}$  and  $y_{\max}$  are the minimum and maximum values of the traffic flow, respectively, and  $y^*$  and  $y$  are the traffic flow data before and after normalization, respectively.

**4.2. Prediction Results of the LSTM Model.** The LSTM rail traffic short-time passenger flow prediction model was established, the results of which are shown in Figure 8.

**4.3. Prediction Results of the Improved PSO-LSTM Model.** The prediction results show that the prediction performance of the LSTM model is poor. Consequently, an improved PSO-LSTM model was introduced, the prediction results of which are shown in Figure 9. By running the model through repeated iterations, its prediction results were found to be optimal with an optimal number of hidden nodes of 167, optimal learning rate of 0.0310, and optimal number of iterations of 30.

**4.4. Prediction Results Based on Chaos Theory and Improved EMD-PSO-LSTM Model**

**4.4.1. EMD of Traffic Flow Change Series.** Because the original rail traffic flow data have characteristics of non-linearity and nonsmoothness [27], the noise in it will have some influence on the prediction results, resulting in inaccurate prediction results. The noise in the time series data can be mitigated through the EMD algorithm, thus improving the predictive power of the model. Based on the EMD algorithm, the EMD of the rail traffic flow variation

series can be divided into nine empirical modal components and one residual component. The EMD results are shown in Figure 10.

It can be observed from the figure that the IMF1, IMF2, and IMF3 empirical modal components have higher frequencies and are high-frequency components of the original rail traffic passenger flow data. IMF4, IMF5, and IMF6 empirical modal components have more obvious periodicity and are low-frequency components of the original rail traffic passenger flow data. The residuals are the overall trend of the time series data and are the trend components of the original rail traffic passenger flow data. The EMD of the rail traffic short-time passenger flow time series provides a clearer understanding of the passenger traffic flow data fluctuation and overall trend.

**4.4.2. Phase Space Reconstruction Based on Chaos Theory.** The EMD of the rail traffic flow change sequence can determine the traffic flow fluctuations more accurately. However, because the three components—that is, IMF7, IMF8, and IMF9—do not show the intrinsic properties of the data, the model only selects the remaining seven components, reconstructing them in phase space by finding the time delay using the mutual information method [28] and the embedding dimension using the Cao method. If phase-space reconstruction parameters are carefully selected, the reconstructed phase space can describe the states in the original system, and the multidimensional phase space can show the dynamic characteristics of the traffic flow change sequence [29]. The

