WILEY | Hindawi

Research Article

Research on Railway Passenger Volume Forecast Based on the Spline Interpolation and IPSO-Gradient Difference Acceleration Rule

Dingyuan Fan ,[1,2] **Fei Yang** ,[2] **Jinghao Ji** ,[3] **and Zexi Zhang** [4]

[1]*School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, China*
[2]*China Railway Engineering Design and Consulting Group Co., Ltd., Beijing 100055, China*
[3]*School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 610031, China*
[4]*China Railway Economic and Planning Research Institute, Beijing 100038, China*

Correspondence should be addressed to Dingyuan Fan; fandy@my.swjtu.edu.cn

In order to solve the problem that the railway passenger volume data are abnormal due to holidays and major events interfering with the prediction accuracy, the spline interpolation method is introduced to replace the abnormal passenger volume data. In addition, an improved particle swarm optimization (IPSO) is proposed to optimize the gradient difference acceleration law to combine and improve the predicted value and further improve the prediction accuracy of the railway passenger traffic. Finally, taking Beijing as the research object, the Holt exponential smoothing method and the BP neural network are selected to verify the effect of spline interpolation and IPSO-gradient difference acceleration law on prediction accuracy. The research results show that the spline interpolation method has a better prediction effect after processing abnormal passenger traffic data, and the improved particle swarm algorithm also shows better optimization ability and convergence speed when solving the double difference postulate. In comparison with the BP neural network, Holt exponential smoothing, simple averaging, and conventional redifference approaches, the IPSO-redifference acceleration method achieves a superior prediction performance, and the absolute values of the forecast error are reduced by 3.320%, 1.518%, 2.419%, and 0.602%.

1. Introduction

As an important basis for preparing macro development strategy and passenger transportation plan, railroad passenger volume forecasting is of great significance for formulating railroad network planning, designing train operation plans, optimizing passenger transportation product structure, and improving passenger transportation service level. There are more research results on railroad passenger traffic forecasting, which are broadly divided into two categories: one is to introduce corresponding methodologies to improve the model according to the shortcomings of the forecasting model itself in order to improve the forecasting accuracy. Dempster and Bui et al. studied the neural network model and proved that the prediction results of this model are more accurate than those of other models [1, 2]. Aihara et al. successfully combined artificial neural networks with the chaos theory to develop a freight volume forecasting model [3]. Li et al. predicted passenger transport volume on railways based on the Grey–Markov chain model [4]. Qiu et al. forecasted China's railway freight volume based on the combined PSO-LSTM model [5].

Another category is the combined prediction model formed by combining the characteristics of several models. Wang combined the Gray model and the neural network to obtain a Gray neural network model, and successfully predicted the highway passenger flow [6]. Ye et al. combined the Markov and the Gray GM (1, 1) model and established the Gray–Markov forecasting model to forecast the freight volume [7]. Ge et al. studied an ARIMA model and

combined it with the FSVR to propose a hybrid prediction method for high-speed railway passenger traffic forecasting [8].

When social emergencies such as epidemics and major social events occur, scholars choose models with a strong adaptability to passenger flow fluctuations to make relevant predictions. Jiao proposed an improved STL-LSTM model to improve the accuracy of bus passenger flow prediction during COVID-19 [9]. Wang et al. showed that the SARIMA-NAR combined model can be used during the epidemic [10].

Based on the current research findings, when there are sudden fluctuations in the training data caused by unexpected social events, scholars select models with a high adaptability to predict passenger flow fluctuations. However, the impact of sudden increases or decreases in the passenger volume caused by epidemics, major social activities, and other unexpected social events on the changes in the trend and pattern of passenger volume has not been effectively addressed. Consequently, it disrupts the accuracy of the prediction [11]. Sudden events such as the 2003 SARS epidemic in Beijing, the 2020 COVID-19 pandemic, and major holidays all cause sudden fluctuations in passenger volume [12]. When incorporating such data for volume prediction, fluctuations in passenger volume over time may impact the model's training parameters, thus resulting in higher prediction errors [13]. In addition, most of the existing forecasting methods improve the forecasting models according to the characteristics of the research object or apply multiple forecasting models for combined forecasting. The prediction accuracy of the weighted combination of multiple models is too dependent on the prediction effect of the selected prediction model and has a certain degree of uncertainty, especially when one of the models has a better prediction accuracy, while another model has a poor prediction accuracy, then there will be a situation that the combined prediction accuracy is worse [14].

Spline interpolation is based on the segmented low-order polynomial interpolation method, and the interpolation condition and smoothness of the curve are achieved by adjusting the coefficients of the interpolation basis function on each interval, which has a high accuracy and stability, and the fitted curve is smooth and not easy to oscillate [15]. Spline interpolation can better explore the rule of smooth change of railroad passenger volume with time series [16]. Therefore, to address the problem that the irregular changes in railroad passenger volume under social emergencies affect the accuracy of passenger volume prediction, this paper introduces spline interpolation to correct the preprocessing of abnormal passenger volume data caused by major events such as new crown pneumonia and epidemics to solve the problem of interference of abnormal data on model prediction accuracy, and applies the Holt exponential smoothing method and the BP neural network for railroad passenger volume forecasting [17]. To address the shortcomings of the existing combined forecasting methods, the idea of the redifference acceleration rule is proposed to combine the forecast results for correction, and the redifference metric parameters are optimized by the improved

particle swarm algorithm. Finally, combining the spline interpolation method and redifference acceleration rule with the Holt exponential smoothing method and the BP neural network forecasting proposes a railroad passenger traffic forecasting method that can adapt to irregular fluctuations of passenger traffic under unexpected events and canimprove the accuracy of combined forecasting.

2. Triple Spline Interpolation

Triple spline interpolation (spline interpolation for short) is an important method used in numerical analysis for function estimation and numerical fitting, which can explore the variation pattern of data series and estimate the function values of certain interpolation nodes in between [18]. The basic idea of applying the spline interpolation method to solve the problem of abnormal railroad passenger volume data caused by holidays or major events is as follows: assuming that the variation law of the railroad passenger volume forecast data series in the study area in the past years is a cubic polynomial function $S(x)$, and by excluding the abnormal data due to holidays or major events, the interpolation nodes (i.e., time nodes) are noted as x_c, then the function is a cubic polynomial on every small interval $[x_j, x_{j+1}]$ of the interval (a, b). If the node is $a = x_0 < x_1 < \cdots < x_n = b$, then $S(x)$ is said to be a node $x_0, x_1, \cdots, x_n$ on the cubic spline function. If a function value y_j is given at node x_j, then the interpolation condition is satisfied [19].

$$S(x_j) = y_j. \tag{1}$$

Then, we call $S(x)$ three times the spline interpolation function.

The expression of the cubic spline interpolation function is

$$S(x) = a + bx + cx^2 + dx^3, \tag{2}$$

where a, b, c, and d are the parameters to be estimated.

To find out $S(x)$, it is necessary to determine the 4 parameters to be estimated on each small interval $[x_j, x_{j+1}]$. Since $S(x)$ is a second-order derivative continuous on the interval (a, b), the continuity condition is satisfied at node x_j.

$$\begin{aligned} S(x_j - 0) &= S(x_j + 0), \\ S'(x_j - 0) &= S'(x_j + 0), \\ S''(x_j - 0) &= S''(x_j + 0). \end{aligned} \tag{3}$$

The natural boundary condition of $S(x)$ at the interpolation node x_0, x_n is

$$S''(x_0) = S''(x_n) = 0. \tag{4}$$

By combining the interpolation condition, the continuity condition and the natural boundary condition can be solved for $S(x)$. The correction value of the abnormal data due to holidays or major events is $S(x_c)$.

