ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202203_5(3).0044

A STL-AESN-SVR Ensemble Model to Forecast Railway Passenger Flow Volume

Hao Tang[1,*]

[1]School of Economics, Jinan University, Guangzhou, Guangdong, 510632, China

Abstract

Railway Traffic is vital to people's daily travelling and life, and the forecast of railway passenger flow is vital to better operate and manage railway system. To improve the precision of railway passenger flow prediction, a new ensemble method STL-AESN-SVR which combines Seasonal-Trend decomposition procedure based on Loess (STL), Support Vector Regression (SVR) and Echo State Networks with Adaboost (AESN) was proposed. This new method first used STL to decompose monthly railway passenger flow into trend component, seasonal component and remainder component and then adapted AESN to predict for trend component, used SVR to predict the remainder component and used the naïve seasonal method to forecast seasonal component. The final prediction is obtained by summing up the prediction of those three components. Using the monthly railway passenger flow data of China, it can be found that this new method can improve the predict precision.

Keywords

Railway traffic; Seasonal adjustment; Echo State Network; SVR.

1. Introduction

Railway is the artery of national economy, and railway transportation is import for people to travel. According to statistical data, till the end of 2019, railway operating mileage in China has reached 139,000 kilometers in total and the total passenger's volume has reached 3.66 billion, accounting for 20.8% of the commercial passenger traffic of all traffic means. The rapid development of railway transportation requires better management of railway system. Forecasting railway passenger volume could provide data support for railway system operation, and it plays an important role in railway system planning. Therefore, in order to continue to optimize the service and operation of China's railway system, it is necessary to accurately forecast the railway passenger volume.

At present, the forecasting methods of railway passenger volume could be divided into two categories: the first category is using a single model to make predictions: Tang et al (2019) [1] has accurately predicted the monthly passenger volume for Chinese railways in 2016 using seasonal differential moving autoregressive (SARIMA) model; Wan et al. (2018) [2] has adapted the particle swarm optimization method into traditional LSTM model, and managed to forecast the annual passenger volume of railway with high accuracy. Wang et al. (2013) [3] taken the influence of traditional holidays such as Spring Festival on railway passenger volume into consideration, and established an seasonal adjustment model suitable for railway passenger volume in China based on X-12-ARIMA model. Zhang et al. (2009) [4] proposed an improved grey prediction model to forecast the railway passenger volume of Hubei Province; Xia et al. (2007) [5] adopted the improved support vector regression (SVR) to forecast the annual passenger volume of railway. Using a single model could ensure good prediction accuracy for annual railway passenger volume. However, the monthly railway passenger volume could be affected by so many factors, and it is often characterized by obvious trends, seasonality and

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202203_5(3).0044

irregularity, so it is difficult to fit with a single model. Therefore, some has combined various forecasting models to form a ensembled model forecasting method of railway monthly passenger volume: Qian et at al(2020)[7]has combined seasonal time series model (SARIMA) with the generalized autoregressive conditional heteroscedasticity (GARCH) model to form a SARIMA-GARCH ensemble model, SARIMA was used to fit the basic trend, and GARCH was used to predict the residual series, which accurately predicts the monthly passenger volume of China's railways.

To improve prediction accuracy, some scholars combined the seasonal decomposition method into prediction method. Firstly, the trend, seasonality and randomness of railway monthly passenger volume are decomposed, and then different models are adopted to forecast them. Qin and LAN (2019) [15] combined seasonal decomposition method with the echo state network, and adopted it to predict railway passenger volume in China. The key to establishing a prediction method combining seasonal decomposition method and model prediction method is to select the most suitable prediction model considering different characteristics of trend, seasonal and random components, and then sum up the prediction results of each component. However, in existing researches, the appropriate model has not been selected according to the different characteristics of each component. Based on this, this paper uses STL seasonal decomposition method to decompose the railway monthly passenger volume into three parts: trend component, seasonal component and residual component. For the three component time series obtained after decomposition, it can be found that the trend component shows some kind of continuous upward nonlinear change, while the adaptive enhanced echo state network AESN has simple calculation in training and good fitting for complex nonlinear trends, so it can be adopted to forecast trend component. The seasonal component changes according to a certain cycle and repeats every year. Moreover, the change intensity of each cycle is roughly the same, so the simple seasonal estimation method can be used to predict the seasonal components. The change of residual component shows obvious nonlinearity and no obvious regularity, while the SVR method of support vector regression can effectively solve the problems of small sample, nonlinearity and high-dimensional regression. The global optimal model can be obtained by support vector regression and the model has strong generalization ability, so it can be used to predict random components.