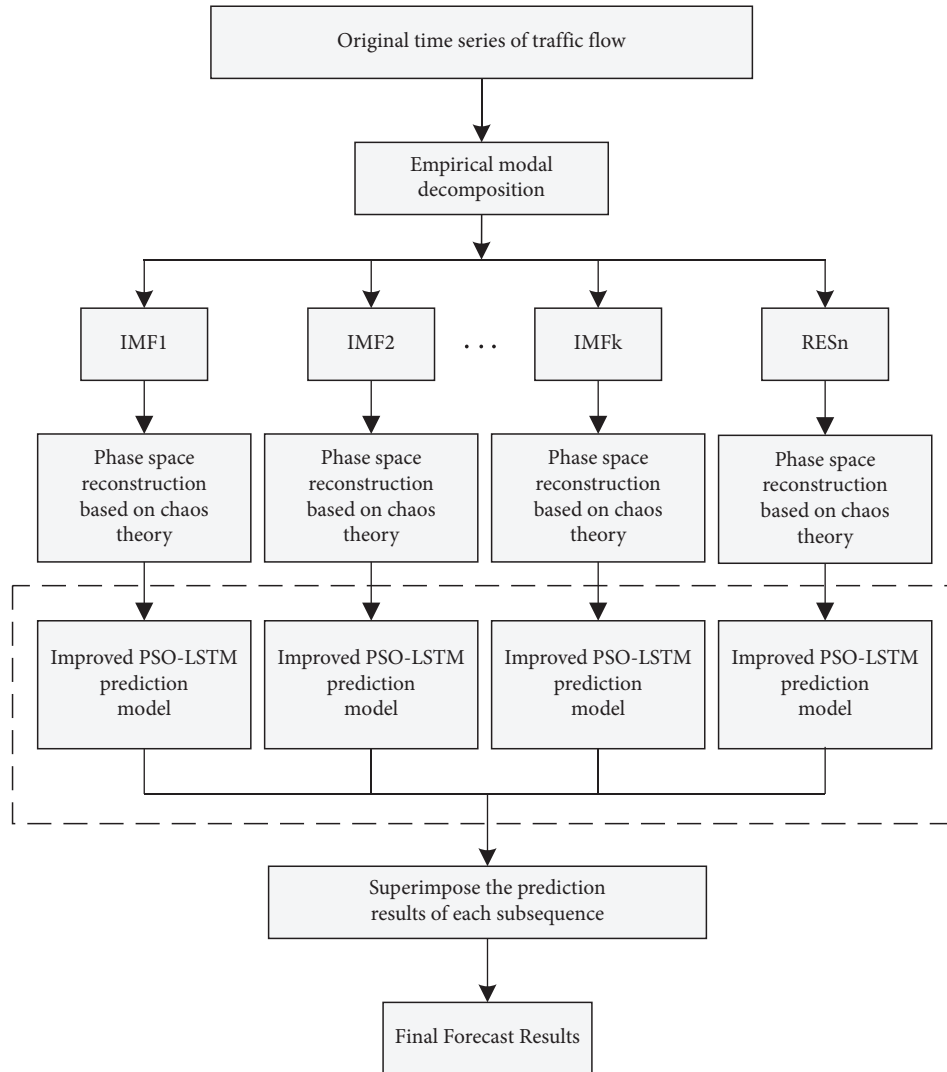


FIGURE 5: Improved EMD-PSO-LSTM model construction process.

remaining seven component phase-space reconstruction parameters are listed in Table 1.

After phase-space reconstruction, the improved PSO algorithm for the components was used to determine the optimal number of hidden nodes, learning rate, and number of iterations of the LSTM prediction model. The optimal parameters of each component are listed in Table 2.

**4.4.3. Prediction Results Based on Chaos Theory and Improved EMD-PSO-LSTM Model.** Owing to the nonlinear and nonsmooth characteristics of the original rail traffic passenger flow data, the noise in its time series data has a certain influence on the prediction results, resulting in inaccurate predictions. Consequently, using the improved PSO-LSTM model, the traffic flow variation sequence could be empirically decomposed. Based on the chaos theory, the decomposed subsequence could be reconstructed in the phase space and an improved EMD-PSO-LSTM-combined optimization model was constructed. The PSO search and prediction results for each component are shown in

Figures 11 more specifically, the superimposed predictions are shown in Figures 11(a)–11(g), which show the plots of the PSO search results for each component, whereas Figures 11(h)–11(n) show the predicted results for each component. Figure 12 shows the predictions for each component overlay.

**4.4.4. Prediction Results of the Improved EMD-PSO-LSTM Model.** To further validate the prediction effect of the combined optimization model based on the chaos theory and the improved EMD-PSO-LSTM model, a set of comparison models were added—that is, phase-space reconstruction of the components without considering the chaos theory. The components obtained after EMD were screened, and the remaining seven components were selected to build the improved PSO-LSTM prediction model. The optimal number of hidden nodes, learning rate, and number of iterations of the LSTM prediction model were determined using the improved PSO algorithm. The optimal parameters of each component are listed in Table 3. The



```

Input: Rail transit short-term passenger flow data
Output: Outputs the final prediction and performs an error analysis
1  Enter the short-term passenger flow data of rail transit
2  EMD empirical modal decomposition of the data
3  For i←0 to n do
4      Modal function INF[i]←Data[i]
5      Filter the individual components after decomposition Choose ()
6      Performing Phase Space Reconstruction Recon ()
7      Establishing an Improved Predictive Model for Components Predict ()
8      Look for the optimal parameter M
9      Train the LSTM model
10 End for
11 For i←0 to n do
12     Overlay the prediction results for each subseries
13     Sum+=Pre[i]
14 End for
15 Outputs the final prediction
    
```

FIGURE 6: Pseudocode of the improved EMD-PSO-LSTM passenger flow forecasting algorithm.

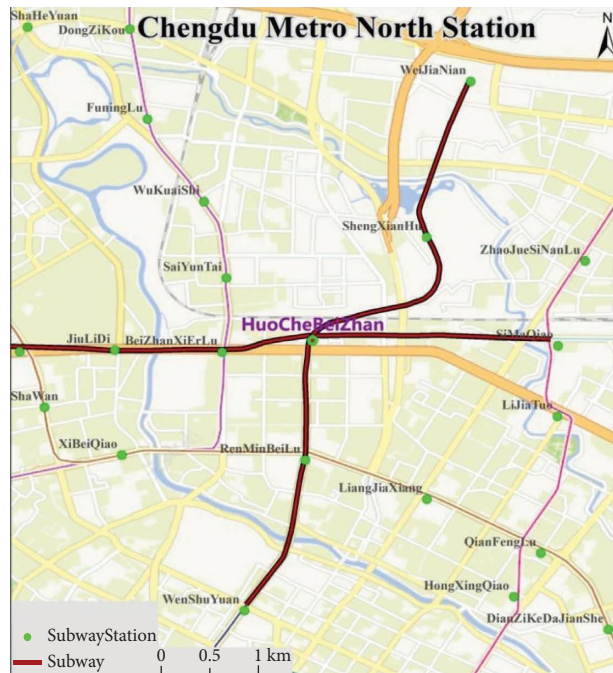


FIGURE 7: Chengdu Metro Rail Transit North Railway Station road network map.