3. IPSO-Redifferential Acceleration Law

3.1. Law of Redifferential Acceleration. The law of redifferential acceleration is an approximation acceleration technique proposed by Liu Hui, a mathematician in the Wei-Jin period of China, to solve higher precision values, the basic principle of which is that in a monotonically bounded approximation series $\{a_n\}$, the existence of $a_1, a_2, \cdots, a_n$ gradually approaches the limiting value a^*, assuming that the approximating series all satisfy the redifference metric δ, the series deviation ratio is a certain value, and the resulting series can be accelerated to approximate the limiting value according to the following formula:

$$\widehat{a}_n = \left(1 + \frac{1}{\delta - 1}\right)a_{n+1} - \frac{1}{\delta - 1}a_n, \tag{5}$$

where $\widehat{a}_n$ is the improved value of a_n and δ is the redifference metric ($\delta > 1$).

The redifference metric is calculated as

$$\delta \approx \frac{a_{n+1} - a_n}{a_{n+2} - a_{n+1}}. \tag{6}$$

According to the redifferential acceleration law, a_{n+2} is the higher accuracy series approximation value, and a_{n+1} and a_n have the next highest accuracy. Applying this idea to the correction of railroad passenger volume forecast results, the following redifference acceleration correction formula is available:

$$\widehat{F} = \left(1 + \frac{1}{\delta - 1}\right)F_b - \frac{1}{\delta - 1}F_a, \tag{7}$$

where $\widehat{F}$ is the improved value of the railroad passenger volume forecast, F_b is the forecast value of model b, which has a higher forecast accuracy, and F_a is the forecast value of model a, which has a lower forecast accuracy than model b.

According to (7), it can be seen that the implementation of the acceleration rule requires the selection of "one main and one auxiliary" forecasting methods for railroad passenger traffic forecasting, and "one main" means the forecasting method with a better forecasting accuracy, and the forecasting result is F_b. The "one auxiliary" is the forecasting method with a poorer forecasting accuracy, and the forecasting result is F_a. The optimization by the redifferential acceleration rule can make the forecast correction value better than the forecasting result of the "one main" model. This design idea maximizes the promotion of the prediction value with a better accuracy and suppresses the prediction value with a poor accuracy, so as to obtain the prediction value with a higher accuracy.

The redifference metric can be calculated according to (6), but this method has two drawbacks: one is that a better or worse accuracy forecasting model is also needed to forecast the railroad passenger traffic, which not only increases the difficulty and complexity of the forecasting work, but also cannot guarantee the forecasting accuracy of the selected model; and the other is that the redifference metric calculated according to (6) is an estimated value, which is not the optimal parameter value to achieve combined

forecasting. Therefore, this paper proposes an improved particle swarm algorithm (IPSO) to solve the optimal redifference metric.

3.2. IPSO Optimization Redifference Metric

3.2.1. Basic Principle of the Particle Swarm Algorithm (PSO). The PSO is a swarm intelligent search algorithm developed by imitating the characteristics of a flock of birds foraging. Its basic idea is to find the final foraging location, i.e., the optimal solution position, by sharing the respective search information of each forager in the flock. In practical application, the particle flies continuously in the set space, and continuously adjusts its position according to its own search experience in the process of finding the optimal target location, until it satisfies the search termination condition to find the optimal solution [20].

Suppose the position and velocity of the ith particle of a certain population at the tth iteration are $x_{i,t}$ and $v_{i,t}$, respectively, then the particle updates the position and velocity by supervising the individual extremes and population extremes to further approximate the optimal solution. The particle velocity and position update are calculated as follows:

$$v_{i,t+1} = w * v_{i,t} + c_1 * \text{rand} * \left(p\text{best}_i - x_{i,t}\right)$$
$$+ c_2 * \text{rand} * \left(g\text{best}_i - x_{i,t}\right), \tag{8}$$
$$x_{i,t+1} = x_{i,t} + \lambda * v_{i,t+1},$$

where w is the weight, taken from 0.4 to 0.9; c_1 and c_2 are the individual learning factor and group learning factor, respectively; rand is a random number generated between 0 and 1; pbest is the individual extreme value; gbest is the group extreme value; and λ is the speed coefficient, generally taken as 1 [21].

3.2.2. Improvement of PSO. The basic PSO generally adopts the method of fixed weights to seek the optimal solution, i.e., a certain fixed value in 0.4~0.9, and the learning factor is generally taken as $c_1 = c_2 = 2$. Depending on different data and simulation environments, its value affects the optimization-seeking ability and convergence speed of the particle swarm algorithm. Therefore, this paper proposes the idea of nonlinear variation to optimize the weights and learning factors of PSO to improve the optimization ability of the algorithm.

The principle of weight optimization is as follows: at the initial iteration, a larger weight is set to ensure that the initial iteration has a larger particle search speed and a better global search capability, and when the number of iterations increases, the weight assignment needs to be reduced to slow down the particle search speed and ensure that the particles have a better local search capability. The optimization formula of weight w is as follows:

$$w = w_{\max} - (w_{\max} - w_{\min}) * \sin\left(\frac{\text{iter}}{\text{ger}} * \frac{\pi}{2}\right), \tag{9}$$

where w_{max} is the maximum weight, w_{min} is the minimum weight, ger is the maximum number of iterations, and iter is the current number of iterations, $0 < \text{iter} \leq \text{ger}, \text{iter} \in N^*$.

The learning factor optimization principle is the optimal value of c_1 ranging from 2.5 to 0.5, and the optimal value of c_2 ranging from 0.5 to 2.5. Therefore, in the optimal value range, in order to ensure the global search ability at the beginning of the iteration, the value of c_1 decreases nonlinearly with the increase of the number of iterations, so that the learning ability of the individual particle itself is larger at the beginning. At the same time, let c_2 increase gradually with the number of iterations to strengthen the group learning ability in the late iteration and avoid the algorithm falling into the local optimum in the late iteration. Based on this idea, this paper selects the nonlinear change function constructed by the power function to improve the learning factor, whose expression is

$$c_1 = 2 * \left(1 - \left(\frac{\text{iter}}{\text{ger}} \right)^3 \right) + 0.5,$$

$$c_2 = 2 * \left(\frac{\text{iter}}{\text{ger}} \right)^3 + 0.5. \tag{10}$$

3.2.3. Algorithm Flow for IPSO Optimization of the Redifference Metric. The specific algorithm for IPSO optimization of the redifference metric taking values is shown in Figure 1.

Step 1. Forecast the railroad passenger volume according to the selected forecasting method.

Step 2. Use the redifference acceleration rule to process the predicted value of the railroad passenger traffic, and at this time, the predicted improved value is a function expression containing.

Step 3. Initialize the population particle parameters. According to the constraints, such as the range of particle parameters, set the initial parameter values.

Step 4. Calculate the particle fitness. With the goal of optimal prediction accuracy, the fitness function $f(\delta)$, which is the minimum absolute value of the relative error of prediction for all prediction years, is determined for the redifference metric optimization. Therefore, the fitness function is defined as follows:

$$f(\delta) = \frac{1}{m} \sum_{t=1}^{m} \frac{|X_t - \widehat{F}_t(\delta)|}{X_t}, \tag{11}$$

where $\widehat{F}_t(\delta)$ is the improved value of the redifferential acceleration rule forecast in a period t, m is the total forecast year, and X_t is the actual value of the passenger traffic in a period t.