Therefore, based on STL seasonal decomposition method, adaptive regression echo state network model AESN and support vector regression SVR, this paper will establish a new combined model STL-AESN-SVR on railway monthly passenger volume. The STL-AESN-SVR method is chosen to fit the monthly passenger volume of China's railways from January 2005 to December 2018, and multi-step prediction is made for the monthly passenger volume of railways. Forecast monthly railway passengers' volume from January 2019 to December 2019, check validity for models by checking the accuracy of the forecast results.

2. Methodology

2.1. Echo State Networks with Adaboost

2.1.1. Brief Introduction of Echo State Network

ESN (Echo State Networks) is a special transition of the traditional circular neural network, which adopted a storage layer to replace the hidden layer in the original RNN. Different from the traditional RNN, only some weights, but not all weights, will be updated during ESN training. Therefore, ESN training is much simpler than traditional RNN. The common echo state network consists of three layers: Input layer, storage layer and output layer. Figure 1 shows an echo state network with k input neurons, n storage neurons and l output neurons.

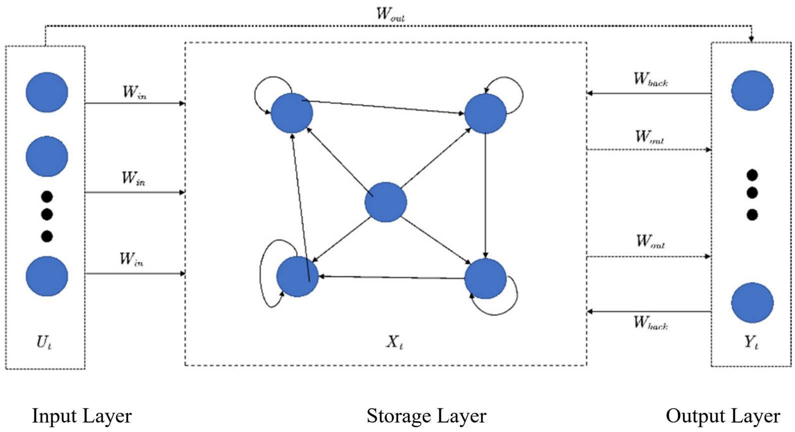

Figure 1. Echo State Network

When the echo network is used for the prediction of time series, for each period t, the input sequence $u_t=[u_{1t},u_{2t},u_{3t}...u_{Kt}]^T$, storage state sequence $x_t=[x_{1t},x_{2t},x_{3t}...x_{Nt}]^T$, output sequence $y_t=[y_{1t},y_{2t},y_{3t}...y_{Lt}]^T$. Storage state at time t+1 x_{t+1}.The update iteration can be performed by the following formula:

$$x_{t+1}=f^{re}(W_{in}u_{t+1}+Ws_t+W_{back}y_t) \tag{1}$$

W_{in} stands for the connecting weight matrix between the input layer and the storage layer, and its dimension is N*K, W stands for the connection weight matrix among neurons in the storage layer, and its dimension is N*N, W_{back} is the feedback connection weight matrix between output layer and storage layer, and its dimension is N*L. W_{in},W ,W_{back} are randomly generated matrices. f^{re} is the activation function of storage layer, which is generally set as hyperbolic tangent function by default. The input sequence and output sequence at time t=0 are initialized to 0, that is $u_0=[0,0,0...0]^T$, $y_0=[0,0,0...0]^T$.

The predicted output sequence of echo state network in T period can be calculated by $\hat{y}_t=W_{out}*f^{out}([x_t;u_t])$, in which f^{out} is the activation function of the output layer. W_{out} is the weight matrix connecting the input layer, the storage layer and the output layer, and is also the training target of the echo network. Normally, W_{out} can be calculated based on the least square method, however, this paper uses ridge regression, that is, the improved least square method to calculate W_{out}, then

$$W_{out}=Y^T(MM^T+\beta I)^{-1}M \tag{2}$$

It should be mentioned that $Y=[y_1,y_2,...,y_T]$, $M=[m_1,m_2,...,m_T]$, $m_t=[x_t;u_t]$, β is the regularization parameter, and t is the length of the training set time series.