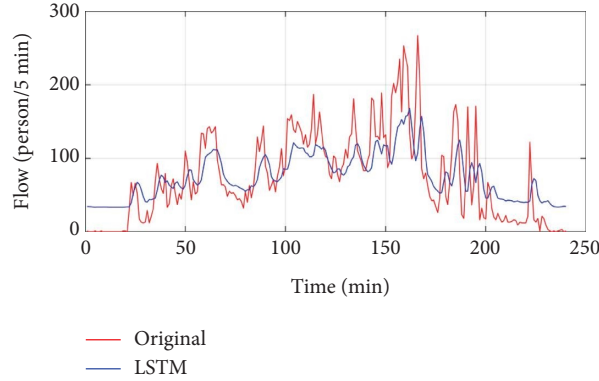


FIGURE 8: LSTM model prediction results.

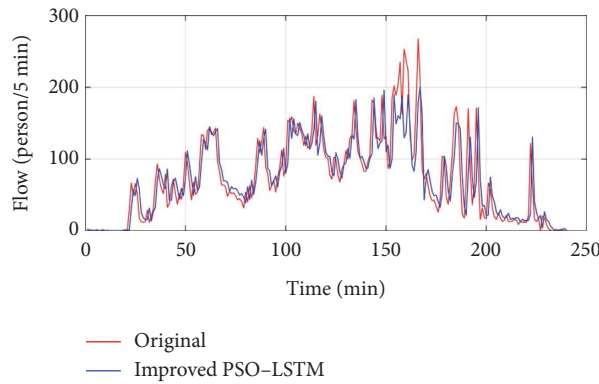


FIGURE 9: Improved PSO-LSTM model prediction results.

prediction results shown in Figures 13(a)–13(g) show the plots of the PSO search results for each component. Figures 13(h)–13(n) show the predicted results for each component. Figure 14 shows the predictions for each component overlay.

**4.5. Evaluation Indicators.** To better compare the predictions between the LSTM model, improved PSO-LSTM model, improved EMD-PSO-LSTM model, and improved EMD-PSO-LSTM model are based on the chaos theory. The mean absolute error (MAE) and root mean square error (RMSE) and the coefficient of determination ( $R^2$ ) metrics were chosen to compare their errors [30]. Their formulas can be expressed as follows:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2}, \\ \text{MAE} &= \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i|, \\ R^2 &= \frac{\sum_{i=1}^n (\hat{x}_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \end{aligned} \quad (13)$$

where  $x_i$  denotes the actual inbound traffic at time  $i$ ,  $\hat{x}_i$  denotes the forecast inbound traffic at time  $i$ ,  $\bar{x}$  indicates the average value of traffic volume, and  $N$  denotes the total volume of inbound traffic in the traffic sequence.

To visually evaluate the prediction results of the four prediction models, the aforementioned errors were used to compare and analyze the strengths and weaknesses of the model predictions [17]. The evaluation indicator values of the prediction results are listed in Table 4, and the percentage improvements in the prediction results are listed in Table 5.

The comparative analysis of the error metrics shown in Tables 4 and 5 indicates that the improved EMD-PSO-LSTM prediction model based on the chaos theory exhibits higher prediction accuracy. In addition, the improved PSO-LSTM and improved EMD-PSO-LSTM models and the improved EMD-PSO-LSTM model based on the chaos theory improve the passenger flow prediction results by using the LSTM model prediction results as a benchmark. The percentage improvement of RMSE values is 25.53, 29.97, and 58.76%, respectively, and the percentage improvement of MAE values is 30.41, 40.13, and 63.08%, and the percentage improvement of  $R^2$  values is 13.36, 16.31, and 32.30%. These results indicate that the proposed model has great potential for short-time passenger flow forecasting applications in rail transit.