Step 5. Compare the particle fitness values of different iterations to find the current optimal particle fitness value and its position, so as to update the particle search position and search speed.
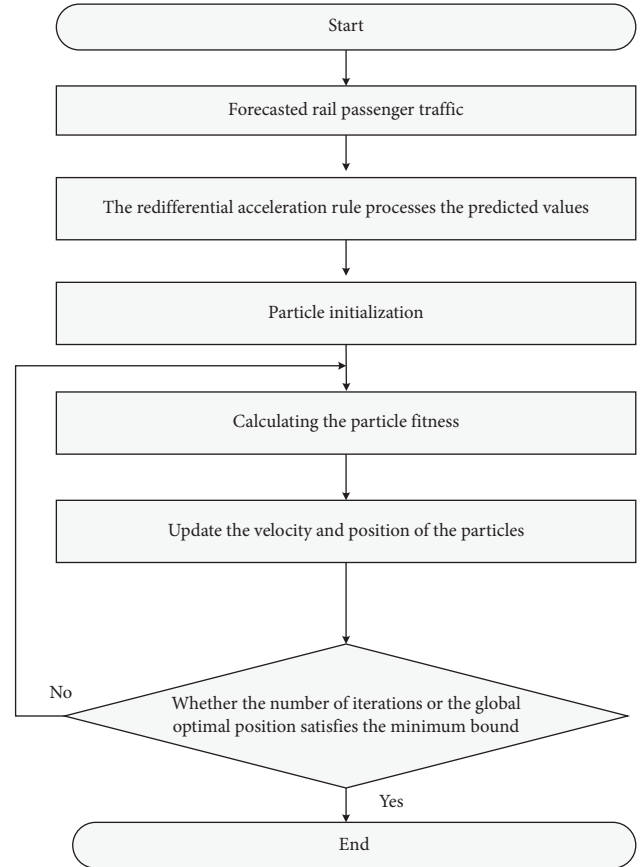

Figure 1: Particle swarm algorithm flow.

Step 6. Determine whether the iteration termination condition of the particle swarm algorithm is satisfied. If satisfied, the algorithm ends; if not, return to step 4.

Step 7. Output the optimal solution.

4. Forecasting Method Selection

Railroad passenger volume forecasting methods can be divided into linear forecasting and nonlinear forecasting models [22]. Among the linear forecasting models, time series forecasting models are represented by ARIMA and the exponential smoothing method, which have the advantage of not needing to study the influence of independent variables (influencing factors) on forecasting results. Among the nonlinear forecasting models, the BP neural network has a better adaptability to nonlinear demand forecasting problems by virtue of its powerful self-adaptability, self-learning, and fault tolerance. Therefore, in this paper, the classical Holt exponential smoothing method and the BP neural network are chosen as the representatives of linear and nonlinear forecasting methods for railroad passenger traffic forecasting.

4.1. Holt Exponential Smoothing Method. The Holt exponential smoothing method is based on the primary exponential smoothing method to forecast the original time series

based on the smoothed and trend values, which strengthens the prediction ability of the exponential smoothing method for trend data [23], and its prediction formula is

$$
\begin{aligned}
F_{t+m} &= S_t + b_t m, \\
S_t &= \alpha X_t + (1 - \alpha)(S_{t-1} + b_{t-1}), \\
b_t &= \gamma(S_t - S_{t-1}) + (1 - \gamma)b_{t-1},
\end{aligned}
\tag{12}
$$

where F_{t+m} is the predicted value in the period $t + m$, S_t is the smoothed value in a period t, X_t is the actual value in a period t, b_t is the trend value in a period t, m is the number of prediction overruns, and α, γ is the exponential smoothing-related parameter [24].

4.2. BP Neural Network. The BP neural network is a multilayer feedforward neural network trained according to the error backpropagation algorithm, which is widely used [25]. The topology of the BP neural network is shown in Figure 2, and its basic idea is the gradient descent method, and the learning process includes forward propagation and backpropagation. In the forward propagation process, the railroad passenger prediction factors are passed from the input layer to the implicit layer as input information, and finally passed to the output layer to output the corresponding railroad passenger volume and training error information [26]. If the output prediction error is larger than the training target, the error is back-propagated from the output layer by repeatedly training and adjusting the weights of w_{ij} and w_{jk} until the error is reduced to an acceptable level or until the maximum number of learning is performed. At this point, the sample data is input again to obtain the output value with the minimum error [27].

There are many types of activation functions for the BP neural network prediction, each with its own advantages and disadvantages. In this paper, we choose the commonly used tansig function and purelin function as the hidden layer neuron and output layer neuron activation functions, respectively, whose expressions are [28] as follows:

$$
\begin{aligned}
\tan \operatorname{sig}(x) &= \frac{2}{1 + e^{-2x}} - 1, \\
\operatorname{purelin}(x) &= x.
\end{aligned}
\tag{13}
$$

5. Railway Passenger Volume Forecast Based on the Spline Interpolation and IPSO-Gradient Difference Acceleration Rule

The combined railroad passenger volume forecasting process combining the spline interpolation method and the redifferential acceleration rule with the Holt exponential smoothing method and BP neural network forecasting is shown in Figure 3.

Step 1: the spline interpolation method is used to make data replacement corrections for the abnormal railroad
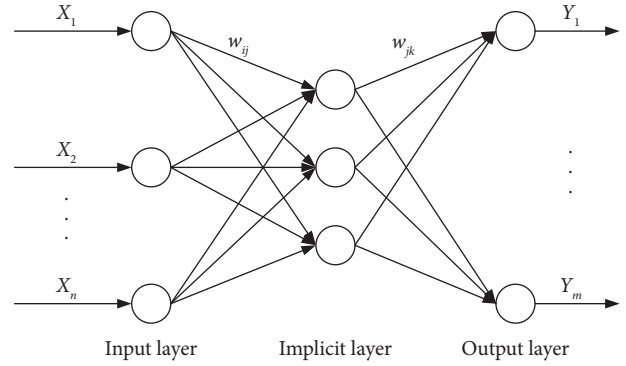

Figure 2: BP neural network's topology.

passenger volume caused by holidays or major events, and the corrected railroad passenger volume data series is formed.

Step 2: based on the corrected railway passenger volume data series, the BP neural network and the Holt exponential smoothing method are applied to make predictions.

Step 3: the adaptation function of the IPSO-optimized redistribution metric was constructed based on the prediction results, and the IPSO was applied to solve the optimal redistribution metric.

Step 4: based on the optimal redifference metric, the redifference acceleration law of equation (7) was applied to improve the prediction results of the BP neural network and Holt exponential smoothing.

6. Case Study

In 2003, Beijing was affected by the SARS epidemic, and the railroad passenger traffic was "abnormally decreasing" [29–33]. Therefore, the research uses Beijing railroad passenger traffic from 2000 to 2019 as the research object, considering the impact of regional GDP, regional resident population, per capita consumption level, and the number of tourists. And the research uses the combined prediction effect of spline interpolation and redifference acceleration method combined with Holt's exponential smoothing method and BP neural network prediction (Table 1).

6.1. Data Preprocessing Based on Spline Interpolation. In the data series of railroad passenger volume from 2000 to 2019 in Beijing, the railroad passenger volume in 2003 is excluded, and three times spline interpolation is performed by calling the spline function with the help of MATLAB, and its function image is shown in Figure 4, and the interpolation calculates the revised value of the railroad passenger volume in 2003 as 52.27 million. The pre-processing situation of the spline interpolation method for railroad passenger traffic in Beijing is shown in Figure 5. It
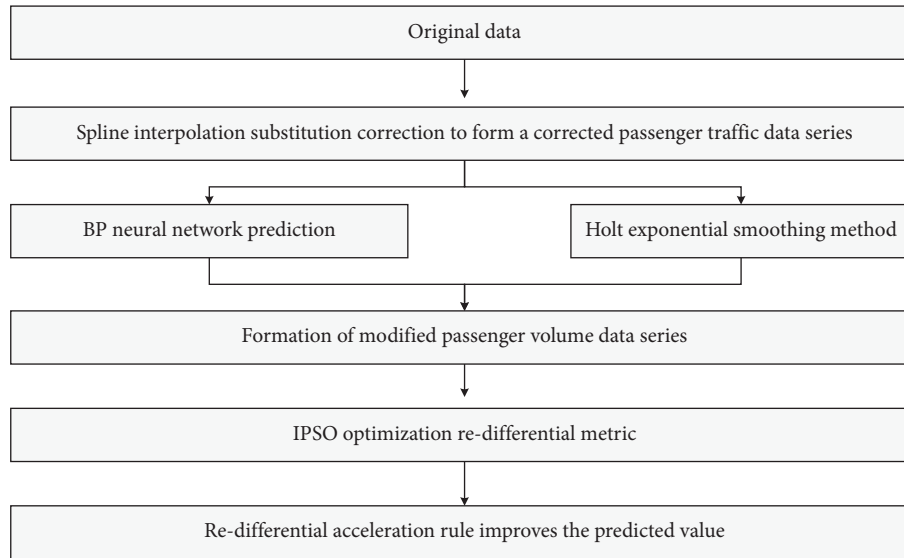
Figure 3: Combined prediction process.