According to formula (2),during the first T_0 periods, the state of the storage layer in each period is affected by the initial weight, while after that, the state of the storage layer is only related to the training samples, so we need to eliminate this influence during the training, that is, to remove all the training samples before T_0. Therefore, T_0 is also known as cleaning time. The final W_{out} could then be calculated by the following formula:

$$W_{out}=Y'^T(M'M'^T+\beta I)^{-1}M'$$

where $Y'=[y_{T_0}...,y_T]$, $M'=[m_{T_0}...,m_T]$.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202203_5(3).0044

2.1.2. Training of echo state Network

To ensure the prediction accuracy of the echo state network, the established echo network must have the nature of echo state, so when the matrixes W, W_{back}, W_{in} are randomly generated, certain procedures should be followed.

(1) Firstly, randomly generate matrixes W_0, W_{back}, W_{in}. Generally, the weight values are randomly generated from the uniform distribution of (-1,1).

(2) Standardize W_0 to get matrix W_1, $W_1 = W_0/|\lambda_{\max}|$, in which $|\lambda_{\max}|$ be the absolute value of the largest eigenvalue of W_0, which is also called the spectral radius of matrix W_0.

(3) Scale W_1 to get the matrix W, $W = \alpha W_0$, in which $\alpha < 1$, α is the spectral radius of the obtained matrix W.

If all these steps are followed during the initialization process of W, W_{back}, W_{in}, then the constructed echo state network will have the nature of echo state.

In addition to ensuring the existence of the nature of the echo state, the parameter setting is also very important for the training of the echo state network. The parameter setting of the echo state network should follow the following principles:

(1) To choose N, the dimension of W_0, we should consider training data size and the complexity of training tasks. Experience shows that N should be between T/10 and T/2.

(2) The selection of the spectral radius α also has a great influence on the properties of the echo state network, so several attempts are needed to find the best value.

(3) The cleaning duration T_0 should also be determined.

Therefore, the final training steps of the echo state network are as follows:

Step 1: determining parameters N,K,L of the echo state network, α, T_0, initialize the weight matrixs W_{in}, W, W_{back} as required.

Step 2: Update the state of the storage layer according to formula (1).

Step 3: Calculate the connection matrix W_{out} according to formula (3)

2.1.3. ESN with Adaboost

Because the connection matrix of echo state network is randomly generated, a single echo state network prediction model naturally has a randomness, which makes its accuracy unstable in the prediction task. Therefore, this paper introduces the AESN model, which combines the echo state network with Adaboost. The specific process of AESN is as follows:

Step 1: Input a training set Test, determine the maximum iteration number Max, use ESN as the base learner, and determine the threshold $\emptyset$ for judging the accuracy of the forecast. Set an initial weight for each training sample $J_i^t = 1/n$, $i = 1, 2 \ldots n$. N is the sample size of the training set.

Step 2: For t=1,...,Max:

(1) Calculate the adjustment error of the base learner model at each sample point. $e_i^t = |y_i - h_t(x_i)|/D_t$, in which: $D_t = \max_{j=1}^{n}|y_j - h_t(x_j)|$

(2) Calculate the total adjustment error: $\epsilon_t = \sum_{i=1}^{n} e_i^t w_i^t$. if $\epsilon_t < \emptyset$, stop the iteration and let Max=t-1.

(3) make $g_t = \epsilon_t/(1 - \epsilon_t)$

(4) Update weight vector $J_i^{t+1} = J_i^t g_t^{1-e_i^t}/Z_t$, where Z_t is the normalization factor.

Step 3: Output the final prediction model. $h_f(x)$, where $h_f(x)$ is the weighted average median of $h_t(x)$, using $\ln(1/g_t)$ as the weights.

2.2. Naïve Seasonal Forecasting Method

For highly seasonal data, we can simply take the observed value of the same season in the previous year as the predicted value of this season series. For example, for monthly data, if the seasonal cycle is one year, the value of January of the previous year can be used as the predicted value of January of this year, which is the idea of naïve seasonal prediction method.