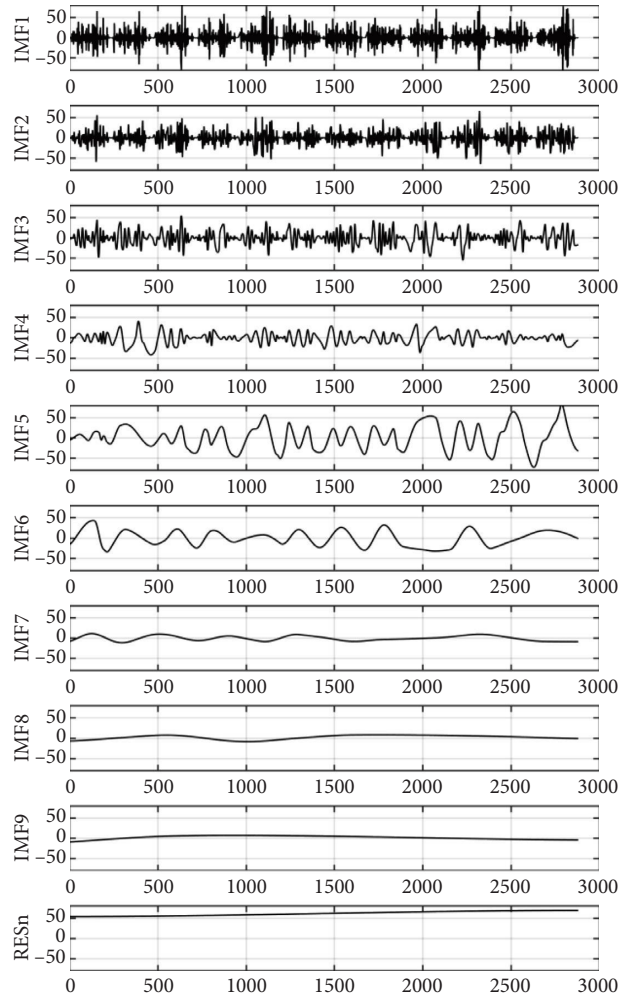


FIGURE 10: EMD results.

TABLE 1: Phase space reconstruction parameters.

Portion size	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	RESn
Time delay $\tau$	99	99	4	8	23	36	99
Embedded dimensions $m$	6	6	6	5	5	4	6

TABLE 2: Optimal parameters of each subsequence after phase-space reconstruction.

Optimal parameters	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	RESn
Hidden nodes	172	184	126	150	200	139	100
Learning rate	0.0051	0.0010	0.0010	0.0010	0.0088	0.0010	0.0010
Number of iterations	30	30	30	30	30	30	30

To further validate the experimental results, the EMD-PSO optimization algorithm based on the chaos theory was compared with combinations of deep learning-based deep belief networks (DBN) and gated recurrent unit (GRU) neural networks. Moreover, combinations of neural network prediction commonly used in the field of traffic flow

prediction, including the model radial basis function neural network (RBF) and multilayer perceptron (MLP), were also compared. The experimental results and error results are presented in Figure 15 and Table 6, respectively.

The experimental results showed that the proposed EMD-PSO-LSTM model of rail transit short-time