Table 1: Basic data related to railroad passenger volume in Beijing.

Year	Passenger volume (million people)	GDP (billions of RMB)	Resident population (million people)	Consumption level per capita (RMB)	Number of tourists (million people)
2000	4458	3278	1364	9835	10468
2001	4750	3862	1385	10313	11293
2002	5032	4526	1423	12380	11810
2003	4352	5267	1456	13734	8885
2004	5437	6253	1493	15328	12266
2005	5779	7150	1538	16636	12863
2006	6269	8387	1601	19003	13590
2007	6915	10426	1676	21769	14716
2008	7644	11813	1771	24005	14560
2009	8161	12901	1860	26353	16670
2010	8903	14964	1962	29483	18390
2011	9755	17189	2019	32924	21404
2012	10315	19025	2069	36028	23135
2013	11588	21135	2115	40270	25189
2014	12609	22926	2152	43008	26150
2015	12821	24779	2171	46604	27279
2016	13380	27041	2173	50301	28532
2017	13873	29883	2171	55121	29746
2018	14273	33106	2154	61429	31094
2019	14755	35371	2154	67756	32210

can be seen that the overall smoothness of the railroad passenger volume after the spline interpolation method is better, which is more consistent with the overall development trend of the passenger volume.

6.2. Single Model Prediction Based on Spline Interpolation. The railroad passenger volume data from 2000 to 2016 are used as training samples, the railroad passenger volume from 2017 to 2019 are used as test samples, and the Holt exponential smoothing method and BP neural network are used for forecasting, respectively. In the research process, the "relative prediction error" is used as the evaluation index of

the model prediction accuracy, which is abbreviated as "prediction error" for the convenience of presentation.

6.2.1. Holt Exponential Smoothing Prediction. The results and prediction errors of Holt exponential smoothing using SPSS are shown in Figures 6 and 7.

It can be seen that Holt exponential smoothing has a good prediction effect. After the spline interpolation method preprocessed the data, the absolute value of the average fitting prediction error for the training samples was 2.053%, and the absolute value of the average prediction error for the test samples was 2.160%.
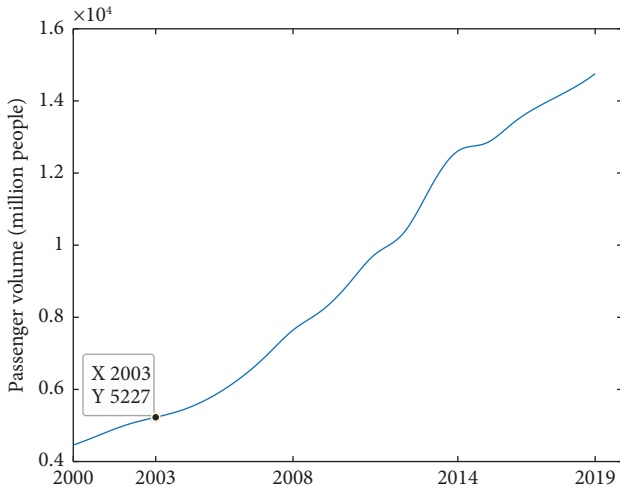
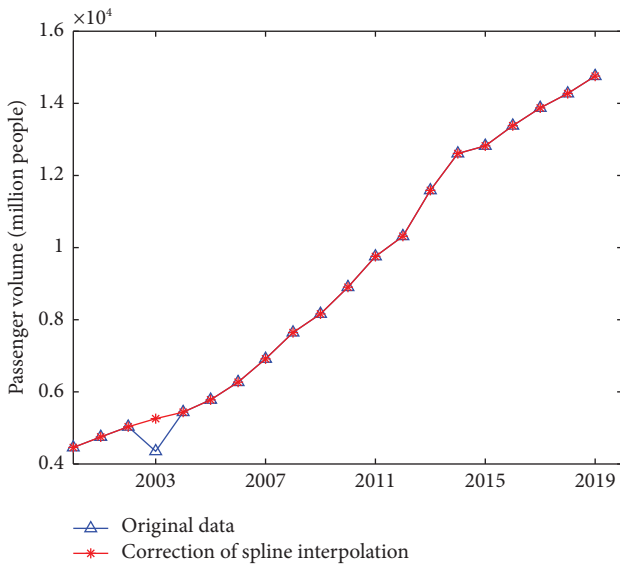FIGURE 4: Spline interpolation function diagram.

FIGURE 6: Holt index smoothing prediction results.

FIGURE 5: Preprocessing of the railroad passenger traffic data.

FIGURE 7: Holt index smoothing prediction error.

The prediction effect of the spline interpolation method is better after the substitution correction of abnormal passenger traffic data. When predicting based on the original data of Beijing railroad passenger volume, the prediction error of the passenger volume in 2003 was as high as 16.016%, and the absolute values of the average prediction errors of its training and testing samples were 5.228% and 5.862%, respectively, which were 2.6 times of the prediction errors after processing the abnormal passenger volume data based on the spline interpolation method.

6.2.2. BP Neural Network Prediction.
The prediction process of the BP neural network is implemented by MATLAB programming, and the main parameters are set as follows: the number of training is 1000 times, the hidden layer is 1 layer, and the neurons are set as 1. The prediction performance of the BP neural network is shown in Figures 8 and 9.
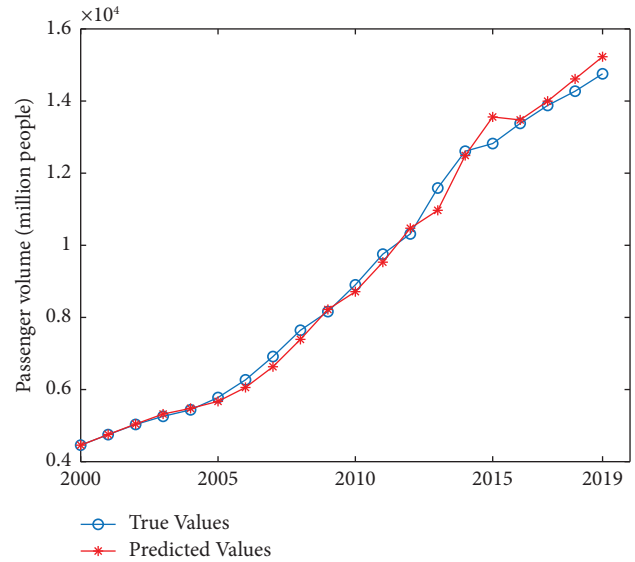
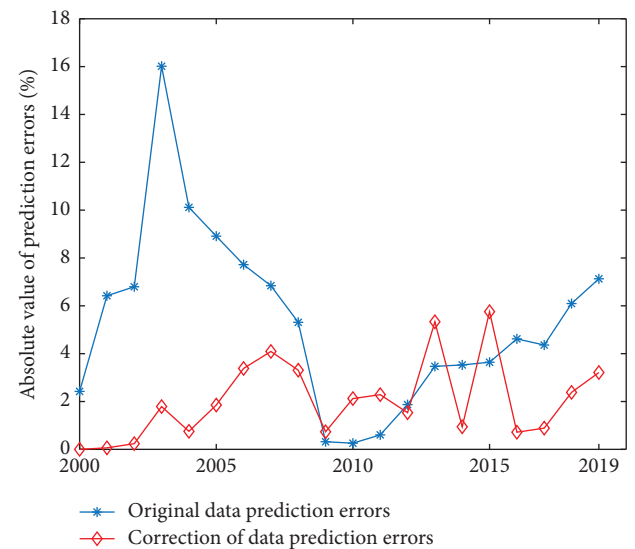Figure 8 shows the variation of prediction error with the number of iterations for different samples, blue is the training set, red is the test set, and green is the validation set generated by the system. The network converges after 14 times of training, and the network error of the validation set is 0.00917, and the network error of the training and test sets is much lower than 0.00917, with a better error accuracy. Figure 9 represents the BP neural network's fit prediction goodness and the fit goodness of the training set, validation set, and the test set and all data are 0.99929, 0.97893, 0.99971, and 0.99777, respectively, which shows that the network training effect is very good.