2.3. Support Vector Regression SVR

SVR (Support Vector Regression) is the special form application of support vector machine to deal with regression problems. The basic idea of support vector regression is to project low-dimensional input features into a high-dimensional space after a nonlinear transformation. Consider a training set. $\{(t_1,z_1),\ldots,(t_l,z_l)\}$, in which $t_i\in\mathbb{R}^n$ is the input vector; $z_i\in\mathbb{R}^1$ is the target output. The SVR function can be written as: $f(t)=\omega^T\Phi(t)+b$, in which $\Phi(t)$ is a nonlinear function of the input vector, $f(t)$ is the predicted value of the model, ω and b are the coefficients of the model. To estimate the coefficients ω and b, loss function could be introduced:

$$R(C)=1/2\|\omega\|^2+C.\frac{1}{n}\sum_{i=1}^{n}|z_i-f(t_i)|_\varepsilon \tag{3}$$

Where $|z_i-f(t_i)|_\varepsilon=\begin{cases}0 & |z_i-f(t_i)|\leq\varepsilon\\ |z_i-f(t_i)|-\varepsilon & otherwise\end{cases}$, C is the penalty coefficient, ε is an insensitive interval.

Introduce two relaxation variables. ξ_i,ξ_i^* .After that, the original optimization problem can be rewritten as:

$$\begin{aligned}\min_{\omega,b,\xi,\xi^*}\ &\frac{1}{2\omega^T\omega}+C\sum_{i=1}^{l}\xi_i+C\sum_{i=1}^{l}\xi_i^* \\ &subject\ \ to{:}\omega^T\Phi(t_i)+b-z_i\leq\epsilon+\xi_i,\\ &\qquad z_{i-}\omega^T\Phi(t_i)-b\leq\epsilon+\xi_i^*,\\ &\qquad \xi_i,\xi_i^*\geq0,i=1,\ldots,l.\end{aligned} \tag{4}$$

The dual problem can be written as:

$$\min_{\alpha,\alpha^*}\ \frac{1}{2(\alpha-\alpha^*)^TQ(\alpha-\alpha^*)}+\epsilon\sum_{i=1}^{l}(\alpha_i+\alpha_i^*)+\sum_{i=1}^{l}z_i(\alpha_i-\alpha_i^*) \tag{5}$$

$$\begin{aligned}&subject\ \ to{:}e^T(\alpha_i-\alpha_i^*)=0,\\ &\qquad 0<\alpha_i,\alpha_i^*\leq C,\\ &\qquad i=1,\ldots,l\end{aligned}$$

where $e^T=[1,1,1\ldots,1]$ is a vector whose elements are all 1, and $Q=K(t_i,t_j)\equiv\Phi(t_i)^T\Phi(t_i)$ is a kernel function.

By solving the dual problem, we can get the regression function as follows:

$$f(t)=\sum_{i=1}^{l}(-\alpha_i+\alpha_i^*)K(t_i,t)+b$$

2.4. Seasonal Trend Decomposition Method STL based on Loess

STL (Seasonal-Trend Decomposition Procedure Based on Loess) is a general and robust method for decomposing time series. STL decomposition method is completed through two internal and external cycles, in which the internal cycle includes six steps: de-trend, periodic subsequence smoothing, low-flux filtering of periodic subsequence, trend removal of subsequence of smoothed period, deseasonalization and trend smoothing. After each internal cycle is executed, Both the seasonal component and the trend component will be updated once. The main function

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202203_5(3).0044

of the outer loop is calculating the robust weight, which will be used in the calculation process of the next inner loop to reduce the influence of outliers. Through STL decomposition, the original time series L_t will be broken down into three parts: trend component T_t, seasonal component S_t, residual component R_t. namely: $L_t = T_t + S_t + R_t$.

2.5. Construction of STL-AESN-SVR Model for Monthly Passenger Volume

When establishing STL-AESN-SVR combined model, the original sequence should be divided into three parts: trend component, seasonal component and random component by STL seasonal decomposition method. AESN model is adopted to forecast the trend component, naïve seasonal prediction method is used to predict the seasonal component, and SVR method is adopted to forecast the residual component, then sum the predicted values of three components to get final forecasting value of the model. Therefore, the STL-AENS-SVR model is used to calculate the railway passenger volume Y_t ,the specific steps of fitting and forecasting are as follows:

Step 1: STL decomposition. Using STL method to decompose the logarithmic time series log(Y_t), and simply calculate the trend component T_t, seasonal component S_t ,and residual component R_t.