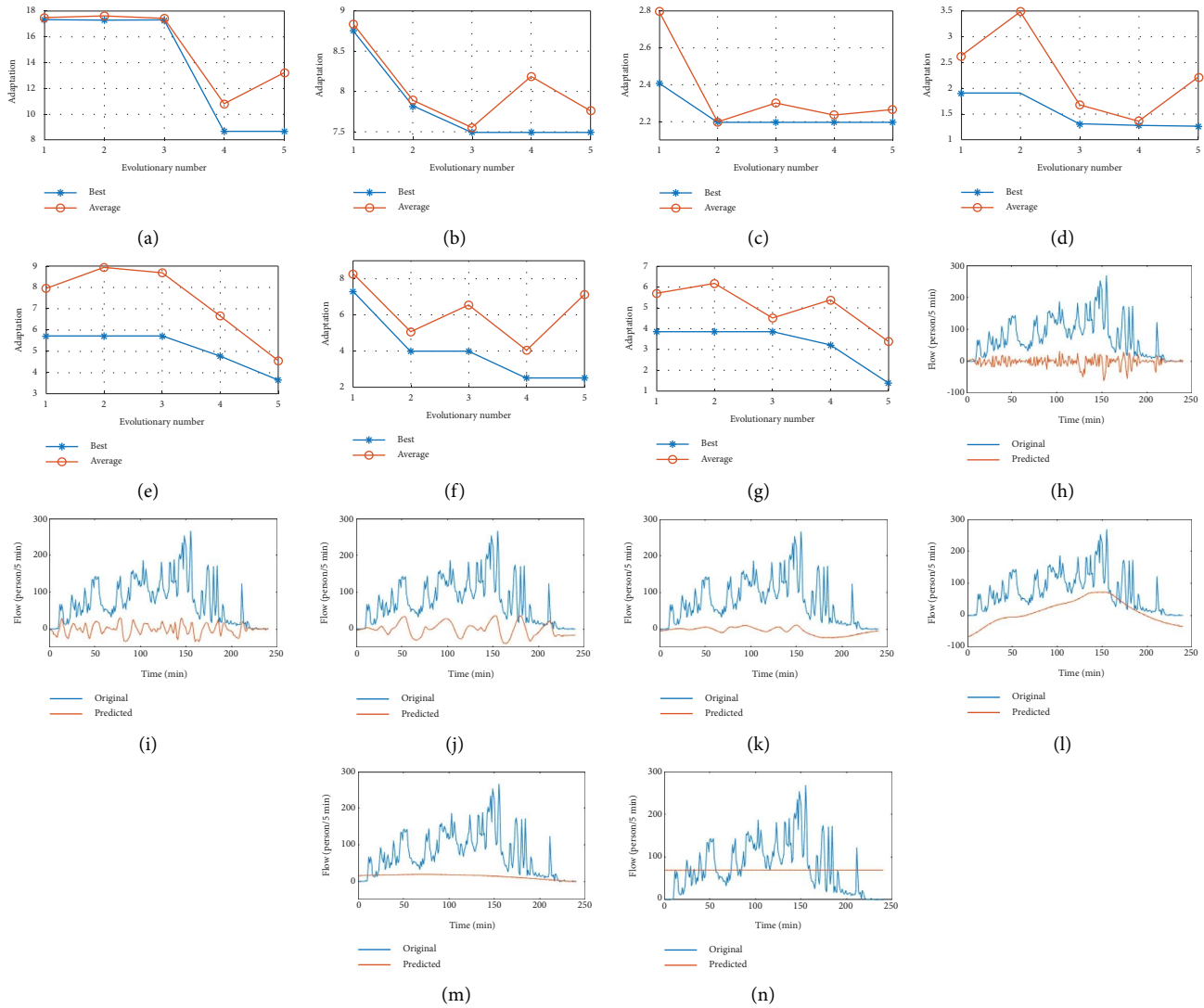


FIGURE 11: Optimization search results and predictions for each component based on chaos theory: (a) IMF1’s search results, (b) IMF2’s search results, (c) IMF3’s search results, (d) IMF4’s search results, (e) IMF5’s search results, (f) IMF6’s search results, (g) RESn’s search results, (h) IMF1’s predicted results, (i) IMF2’s predicted results, (j) IMF3’s predicted results, (k) IMF4’s predicted results, (l) IMF5’s predicted results, (m) IMF6’s predicted results, and (n) RESn’s predicted results.

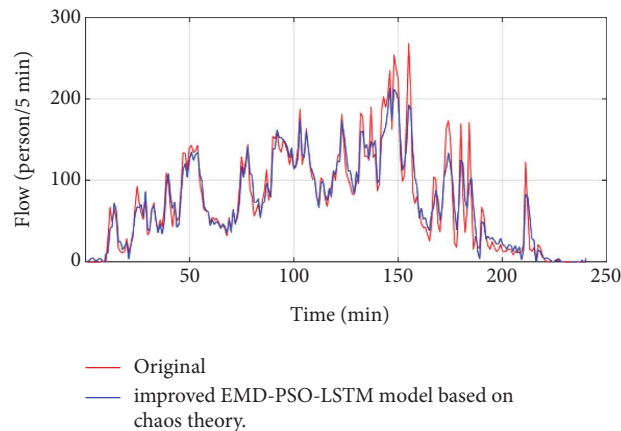


FIGURE 12: Prediction results of the improved EMD-PSO-LSTM model based on chaos theory.

TABLE 3: Optimal parameters of each subsequence after phase-space reconstruction.

Optimal parameters	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	RESn
Hidden nodes	126	116	139	100	182	100	200
Learning rate	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0013
Number of iterations	30	30	30	30	30	30	30