The absolute values of the predicted values and prediction errors of the BP neural network are shown in Figures 10 and 11. Figure 10 demonstrates that the predicted and true values of the BP neural network are better fitted. It
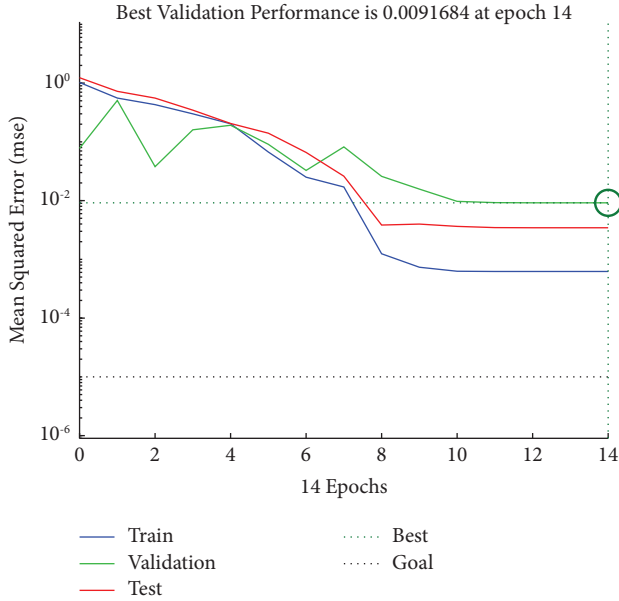
Figure 8: Convergence process of the BP neural network.

can also be seen from Figure 11 that the spline interpolation method has a better prediction effect after processing abnormal passenger traffic data, and the average absolute value of prediction error for training and test samples is reduced by 1.712% and 0.860%, respectively.

6.3. Combined Prediction Effect Based on the Redifferential Acceleration Law

6.3.1. Simulation Parameter Setting. The algorithm experiments in this section are conducted in the MATLAB 2019 environment, and the improved particle swarm algorithm code is written to find out the optimal redifference metric. The main parameters of the particle swarm algorithm are initialized as follows: the population size is 30, the maximum number of iterations is 300, and the speed search interval is -0.5 to 0.5. To further verify the effect of the nonlinear change optimization method proposed in this paper on the superiority-seeking ability and convergence speed of the particle swarm algorithm, the superiority-seeking situation and convergence speed of the basic particle swarm algorithm are compared with those of the basic particle swarm algorithm. The fixed weight of the basic particle swarm algorithm is taken as 0.9, and the learning factor is $c_1 = c_2 = 2$.

6.3.2. Prediction Results and Analysis

(1) Analysis of IPSO Algorithm Advantage. The variation of the fitness values of different particle swarm algorithms for solving the redifference metric with the number of iterations is shown in Figure 12. From the experimental results, it can be seen that the improved particle swarm algorithm with nonlinearly varying weights and learning factors proposed in this paper has a better finding ability and convergence speed, which is reflected in the following: the fitness value of the

improved particle swarm algorithm for solving the optimal heavy difference metric is 0.00640188456, which converges after the 113th iteration, while the fitness value of the basic particle swarm algorithm for solving the optimal heavy difference metric is 0.00640189062, which converges after the 234th iteration. In contrast, the improved particle swarm algorithm has a better seeking ability and convergence speed.

(2) Combined Prediction Effect of the IPSO-Heavy Difference Acceleration Law. The IPSO-optimized heavy difference metric is taken as shown in Figure 13, and the optimal heavy difference metric is solved as 1.9088.

Section 4.2 shows that the prediction accuracy of the Holt exponential smoothing is better than that of the BP neural network. The prediction results of the BP neural network and Holt exponential smoothing from 2017 to 2019 after processing abnormal passenger traffic data by the spline interpolation method are improved by the redifference acceleration rule equation (7). To verify the improvement effect, the prediction results are compared with those of the average weighted combination approach, i.e.,

$$F_w = w_1 F_{bp} + w_2 F_{\text{holt}}, \tag{14}$$

where F_w is the average weighted combination of the passenger traffic forecast, F_{bp} is the BP neural network forecast, F_{holt} is the Holt exponential smoothing forecast, and w_1 and w_2 are the weights, $w_1 = w_2 = 0.5$.

The predicted values of the railroad passenger traffic in Beijing from 2017 to 2019 are shown in Table 2.

The analysis shows that the improved prediction values of the redifference acceleration rule have a better prediction accuracy, which is reflected in the following.

After the improvement of the prediction results of the heavy difference acceleration rule on the BP neural network and the Holt exponential smoothing method, the mean value of the absolute value of the prediction error from 2017 to 2019 is 0.642%, which is 3.320% and 1.518% less than that of the BP neural network and the Holt exponential smoothing method, respectively.

Comparing the forecast results of the redifferential acceleration rule with the average weighted combination forecast results, the absolute value of its average forecast error of the railroad passenger traffic from 2017 to 2019 is reduced by 2.419%. In addition, the mean value of the absolute forecast error of the average weighted combination is 3.061%, which is higher than the mean absolute forecast error of the Holt exponential smoothing method of 2.160%. This also indicates that the direct weighted combination of multiple model prediction results has some uncertainty, i.e., there is no guarantee that its combined prediction effect is necessarily better than that of a single prediction model. The IPSO-redifference acceleration method effectively avoids this problem.

