Short-term Forecasting of High-Speed Rail Passenger Flow

Pei Zhang[1, a], Xiao-long Li [2,b] and Qin-zhao Wang [3,c]

[1]School of Academy of Armored Force Engineering, Beijing 100072, China

[2]School of Academy of Armored Force Engineering, Beijing 100072, China

[3]School of Academy of Armored Force Engineering, Beijing 100072, China

[a]13120819@bjtu.edu.cn, [b]13120817@bjtu.edu.cn, [c]11125065@bjtu.edu.cn

Abstract. Short-term forecasting of high-speed rail (HSR) passenger flow is the key to high-speed passenger rail planning decision-making. Furthermore, accurate short-term demand estimates are sure to success of rail revenue management. Different station passenger flow is effect each other, however, previous researches are lack of consideration. This paper proposed an approach based on the ensemble empirical mode decomposition and grey support vector machine. Then we use the approach to forecast the passenger flow of multiple stations. Application results indicate that the approach is effective in terms of prediction accuracy.

Introduction

Short-term forecasting of high-speed rail (HSR) passenger flow is necessary to planning decision-making of HSR, as it can provide daily ridership estimations in the next week. Furthermore, accurate short-term passenger flow is sure to success of improving the quality of high-speed rail service and rail revenue management. The forecast of high - speed rail short-term passenger flow traffic has become a hotspot.

Many foreign scholars have made a thorough study on the passenger flow of high-speed rail, and a number of prediction methods about high-speed railway passenger flow have been built. The first person, who built the Box-Jenkins time series model to forecast the high-speed rail passenger flow, is Cook [1]. MD model, which was first put forward in Japan and widely used to forecast passenger flow of passenger lines, suitable for Chinese high-speed rail has been proposed by Kiyoham et al. [2] on the base of describing the application of the model in Japan's Shinkansen. As many scholars working on neural networks, the development of modern science neural network is rapid. On the basis of the research, artificial neural networks mathematical model have been proposed for predicting traffic. The combination forecasting model of neural network model and logit model was established by Peter [3]. Anthony Stathopoulos and Matthew G. Karlaftis [4] has carried out the spectrum analysis and cross spectrum analysis to the traffic flow at different places. Yibing Wang and Markos Papageorgious [5] promoted the application of Extended Kalman Filter (EKF) method combined with macroscopic traffic flow model for real-time traffic flow forecasting motorway, and discuss the process of the model and parameter estimation in detail.

Methodology

Based on the analysis of multiple forecasting methods, this paper establishes a passenger flow forecasting model based on empirical modal decomposition and gray support vector machine. Firstly, the empirical modal decomposition method can decompose the original data into several intrinsic modal components, which can eliminate the effects of noise. Then the gray support vector machine model is used to "isolate" the prediction. Finally, the passenger flow forecasting result can be get based on the modal reconstruction of each mode.

Empirical Modal Decomposition Method

Any signal $x(t)$ can be decomposed as follows using the EEMD method.

$$x(t) - m_1 = h_1 . \tag{1}$$

$$r_1 = x(t) - c_1. \tag{2}$$

$$r_n = r_{n-1} - c_n. \tag{3}$$

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t). \tag{4}$$

m_1 is the mean line, c_1 is the component of m_1, r_1 is the components removed high frequency, $r_n(t)$ is the remaining components, representing the average trend of the original signal, $c_i(t)$ represents the components of the signal from high to low different frequency segments.

The process of the empirical modal decomposition method is constantly "filtering", and its goal is to decompose the original time series curve into several different IMF component curves according to the algorithm. Each curve has its own range and is a non-overlapping frequency Interval. The decomposition of the components obtained by decomposition is "simple", which will help to analyze the fluctuation rules of the original sequence and the inherent laws of the original sequence. It can also better apply each IMF component to other algorithms to analyze the original sequence.

Gray Support Vector Machine Model

Gray sequence generation Model

Accumulate generated

$$x^{(0)}(k)d = \sum_{i=1}^{k} x^{(0)}(i); \; k = 1, 2, \mathbf{L}, n. \tag{5}$$

$$X^{(0)}D^r = X(r)(x^{(r)}(1)d, x^{(r)}(2)d, \mathbf{L}, x^{(r)}(n)d). \tag{6}$$

$$x(r)(k) = \sum_{i=1}^{k} x(r-1)(i); \; k = 1, 2, \mathbf{L}, n. \tag{7}$$

$X^{(0)}$ is the original time series, D is a cumulative generation operator of $X^{(0)}$, D^r is the nth cumulative generation operator of $X^{(0)}$.