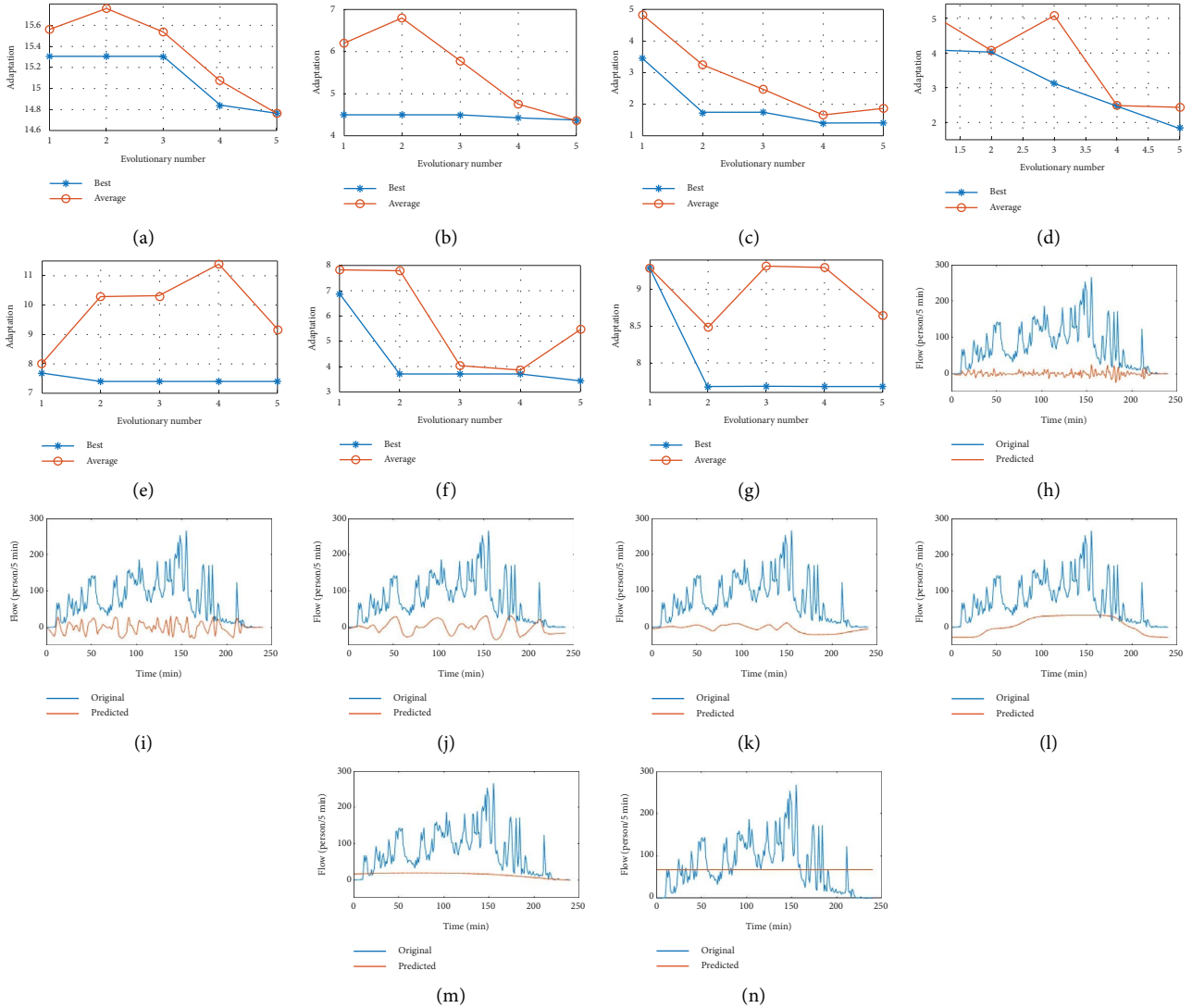


FIGURE 13: Optimization search results and predictions for each component: (a) IMF1’s search results, (b) IMF2’s search results, (c) IMF3’s search results, (d) IMF4’s search results, (e) IMF5’s search results, (f) IMF6’s search results, (g) RESn’s search results, (h) IMF1’s predicted results, (i) IMF2’s predicted results, (j) IMF3’s predicted results, (k) IMF4’s predicted results, (l) IMF5’s predicted results, (m) IMF6’s predicted results, and (n) RESn’s predicted results.

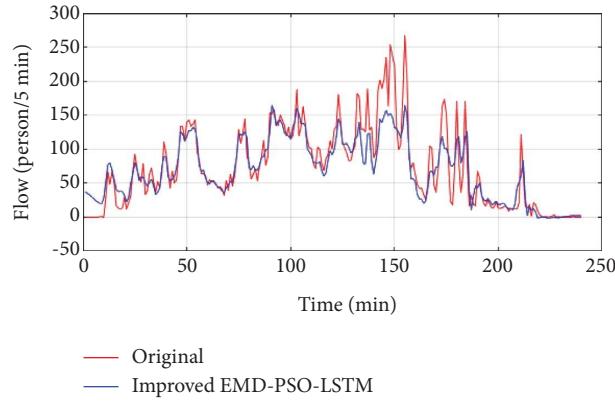


FIGURE 14: Improved EMD-PSO-LSTM model prediction results.

TABLE 4: Results of various models.

Predictive models	Evaluation indicators		
	RMSE	MAE	$R^2$ (%)
LSTM	39.0150	30.7984	71.02
Improved PSO-LSTM	29.0557	21.4331	80.51
Improved EMD-PSO-LSTM	27.3210	18.4379	82.60
Improved EMD-PSO-LSTM based on chaos theory	16.0908	11.3704	93.96

TABLE 5: Percentage improvements in the prediction results.