To further verify the advantage of the IPSO-redifference acceleration rule over the traditional redifference acceleration rule, the estimated value of the redifference metric is calculated according to (6). It is known that this method needs to introduce another prediction method with a worse
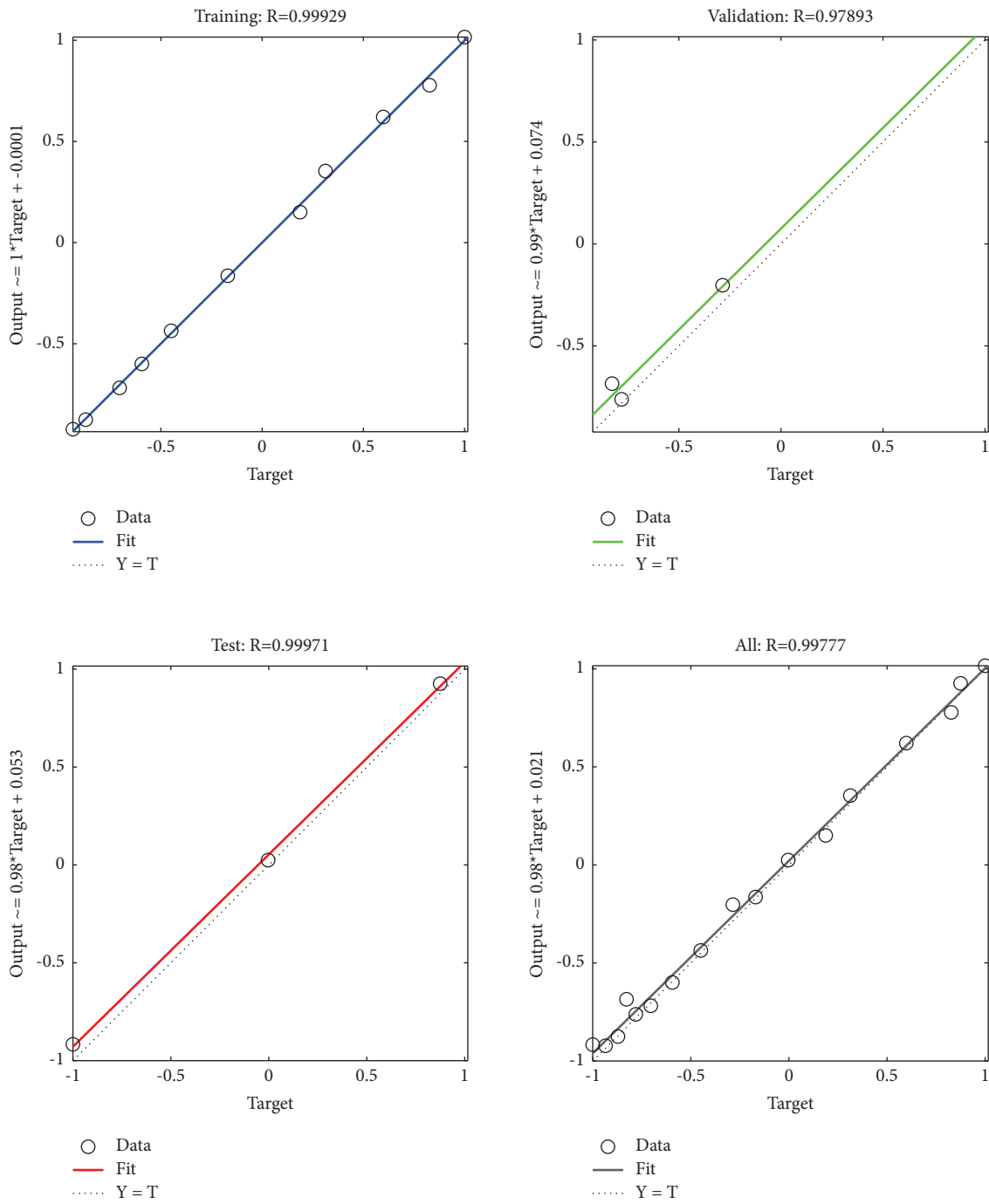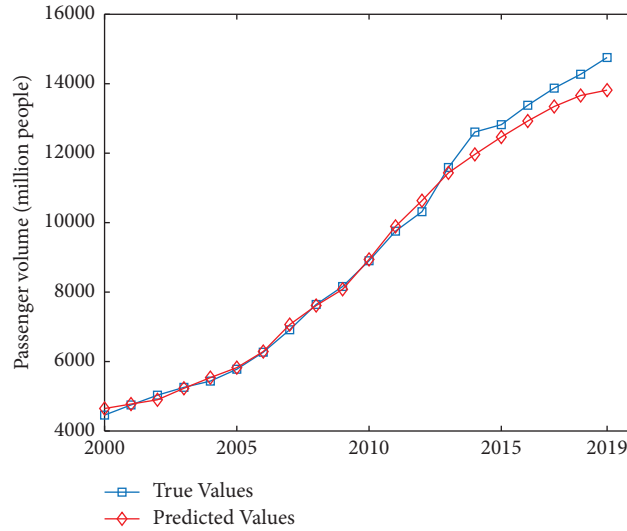
Figure 9: Goodness of fit of the BP neural network.

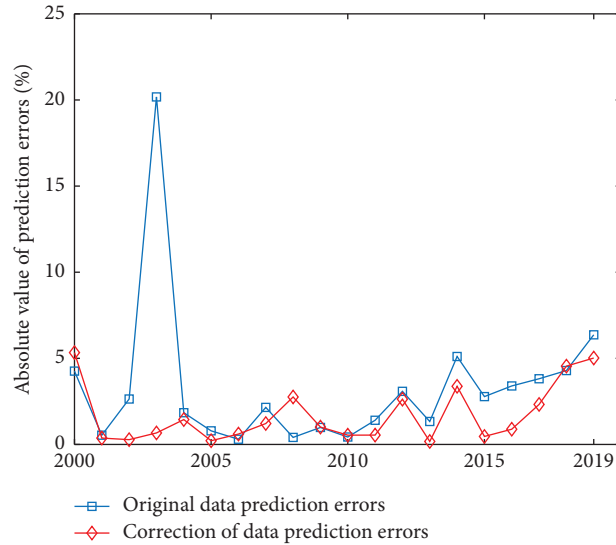Figure 10: BP neural network's prediction results.

Figure 11: BP neural network's prediction error of absolute value.

prediction accuracy compared to the BP neural network and the Holt exponential smoothing method. After several experiments, the ARIMA (3,0,1) prediction model was selected, and its prediction results in the SPSS25 environment are shown in Figure 14. The traditional redifference acceleration method improves the prediction value with the following idea: the mean value of the prediction value of the BP neural network, the Holt exponential smoothing method, and the ARIMA (3,0,1) prediction model from 2017 to 2019 is substituted into (6) to estimate the redifference metric, and then the prediction results of the BP neural network and the Holt exponential smoothing method are improved according to (7).

The predicted values of the ARIMA (3,0,1) model from 2017 to 2019 are 143.64 million, 156.93 million, and 168.56 million. The mean values of the BP neural network, the Holt exponential smoothing method, and ARIMA (3,0,1)

predicted values from 2017 to 2019 are 148.71 million, 146.13 million, and 156.38 million, respectively. Therefore, the redifference metric is calculated as

$$\delta = \frac{14871 - 15638}{14613 - 14871} = 2.973. \quad (15)$$

According to (7), the improvement values of the traditional redifferential acceleration rule on the prediction results of the BP neural network and the Holt exponential smoothing method from 2017 to 2019 were calculated as 138.95 million, 144.56 million, and 150.94 million, and the absolute values of prediction errors were 0.157%, 1.282%, and 2.296%.