Step 2: Data preprocessing. Normalize the three components in step 1 and divide them into training set and test set.

Step 3: Single model construction. Train a suitable AESN model for the trend component and a SVR model for the residual component based on the training set , and using naïve seasonal prediction method to predict the seasonal component.

Step 4: Build the ensemble model. Sum up the results of each component in step 3 to obtain the forecasting value of STL-AESN-SVR.

Step 5: Prediction and model evaluation of the ensemble model. The STL-AESN-SVR combined model is used for prediction, and the forecasting results are compared with the values of the test set to measure the accuracy of the model.

3. Model Establishment

In this paper, the monthly railway passenger volume data from January 2005 to December 2019 published by the Chinese National Bureau of Statistics official website is used as the data basis. All time series data are used in seasonal decomposition. When establishing the forecasting model, the monthly passenger traffic data from January 2005 to December 2018 are used as training data, and the monthly passenger traffic data from January 2019 to December 2019 are used as test data. RMSE (root of mean square error) and MAPE are used as evaluation indexes to measure the accuracy of the model.

(1) Root mean square error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y_i}\right)^2} \tag{6}$$

(2) Mean absolute percentage error (MAPE)

$$MAPE = \left[\frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i - \hat{y_i}\right|}{|y_i|}\right]*100 \tag{7}$$

3.1. STL Seasonal Decomposition

As can be seen from Figure 2, the monthly passenger volume of China's railways has a nonlinear long-term trend of upward growth, and at the same time, it also has great seasonal fluctuation, so it is difficult to obtain good results by directly modeling and forecasting it, so it is necessary to decompose it seasonally. In this paper, STL seasonal decomposition method is used to decompose it. STL is a decomposition model based on additive model. Through observation, it could be found that the relationship between monthly railway passenger volume and time is exponential, so the multiplication model should be more in line with the change trend of railway monthly passenger volume. Therefore, when the model is used to fit and forecast the monthly railway passenger volume, it is necessary to change the original STL method. The monthly passenger volume of the railway is L_t ,the trend component is T_t ,the seasonal component is S_t,the random component is R_t . Then the multiplication model can be expressed as $L_t = T_t * S_t * R_t$. Since the values of the time series of railway monthly passenger volume are greater than 0, the simple logarithmic transformation can be made at both ends of the original multiplication model $\log(L_t) = \log(T_t) + \log(S_t) + \log(R_t)$.The original multiplication model can be transformed into an addition model. Then the logarithmic time series $\log(L_t)$ could be decomposed by STL method, and the trend component T_t, seasonal componentS_t, residual component R_t could be obtained. And then the three components could be predicted respectively.

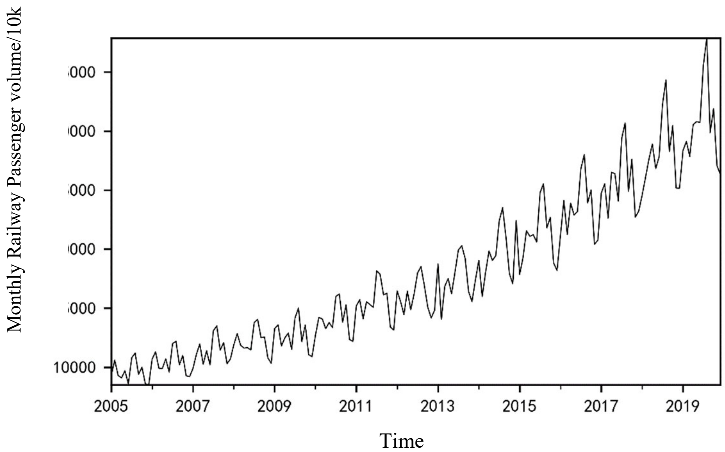

Figure 2. Monthly Railway Passenger volume

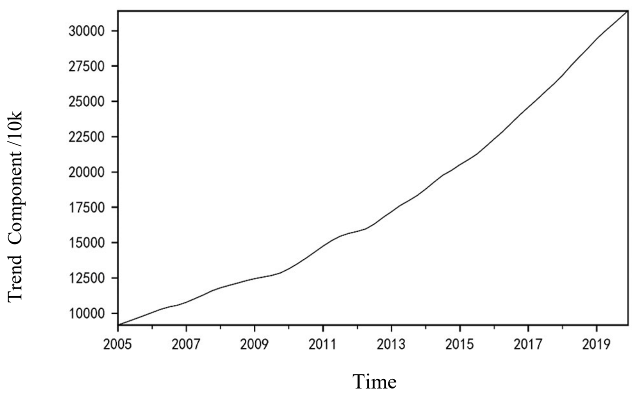

Figure 3. Trend Component

In this paper, STL decomposition is carried out by STL function in R language. The parameter to be set is s.windows = 13, and other parameters are taken as the default value of the function.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202203_5(3).0044