Predictive models	Evaluation indicators (%)		
	RMSE	MAE	$R^2$ (%)
Improved PSO-LSTM	25.53	30.41	13.36
Improved EMD-PSO-LSTM	29.97	40.13	16.31
Improved EMD-PSO-LSTM based on chaos theory	58.76	63.08	32.30

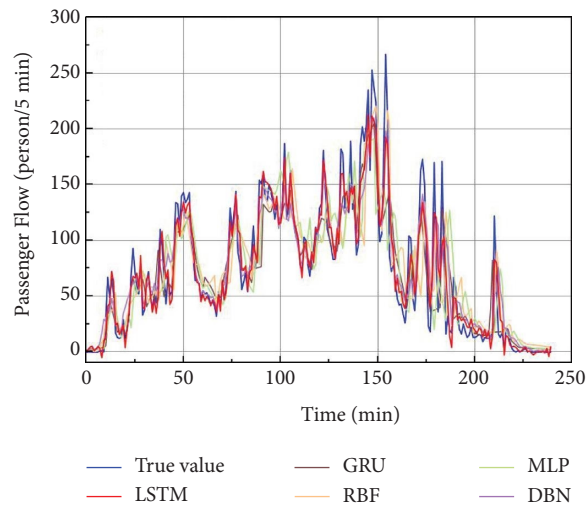


FIGURE 15: Experimental results of the five models.

TABLE 6: Results of various models.

Predictive models	Evaluation indicators			Algorithm run time
	RMSE	MAE	$R^2$ (%)	Training and testing (s)
Improved EMD-PSO-MLP	25.4751	17.5730	84.38	984.65
Improved EMD-PSO-RBF	22.2670	17.6294	86.70	1092.47
Improved EMD-PSO-DBN	17.1435	13.4770	92.15	1168.54
Improved EMD-PSO-GRU	18.5085	14.3520	91.32	945.28
Improved EMD-PSO-LSTM	16.0908	11.3704	93.96	967.53

passenger flow prediction based on the chaos theory obtained better prediction accuracy than the deep belief network (DBN), gated recurrent unit (GRU) neural network, radial basis function (RBF) neural network, and multilayer perceptron (MLP) models.

## 5. Conclusions

The search performance was significantly improved by improving the PSO algorithm. Moreover, by decomposing the time series using EMD and recombining them in a phase-space based on the chaos theory, the characteristic information in the traffic data series could be fully captured, thereby improving the accuracy of the prediction results. It was also found that the chaos theory could further explore the intrinsic characteristics of the time series data, which considerably improved the prediction results. Therefore, combining chaos theory with EMD and an improved PSO-LSTM model for optimization is an effective method for short-time passenger flow forecasting of rail traffic. However, the prediction of short-term traffic flows from the perspective of historical flows alone is somewhat one-dimensional and ignores the impact of data uncertainty. Future research must focus on exploring deep neural architectures to address data uncertainty and making additional improvements to model prediction capabilities. These include rough autoencoder (RAE), interval probability distribution learning (IPDL), and deep temporal dictionary learning (DTDL).

## Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## References

- [1] M. Levin and Y. D. Tsao, "On forecasting freeway occupancies and volumes (abridgment)," *Transportation Research Record*, vol. 773, pp. 47–49, 1980.
- [2] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664–672, 2003.
- [3] J. Guo, W. Huang, and B. M. Williams, "Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 50–64, 2014.
- [4] Y. Zhang and Y. Liu, "Traffic forecasting using least squares support vector machines," *Transportmetrica*, vol. 5, no. 3, pp. 193–213, 2009.
- [5] B. L. Smith and M. J. Demetsky, "Short-term traffic flow prediction: neural network approach," *Transportation Research Record*, vol. 1453, pp. 98–104, 1994.
- [6] A. Csikós, Z. J. Viharos, K. B. Kis, T. Tettamanti, and I. Varga, "Traffic speed prediction method for urban networks—an ANN approach," in *Proceedings of the 2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pp. 102–108, Budapest, Hungary, June, 2015.
- [7] J. Wang, W. Deng, and Y. Guo, "New Bayesian combination method for short-term traffic flow forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 79–94, 2014.
- [8] Z. Leng, J. Gao, B. Zhang, X. Liu, and Z. Ma, "Short-term traffic flow forecasting model of optimized BP neural network based on genetic algorithm," in *Proceedings of the 32nd Chinese Control Conference*, pp. 8125–8129, Xi'an, China, July, 2013.
- [9] X. Ma, H. Yu, Y. Wang, and Y. Wang, "Large-scale transportation network congestion evolution prediction using deep learning theory," *PLoS One*, vol. 10, no. 3, Article ID e0119044, 2015.
- [10] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.
- [11] H. Shao and B. H. Soong, "Traffic flow prediction with long short-term memory networks (LSTMs)," in *Proceedings of the IEEE Region 10 Conference*, pp. 2986–2989, Singapore, November, 2016.
- [12] J. Huang, F. Shao, and S. Yang, "Passenger flow prediction based on recurrent neural networks and wavelet transform," *Journal of Physics: Conference Series*, IOP Publishing, vol. 1486, no. 2, Article ID 22021, 2020.
- [13] J. Zhang, F. Chen, Y. Guo, and X. Li, "Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit," *IET Intelligent Transport Systems*, vol. 14, no. 10, pp. 1210–1217, 2020.
- [14] X. Luo, D. Li, Y. Yang, and S. Zhang, "Spatiotemporal traffic flow prediction with KNN and LSTM," *Journal of Advanced Transportation*, vol. 2019, Article ID 4145353, 10 pages, 2019.
- [15] H. Zhai, R. Tian, L. Cui, X. Xu, and W. Zhang, "A novel hierarchical hybrid model for short-term bus passenger flow