Subtraction generation

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \mathbf{L}, x^{(0)}(n)). \tag{8}$$

$$X^{(0)}D = (x^{(0)}(1)d, x^{(0)}(2)d, \mathbf{L}, x^{(0)}(n)d). \tag{9}$$

$$x^{(0)}(k)d = x^{(0)}(k) - x^{(0)}(k-1); \; k = 1, 2, \mathbf{L}, n. \tag{10}$$

Support Vector Machine Model

We assume that the total amount of the training sample set is 1 , $\{(x_i, y_i), i = 1, 2, \mathbf{L}\} \subset R^l \times R$.

$$f(x) = w\Phi(x) + b. \tag{11}$$

$$L(f(x), y, e) = \begin{cases} 0, & |y - f(x)| \le e \\ |y - f(x)| - e, & |y - f(x)| > e \end{cases}. \tag{12}$$

$$\begin{cases} \min \dfrac{1}{2}\|w\|^2 + C\sum\limits_{i=1}^{l}(\boldsymbol{x}_i + \boldsymbol{x}_i^*) \\ \text{s.t.} \begin{cases} y_i - w\Phi(x_i) - b \le e + \boldsymbol{x}_i \\ -y_i + w\Phi(x_i) + b \le e + \boldsymbol{x}_i^* \\ \boldsymbol{x}_i \ge 0, \boldsymbol{x}_i^* \ge 0 \end{cases}, i=1,2,\mathbf{L},l. \end{cases} \tag{13}$$

$$\begin{cases} \max\limits_{a,a^*} \left[-\dfrac{1}{2}\sum\limits_{i=1}^{l}\sum\limits_{j=1}^{l}(\boldsymbol{a}_i - \boldsymbol{a}_i^*)(\boldsymbol{a}_j - \boldsymbol{a}_j^*)\Phi - \sum\limits_{i=1}^{l}(\boldsymbol{a}_i + \boldsymbol{a}_i^*)e + \sum\limits_{i=1}^{l}(\boldsymbol{a}_i - \boldsymbol{a}_i^*)y_i \right] \\ \text{s.t.} \begin{cases} \sum\limits_{i=1}^{l}(\boldsymbol{a}_i - \boldsymbol{a}_i^*) = 0 \\ 0 \le \boldsymbol{a}_i, \boldsymbol{a}_i^* \le C \end{cases} \end{cases}. \tag{14}$$

$$f(x) = \sum_{i=1}^{l}(\boldsymbol{a}_i - \boldsymbol{a}_i^*)K(x_i, x) + b^*. \tag{15}$$

x_i is importation, y_i is forecasting target, e is linear insensitive loss function
$f(x)$ is predictive value, y is true value, C is penalty factor.

$\Phi = K(x_i, x_j) = \Phi(x_i)\,\Phi(x_j)$ is kernel function. We choose the radial basis function to solve the model. And we use the particle swarm optimization algorithm to optimize the parameters.

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2s^2}\right). \tag{16}$$

The model structure is shown as Fig. 1.

$$f(x) = \sum_{i=1}^{l} \left(\alpha_i - \alpha_i^* \right) K\left(x_i, x\right) + b^*$$

$$f(x)$$

$\alpha_1^* y_1$ $\alpha_2^* y_2$ $\alpha_2^* y_2$

$K(x_1,x)$ $K(x_2,x)$ Ă $K(x_2,x)$
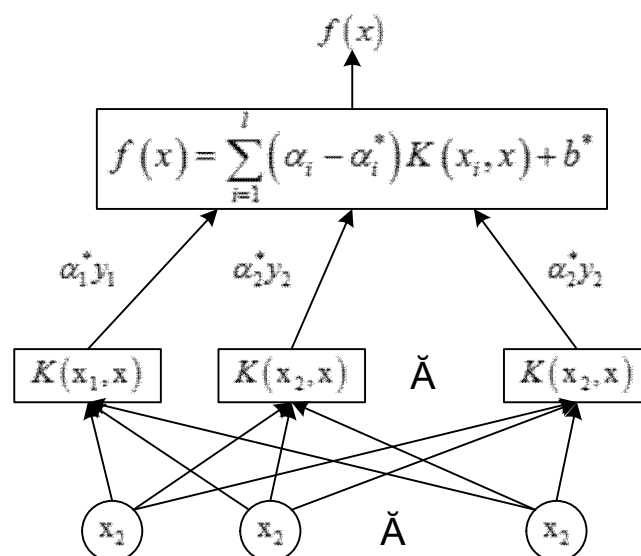
x_1 x_2 Ă x_1