The results of STL decomposition of logarithmic series of railway passenger volume are shown in figures 3 to 5. After STL decomposition, the trend component T_t (Figure 3), seasonal component S_t (Figure 4), residual component R_t (Figure 5) are obtained. It can be seen from the decomposition results that the original time series not only has obvious trend fluctuation and seasonal variation, but also has obvious random disturbance.

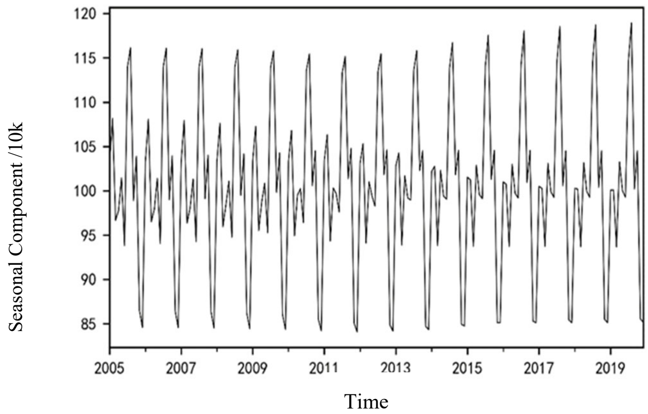

Figure 4. Seasonal Component

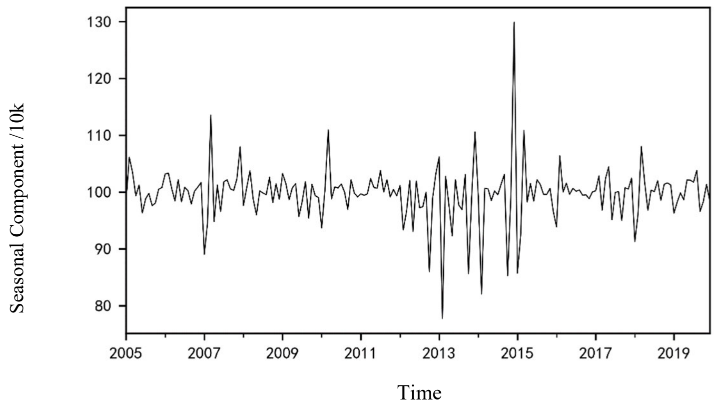

Figure 5. Residual Component

3.2. Prediction of Trend Components

The trend component T_t obtained by seasonal decomposition will be predicted by AESN model.

3.2.1. Data Normalization

The data normalization was carried out according to the following formula:

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{8}$$

Where $X_{\max}$ Is the largest among sample data,$X_{\min}$ is the smallest among sample data.

3.2.2. Construction of AESN Model

In this paper, the construction of the AESN model is realized by Python language, and the construction of the AESN model is realized by combining AdaboostRegressor in SKlearn library of Python with ESN function written in Python. When training the AESN model, the data from January 2005 to December 2018 was adopted as the training set and the data from January

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202203_5(3).0044

2019 to December 2019 was adopted as test set to measure the final accuracy of method. The model uses the lag term of the trend component as input and the output is the predicted value of the current trend component. After cross-validation, the final number of lag periods is 12, that is, the data from lag period 1 to lag period 12 is used as the input sequence. After the parameters are optimized by cross-validation, the number of base learners in the AESN model constructed in this paper is n = 10, and the maximum iteration times and threshold settings of AdaBoost algorithm adopt the default values of AdaboostRegressor. The number of neurons in the input layer K=12, the number of neurons in the storage layer N=10, and the spectral radius of the echo state network of the basic learner ESN,α should be set to 0.95, cleaning time T_0=25. Use AESN model to get the predicted value of the trend component, and denormalize the value to get the final forecasting value of the trend component, which is recorded as $\hat{T_t}$.