- forecasting,” *Journal of Advanced Transportation*, vol. 2020, Article ID 7917353, 16 pages, 2020.
- [16] G. Heyao, J. Hongfei, and Y. Lili, “An improved CEEMDAN-FE-TCN model for highway traffic flow prediction,” *Journal of Advanced Transportation*, vol. 2022, Article ID 2265000, 20 pages, 2022.
- [17] X. Wang, W. Li, C. Yin, S. Zeng, and P. Liu, “A multiple-parameter approach for short-term traffic flow prediction,” *Modern Physics Letters B*, vol. 35, no. 14, Article ID 2150245, 2021.
- [18] L. L. Wu, L. L. Yin, and Q. L. Ren, “A combined EMD and DE-BPNN optimization method for short-time traffic flow prediction,” *Journal of Chongqing University of Technology (Natural Sciences)*, vol. 35, no. 12, pp. 155–163, 2021.
- [19] Y. Wang, C. Li, and L. Zhang, “Short-term data forecasting based on wavelet transformation and chaos theory,” *IOP Conference Series: Materials Science and Engineering*, vol. 242, no. 1, Article ID 12121, 2017.
- [20] Y. Wei and M. C. Chen, “Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks,” *Transportation Research Part C: Emerging Technologies*, vol. 21, no. 1, pp. 148–162, 2012.
- [21] Y. Liu, L. Wang, and L. Zhao, Z. Yu, *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery*, Springer, Cham, Switzerland, 2022.
- [22] T. Wang, Y. Zhang, N. N. Xiong, S. Wan, S. Shen, and S. Huang, “An effective edge-intelligent service placement technology for 5G-and-beyond industrial IoT,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 6, pp. 4148–4157, 2022.
- [23] F. Peng, J. Wen, Y. Zhang, and J. Jin, “Monthly streamflow prediction based on random forest algorithm and phase space reconstruction theory,” *Journal of Physics: Conference Series*, vol. 1637, no. 1, Article ID 12091, 2020.
- [24] Z. Chen, W. Li, X. Shu, J. Shen, Y. Zhang, and S. Shen, “Operation efficiency optimization for permanent magnet synchronous motor based on improved particle swarm optimization,” *IEEE Access*, vol. 9, pp. 777–788, 2021.
- [25] L. Lin, J. C. Handley, Y. Gu, L. Zhu, X. Wen, and A. W. Sadek, “Quantifying uncertainty in short-term traffic prediction and its application to optimal staffing plan development,” *Transportation Research Part C: Emerging Technologies*, vol. 92, pp. 323–348, 2018.
- [26] W. Wang, M. Tong, and M. Yu, “Blood glucose prediction with VMD and LSTM optimized by improved particle swarm optimization,” *IEEE Access*, vol. 8, pp. 217908–217916, 2020.
- [27] P. Jiao, R. Li, T. Sun, Z. Hou, and A. Ibrahim, “Three revised kalman filtering models for short-term rail transit passenger flow prediction,” *Mathematical Problems in Engineering*, vol. 2016, Article ID 9717582, 10 pages, 2016.
- [28] F. Yang, J. Chen, and Y. Liu, “Improved and optimized recurrent neural network based on PSO and its application in stock price prediction,” *Soft Computing*, vol. 27, no. 6, pp. 3461–3476, 2021.
- [29] H. Tian, L. Yang, H. Wang, and K. Li, “Compressor health prediction based on multi-source information fusion and LSTM,” in *Proceedings of the 2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS)*, pp. 877–882, Suzhou, China, May, 2021.
- [30] L. Liu and R. C. Chen, “A novel passenger flow prediction model using deep learning methods,” *Transportation Research Part C: Emerging Technologies*, vol. 84, pp. 74–91, 2017.