Fig. 1 The support vector machine model structure

Case Study

In order to verify the validity of the model, this paper selects several stations of Wuhan-Guangzhou high-speed railway as the research object. A total of 638-day (between February 1, 2010 and October 31, 2011) long daily passenger flows have been used. The predicted results are as following.
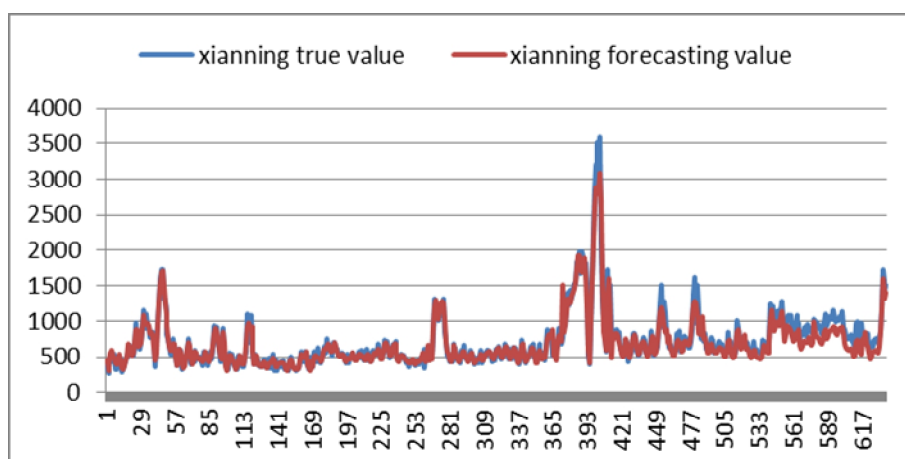

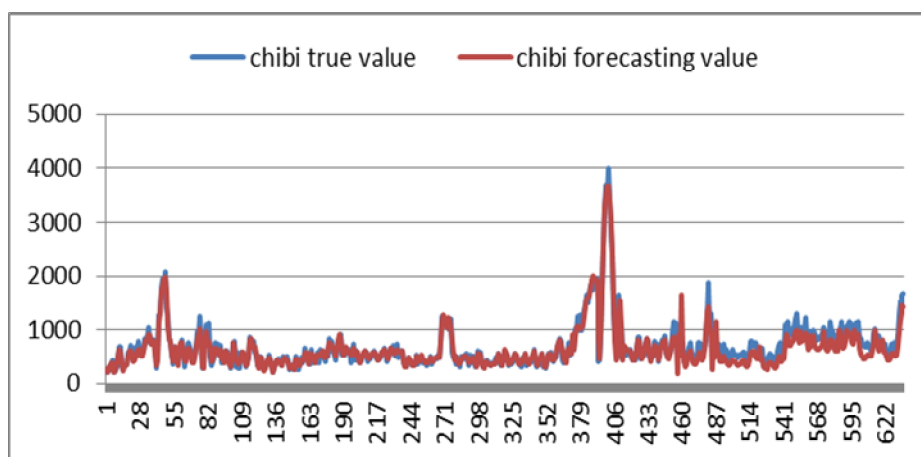Fig. 2 The comparison of predicted results of Xianning station