3.3. Prediction of Random Components

Support vector regression (SVR) could be adopted to forecast the random component, which is consistent with the prediction of the trend component, and it needs to be normalized before training. Similarly, the data from January 2005 to December 2018 will be adopted as the training set and the data from January 2019 to December 2019 will be adopted as the test set. The input of SVR model is the lag term of random component, and the output is the predicted value of current random component. After cross-validation, the parameters of SVR model are set as follows: kernel function $K(t_i,t_j)$ should be set to' RBF', penalty parameter C=0.05, relaxation coefficient ε=0.02, and the lag period of the random component is 8. After the SVR prediction result is denormalized, the final prediction value is recorded as $\hat{R_t}$.

3.4. Prediction of Seasonal Components.

For the prediction of seasonal components, we use naïve time series prediction method. That is, the seasonal component of the same month of last year is taken as the forecasting value of the seasonal component of current month, and the predicted value of the seasonal component is recorded as $\hat{S_t}$.

3.5. Use the Ensemble Model STL-AESN-SVR for Prediction

Table 1. Forecast results of different models for railway monthly passenger volume in 2019

Month	reality	SARIMA	ESN	SVR	STL-ESN	STL-ESN-SVR	STL-AESN	STL-AESN-SVR
Jan	28342	27797	25639	24719	26235	29138	26350	29327
Feb	29112	28928	27539	28326	27704	29090	27898	29367
Mar	27860	29882	28294	28065	29165	27981	29445	28258
Apr	30536	31323	29393	28161	30394	29997	30720	30326
May	30801	29565	29869	27570	28123	29792	28457	30171
Jun	30735	30266	29463	27605	29051	28761	29451	29204
Jul	35570	34407	30988	31092	33561	33781	34066	34339
Aug	37884	36264	33217	30537	35608	35074	36161	35718
Sep	29873	30703	33324	26447	29210	30067	29670	30694
Oct	31903	32775	31399	28726	31371	31574	31917	32253
Nov	27080	27953	30734	24627	25787	26335	26269	26966
Dec	26306	27982	28765	25121	25609	24996	26188	25668

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202203_5(3).0044

The final STL-AESN-SVR combination model can be obtained by integrating the prediction models established for trend components, seasonal components and random components. The model is adopted to forecast the railway passengers volume for 12 periods, and the forecasting value of the trend component obtained by AESN is recorded as $\hat{T_t}$, the forecasting value of the random component obtained by SVR can be recorded as $\hat{S_t}$, the forecasting value of seasonal component obtained by simple seasonal prediction method is recorded as $\hat{R_t}$, the final forecasting value of STL-AESN-SVR combined model can be determined by $\hat{Y_t} = \hat{T_t} * \hat{S_t} * \hat{R_t}$, the final forecasting results are shown in Table 1.

3.6. Conclusion

Compare the forecasting results of the combined model with those of the seasonal autoregressive model SARIMA, single ESN model, single SVR model, STL-ESN model, STL-ESN-SVR model and STL-AESN model. The comparison results are shown in Table 2. From the table, it can be seen that the prediction values of STL-AESN-SVR model proposed in this paper are closer to reality than those of other models.

In order to measure the prediction accuracy of each model more clearly, RMSE and MAPE of each model can be compared. As shown in Table 2, when single ESN model and SVR model are used to predict railway passenger volume, the values of error MAPE and RMSE are both larger. Compared with single prediction model, STL-AESN-SVR ensemble prediction model proposed in this paper greatly reduces RMSE and MAPE.It can be seen that the application of ensemble method improves the accuracy of single prediction model; Compared with combined forecasting methods such as STL-ESN-SVR ensemble forecasting model, STL-AESN-SVR ensemble forecasting model also performs better in RMSE,MAPE, which shows that combining echo state network with Adaboost can improve the forecasting accuracy of echo state network. Compared with STL-AESN model, STL-AESN-SVR model has reduced MPAE by nearly one percentage point, and its performance in RMSE index has also been improved to a certain extent. It can be seen that the introduction of SVR has played a positive role in improving the model accuracy. Compared with the traditional SARIMA prediction method, the STL-AESN-SVR model proposed in this paper also performs better in RMSE and MPAE.