By comparing the calculated results with Table 1, it can be seen that, after improving the forecasting results of the BP neural network and the Holt exponential smoothing method, the mean value of the absolute forecast error from
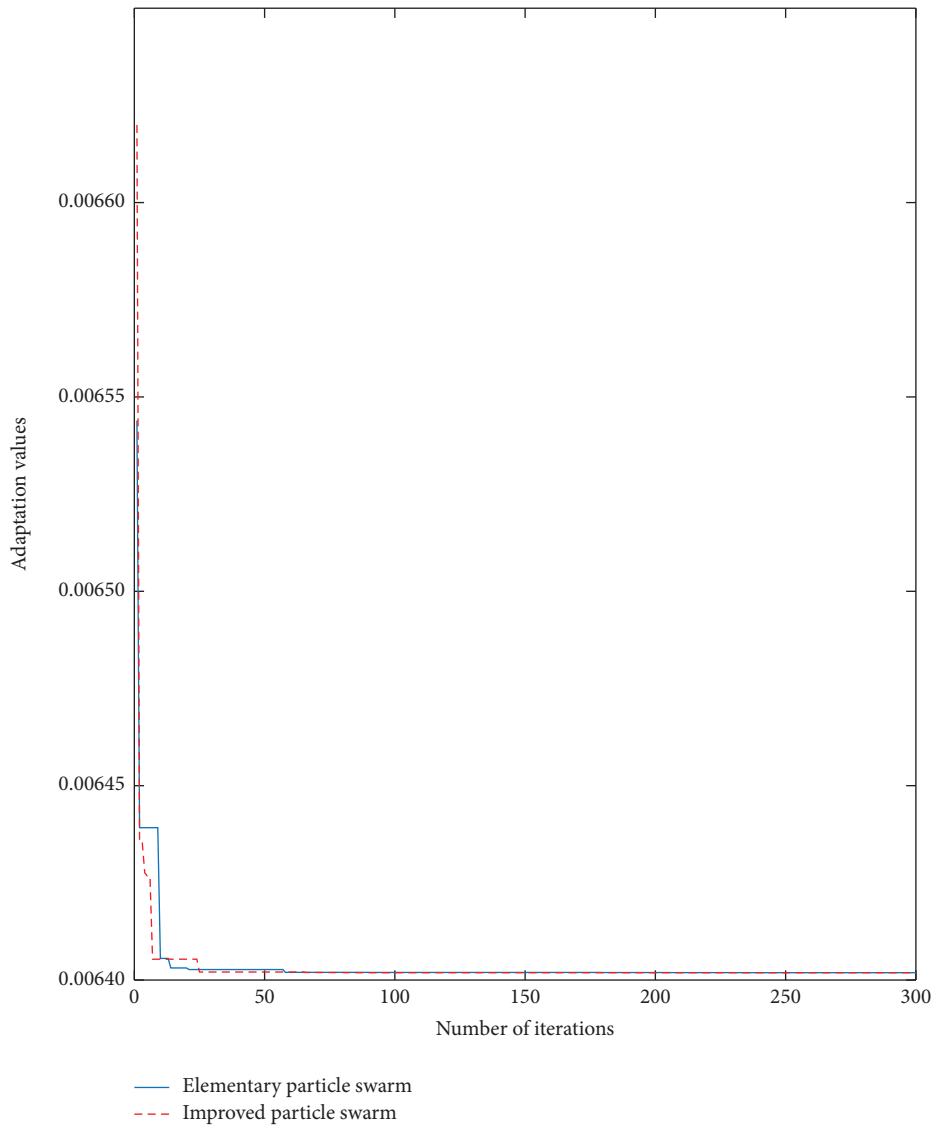
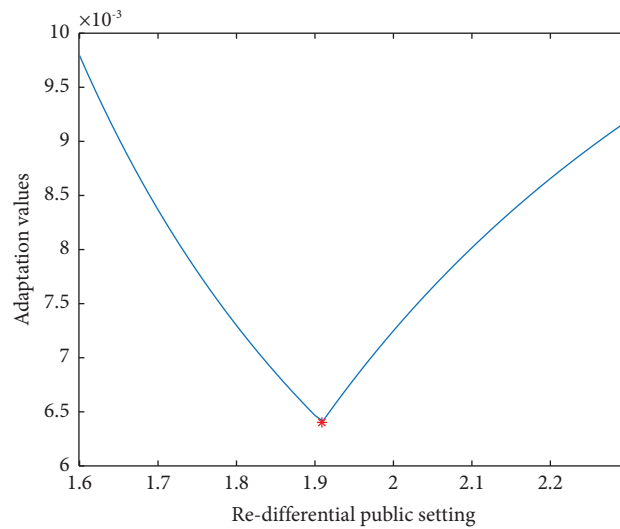Figure 12: Optimization effect of different particle swarm algorithms.

Figure 13: Redifferential metric values.

Table 2: Forecast values of the railroad passenger traffic in Beijing for different forecasting methods.

| Forecast year	BP neural network prediction method		Holt exponential smoothing method		Average weighted portfolio method		IPSO-redifferential acceleration law method	
	Forecast value (million people)	Absolute value of prediction errors (%)	Forecast value (million people)	Absolute value of prediction errors (%)	Forecast value (million people)	Absolute value of prediction errors (%)	Forecast value (million people)	Absolute value of prediction errors (%)
2017	14195	2.323	13996	0.887	14096	1.605	13777	0.694
2018	14922	4.548	14613	2.382	14768	3.465	14272	0.007
2019	15495	5.017	15229	3.212	15362	4.115	14936	1.227
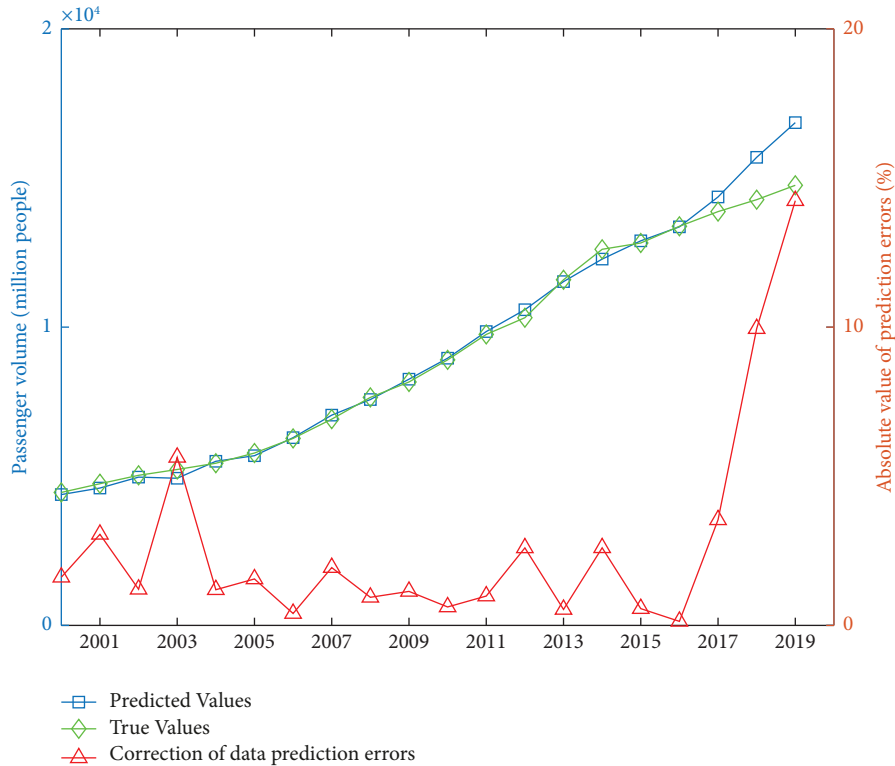
FIGURE 14: ARIMA (3,0,1)'s forecast results.

2017 to 2019 is 0.642%, which is 1.518% less than the forecast error of the Holt exponential smoothing method's "main" forecasting model.

The IPSO-redifferential acceleration method has a higher prediction value improvement accuracy than that of the traditional redifferential acceleration method, and its average prediction error is reduced by 0.602% in absolute value.

7. Conclusion

To address the problem of abnormal railroad passenger volume data due to holidays or major events and the uncertainty of weighted combination forecasting, this paper introduces the spline interpolation method to make replacement corrections for abnormal railroad passenger volume and reduce the interference of abnormal data on forecasting accuracy. In addition, an improved particle swarm algorithm is proposed to optimize the redifference acceleration rule to improve the railroad passenger volume forecasting results, and combine the BP neural network and the Holt exponential smoothing method to forecast the railroad passenger volume in Beijing. The results show that the spline interpolation method has a better prediction effect after the substitution correction of anomalous data, and the improved particle swarm algorithm also shows a better finding ability and convergence speed when solving the optimal redifference metric. Compared with the

prediction results of the BP neural network, the Holt exponential smoothing method, the average weighted combination method, and the traditional redifference acceleration rule, the prediction accuracy of the IPSO-redifference acceleration rule to improve the prediction results is higher. It is worth mentioning that there are more interpolation methods in the field of numerical analysis, such as Newton interpolation and Lagrange interpolation, and which interpolation method has the best effect in solving the problem of data anomalies caused by holidays or major events is the focus of future research.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. A. H. Dempster and A. M. Ireland, "Object-oriented model integration in a financial decision support system," *Decision Support Systems*, vol. 7, no. 4, pp. 329–340, 1991.