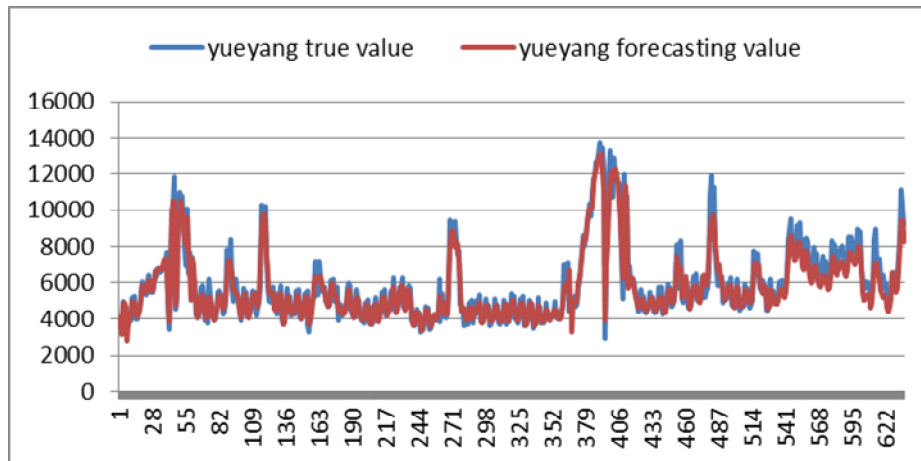Fig. 3 The comparison of predicted results of Chibi station

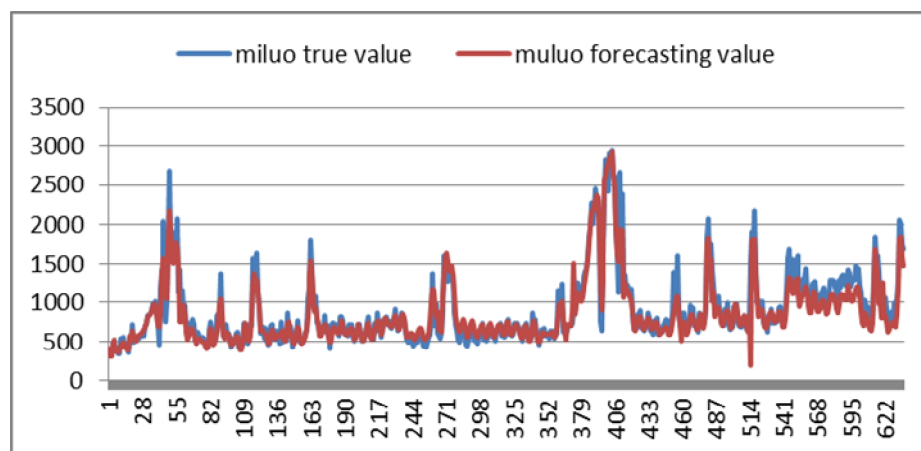Fig. 4 The comparison of predicted results of Yueyang station

Fig. 5 The comparison of Between 1 and 2 results of Miluo station

Table.1 a comparison of indicators of four types of prediction error method

Section	MAPE	MAD	MSE	Squared correlation
Xianning	9.86%	112.43	0.0168	0.8721
Chibi	7.32%	95.76	0.0039	0.9655
Yueyang	8.06%	421.65	0.0072	0.8681
Miluo	8.43%	107.45	0.0065	0.8237

Conclusions

From the case study we can see the errors are small between predicted value and the true value of the high-speed rail passenger flow. The errors are 9.86% ,7.32% ,8.06% and 8.43%, meeting the accuracy requirements of less than 10%. It shows that the short-term forecasting model of high-speed rail (HSR) passenger flow established based on the ensemble empirical mode decomposition and grey support vector machine in this paper is feasible and effective.

References

[1] A. R. Cook. Analysis of Railway Time-series Data Using Box-Jerkins Technique. InTransportation Research 733,National Research Council,Washington,D.C.,1979.

[2] Kiyoharu Takagi,Ryujiro Yamagishi,Li Qun - ren. Demand Forecast for China High – SpeedRailway. Traffic and Transportation Studies; Proceedings of Ictts,pp. 899-909. 2000

[3] Peter Nijkamp, Aura Reggiani, Tommaso Tritapepe.Modeling Railway Passenger Flows in Italya Comparison between Neural Network Analysis and Logit Analysis. Transportation Research Part C:Emerging Technologies, Volume 4,Number 6, December 1996,pp. 323-338(16).

[4] Stathopoulos A, Karlaftis M G. A multivariate state space approach for urban traffic flow modeling and prediction[J]. Transportation Research Part C Emerging Technologies, 2003, 11(2):121-135.

[5] Wang Y, Papageorgiou M. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach[J]. Transportation Research Part B: Methodological, 2005, 39(2): 141-167.