Table 2. Comparison of prediction accuracy of different models

model	RMSE	MAPE(%)
SARIMA	1147.08	3.38
ESN	2709.65	7.38
SVR	3451.01	9.30
STL-ESN	1587.35	4.55
STL-ESN-SVR	1269.97	3.05
STL-AESN	1324.92	3.52
STL-AESN-SVR	974.56	2.45

4. Conclusion

Aiming at the important issue of railway passenger volume forecast, fully considering the trend, seasonality and randomness of railway passenger volume time series, an ensemble forecasting method of railway monthly passenger volume STL-AESN-SVR is proposed by combining STL seasonal adjustment method, echo state network with Adaboost, support vector regression and naïve seasonal forecasting method. By proposing this new model, the ways to forecast railway

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202203_5(3).0044

passenger volume are broadened. The empirical results show that STL-AESN-SVR ensemble forecasting method has high applicability, and it effectively improves the accuracy of railway monthly passenger volume forecasting. Therefore, its forecasting results can provide effective guidance for railway departments to carry out various passenger service management.

References

[1] TANG Yinying, ZHU Xinglong, LI Long. Monthly Railway Passenger Traffic Volume Forecasting Based on SARIMA Model [J]. Journal of Transportation Engineering and Information, 2019, 17(01): 25-32.

[2] LI Wan, FENG Fenling, JIANG Qiwei. Prediction for railway passenger volume based on modified PSO optimized LSTM neural network [J]. Journal of Railway Science and Engineering, 2018, 15(12): 3274-3280.

[3] WANG Zhi-hong, WANG Qian-yuan. Seasonal Adjustment Model of China Monthly Passenger Traffic Volume Based on Spring Festival Factors[J].Journal of the China Railway Society, 2013, 35(07):9-13.

[4] ZHANG Da-bin, ZHU Hou, Li Wei,et al. Application of a grey model based on Genetic Algorithm in Railway Passenger Flow Prediction [J]. Statistics & Decision, 2009 (24): 24-26.

[5] XIA Guoen, JIN Weidong, ZHANG Gexiang. Improved Support Vector Regression And Its Application to Prediction of Railway Passenger Traffic Volume [J]. Journal of Southwest Jiaotong University, 2007(04): 494-498.

[6] LU Xiaolan, ZHANG Kexin. An Integrated Prediction of Railway Passenger Traffic Volume Based on IOWGA Operator [J] Mathematics in Practice and Theory, 2020, 50(06): 120-131.

[7] QIAN Mingjun, LI Yinzhen, E Runa. Method on Monthly Railway Passenger Forecast Based on Seasonal Decomposition and SARIMA-GARCH Model [J]. Journal of the China Railway Society, 2020, 42(06): 25-34.

[8] Shiskin, J., Young, et al. The X-11 Variant of the Census Method II Seasonal Adjustment Program[R]. Technical Paper No. 15, U.S. Department of Commerce, Bureau of Economic Analysis,1967.

[9] Findley, D.F., B.C. Monsell, et al. New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program[J]. Journal of Business and Economic Statistics, 1998, (16).

[10] S Hylleberg, R Engle, C Granger, et al. Seasonal Integration and Cointegration[J]. Journal of Econometrics, 1990, (41).

[11] Cleveland R B, Cleveland W S, McRae J E, et al. STL: A seasonal-trend decomposition[J]. Journal of official statistics, 1990, 6(1): 3-73.

[12] Jaeger H. Echo State Network[J]. Scholarpedia, 2007, 2(9): 2330.

[13] ZHANG Ning-zhi, ZHOU Jia-li, SUN Wu-jun. Stock Index Prediction Based on Optimized Echo State Network[J].Journal of Industrial Engineering and Engineering Management,2014,28(01):94-101

[14] Chouikhi N, Ammar B,Rokbani N, et al. Pso-based Analysis of Echo State Network Parameters for Time Series Forecasting [J]. Applied Soft Computing, 2017, 55: 211-225.

[15] Wang L,Lv S, Zeng Y. Effective Sparse Adaboost Method with Esn and Foa for Industrial Electricity Consumption Forecasting in China[J]. Energy, 2018, 155: 1013-1031.

[16] Qin L, Li W, Li S. Effective Passenger Flow Forecasting Using Stl and Esn Based on Two Improvement Strategies[J]. Neurocomputing, 2019, 356: 244-256.