[2] T. Bui and J. Lee, "An agent-based framework for building decision support systems," *Decision Support Systems*, vol. 25, no. 3, pp. 225–237, 1999.

[3] K. Aihara, T. Takabe, and M. Toyoda, "Chaotic neural networks," *Physics Letters A*, vol. 144, no. 6-7, pp. 333–340, 1990.

[4] W. Li, F. Feng, and Q. Jiang, "Prediction for railway passenger volume based on modified PSO optimized LSTM neural network," *Journal of Railway Science and Engineering*, vol. 15, no. 12, pp. 260–266, 2018.

[5] Y. Y. Qiu, Q. Zhang, and M. Lei, "Forecasting the railway freight volume in China based on combined PSO-LSTM model," in *Proceedings of the 2020 second International Conference on Artificial Intelligence Technologies and Application (ICAITA)*, Dalian, China, August 2020.

[6] H. Wang, Y. Wang, and D. Wu, "A new seasonal cycle GM (1, 1) model and its application in railway passenger volume forecasting," *Grey Systems: Theory and Application*, vol. 12, no. 2, pp. 293–317, 2022.

[7] J. Ye, Z. Xu, and X. Gou, "An adaptive Grey-Markov model based on parameters Self-optimization with application to passenger flow volume prediction," *Expert Systems with Applications*, vol. 202, Article ID 117302, 2022.

[8] M. Ge, Z. Junfeng, W. Jinfei, H. Huiting, S. Xinghua, and W. Hongye, "ARIMA-FSVR hybrid method for high-speed railway passenger traffic forecasting," *Mathematical Problems in Engineering*, vol. 2021, Article ID 9961324, 5 pages, 2021.

[9] F. Jiao, L. Huang, R. Song, and H. Huang, "An improved STL-LSTM model for daily bus passenger flow prediction during the COVID-19 pandemic," *Sensors*, vol. 21, no. 17, p. 5950, 2021.

[10] Y. Wang, R. Jia, F. Dai, and Y. Ye, "Traffic flow prediction method based on seasonal characteristics and SARIMA-NAR model," *Applied Sciences*, vol. 12, no. 4, p. 2190, 2022.

[11] L. Qin, W. Li, and S. Li, "Effective passenger flow forecasting using STL and ESN based on two improvement strategies," *Neurocomputing*, vol. 356, pp. 244–256, 2019.

[12] W. Guo, Y. Feng, W. Luo et al., "The impacts of COVID-19 and policies on spatial and temporal distribution characteristics of traffic: two examples in Beijing," *Sustainability*, vol. 14, no. 3, p. 1733, 2022.

[13] R. Adhikari and R. K. Agrawal, "An introductory study on time series modeling and forecasting," 2013, https://arxiv.org/abs/1302.6613.

[14] Z. Changqing and L. Chengzhi, "Research on railway passenger traffic volume forecasting method," in *Proceedings of the 2009 International Forum on Information Technology and Applications*, IEEE, Chengdu, China, May 2009.

[15] C. De Boor and C. De Boor, *A practical guide to splines*, Springer-Verlag, New York, NY, USA, 1978.

[16] G. Birkhoff and C. De Boor, "Error bounds for spline interpolation," *Indiana University Mathematics Journal*, vol. 13, no. 5, pp. 827–835, 1964.

[17] A. L. S. Maia and F. D. A. de Carvalho, "Holt's exponential smoothing and neural network models for forecasting interval-valued time series," *International Journal of Forecasting*, vol. 27, no. 3, pp. 740–759, 2011.

[18] H. Yue, W. Chen, W. Chen, and X. Wu, "Spline-interpolation based PVT algorithm and application in a bionic cockroach robot," in *Proceedings of the 2010 11th international conference on control automation robotics & vision*, pp. 1742–1747, IEEE, Singapore, December 2010.

[19] M. G. Cox, "An algorithm for spline interpolation," *IMA Journal of Applied Mathematics*, vol. 15, no. 1, pp. 95–108, 1975.

[20] F. Marini and B. Walczak, "Particle swarm optimization (PSO). A tutorial," *Chemometrics and Intelligent Laboratory Systems*, vol. 149, pp. 153–165, 2015.

[21] M. Jain, V. Saihjpal, N. Singh, and S. B. Singh, "An overview of variants and advancements of PSO algorithm," *Applied Sciences*, vol. 12, no. 17, p. 8392, 2022.

[22] X. Yang, Q. Xue, M. Ding, J. Wu, and Z. Gao, "Short-term prediction of passenger volume for urban rail systems: a deep learning approach based on smart-card data," *International Journal of Production Economics*, vol. 231, Article ID 107920, 2021.

[23] E. S. Gardner Jr, "Exponential smoothing: the state of the art," *Journal of Forecasting*, vol. 4, no. 1, pp. 1–28, 1985.

[24] P. Goodwin, "The holt-winters approach to exponential smoothing: 50 years old and going strong," *Foresight*, vol. 19, no. 19, pp. 30–33, 2010.

[25] J. Li, J. H. Cheng, J. Y. Shi, and F. Huang, "Brief introduction of back propagation (BP) neural network algorithm and its improvement," in *Proceedings of the Advances in Computer Science and Information Engineering*, Zhengzhou, China, May 2012.

[26] M. Buscema, "Back propagation neural networks," *Substance Use & Misuse*, vol. 33, no. 2, pp. 233–270, 1998.

[27] R. Xie, X. Wang, Y. Li, and K. Zhao, "Research and application on improved BP neural network algorithm," in *Proceedings of the 2010 5th IEEE Conference on Industrial Electronics and Applications*, pp. 1462–1466, IEEE, Taichung, Taiwan, June 2010.

[28] Z. Laabid, A. Moumen, K. Mansouri, and A. Siadat, "Numerical study of the speed's response of the various intelligent models using the tansig, logsig and purelin activation functions in different layers of artificial neural network," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 1, p. 155, 2023.

[29] P. Beutels, N. Jia, Q. Y. Zhou, R. Smith, W. C. Cao, and S. J. De Vlas, "The economic impact of SARS in Beijing, China," *Tropical Medicine and International Health*, vol. 14, pp. 85–91, 2009.

[30] K. Leung, E. H. Lau, C. K. Wong, G. M. Leung, and J. T. Wu, "Estimating the transmission dynamics of SARS-CoV-2 Omicron BF. 7 in Beijing after adjustment of the zero-COVID policy in November–December 2022," *Nature Medicine*, vol. 29, no. 3, pp. 579–582, 2023.

[31] R. Liu, D. Li, and S. Kaewunruen, "Role of railway transportation in the spread of the coronavirus: evidence from wuhan-beijing railway corridor," *Frontiers in Built Environment*, vol. 6, Article ID 590146, 2020.

[32] O. Dombey, "The effects of SARS on the Chinese tourism industry," *Journal of Vacation Marketing*, vol. 10, no. 1, pp. 4–10, 2004.

[33] H. Zhou, Y. Wang, J. R. Huscroft, and K. Bai, "Impacts of COVID-19 and anti-pandemic policies on urban transport—an empirical study in China," *Transport Policy*, vol. 110, pp. 135–149, 2021.