International Journal of u- and e- Service, Science and Technology
Vol.7, No.2 (2014), pp.101-112
http://dx.doi.org/10.14257/ijunesst.2014.7.2.10

A Partial Least Square Based Support Vector Regression Rail Transit Passenger Flow Prediction Method

Huijuan Zhou[1], Yong Qin[2] and Yinghong Li[1]

[1]Beijing Key Laboratory of Urban Intelligent Traffic Control Technology, North China University of Technology
No. 5 Jinuanzhuang Road, Shijingshan District, Beijing, China 100144
[2]State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University
No. 3 Shangyuancun, Haidian District, Beijing, China, 100044
zhhjuan@sina.com,qinyiong2146@126.com, lyh@ncut.edu.cn

Abstract

In this article, aiming at complex prediction problems of the rail transit passenger flow and prediction problems combined with the actual situation of rail transit in Beijing. We will propose a fusion model of passenger flow for predicting. It can improve the prediction accuracy for passenger flow forecasting. We use the partial least squares regression method to solve multicollinearity between the dependent variable. The method of principal component analysis can rescreen the all the factors which are affect the passenger flow. To extract comprehensive variable this has the best ability to explain the passenger flow from all of the information, in order to solve the relevance of statistics, noise and information redundancy. The nonlinear prediction model which is between comprehensive variable and passenger flow will be established. Finally, through the passenger flow forecasting of exchange station of Beijing Metro Line 1 to verify the effectiveness of the method.

Keywords: *Partial Least Square, Rail Transit Passenger Flow, Prediction, kernel function*

1. Introduction

Rail transportation is the main way to public transportation in big city .Analysis of actual data of rail transit passenger flow and forecast the future passenger flow of the road network that is a very important way to improve the operation organization efficiency or rail transportation management. That also can enhance transport capacity of the peak passenger flow, and provide the data support for the operation safety decision.

Researchers will divided the micro prediction method of traffic flow into linear prediction methods (such as the historical average forecast method) and nonlinear prediction methods (such as grey prediction method and so on).Linear prediction method has the advantages of simple calculations, but the accuracy is not high; Nonlinear prediction method has characteristic of high accuracy, but the computational complexity, and the influence of coping with multiple factors is not very good [1-9].

There exist many methods of linear sequence forecasting in the open literature, such as the regression forecasting method [10], the time series prediction method [11], *etc*. Almost all existing methods infer future trends on the basis of a changing time sequence in the past. Non-linear problems are the focus of research in many fields currently, and the forecasting methods includes the Jenkins forecasting method [12], the Markov forecasting method [13], the Grey forecasting method [14], the multi-objective forecasting method [15], *etc*. Recently, the neural network and the wavelet network have been applied in the transportation field.

ISSN: 2005-4246 IJUNESST

Though much progress has been made in non-linear forecasting, there are few studies discussing how to forecast the passenger flow in a transit system. Because of the regularity and randomness of the passenger flow rate, and traditional forecasting methods always lead to inaccurate results.

In this article, aiming at complex prediction problems of the rail transit passenger flow, and combined with the actual situation of rail transit in Beijing. We will propose a fusion model of passenger flow for predicting. It can improve the prediction accuracy for passenger flow forecasting. We use the partial least squares regression method to solve multicollinearity between the dependent variable. The method of principal component analysis can rescreen the all the factors which are affect the passenger flow. To extract comprehensive variable this has the best ability to explain the passenger flow from all of the information, in order to solve the relevance of statistics, noise and information redundancy. The nonlinear prediction model which is between comprehensive variable and passenger flow will be established. Finally, through the passenger flow forecasting of exchange station of Beijing Metro Line 1 to verify the effectiveness of the method.

2. The Basic Model of SVR

$LS-SVR$ expands standard SVR by optimizing the square of relaxation factors and converting the constraints of inequality to equality, so the quadratic programming problem in traditional SVR becomes linear simultaneous equations, thus the calculating difficulty reduces a lot in company with the solution high efficiency and convergence speeding up.

The basic method of SVR :

Define $x \in R^n$ and $y \in R$, let R^n be the input space, by nonlinear transformation $\phi(\cdot)$, we let in the input space x map into a high dimensional characteristic space where we use the linear function to fit sample data while making sure the generalization.

In the characteristic space, the linear estimation function is defined as:

$$y = f(x,\omega) = \omega^T \phi(x) + b \qquad (1)$$

Where ω is the weight and b is the skewness.

The aim function is:

$$\min_{\omega,b,\xi} J(\omega,b,\xi) = \frac{1}{2}\omega^T\omega + \frac{1}{2}C\sum_{i=1}^{N}\xi_i^2 \qquad (2)$$

s.t.

$$y_i = \phi(x_i)\omega + b + \xi_i \quad i = 1,\cdots,N \qquad (3)$$

Where $\omega \in R^h$ is the weight vector and $\phi(\cdot)$ is non-linear mapping function, $\xi_i \in R^{N\times1}$ is relaxation factor, $b \in R$ is the skewness while $C > 0$ is penalty factor.

Importing factors, $\alpha_i \in R^{N\times1}$, we can easily get the function as:

$$L(\omega,b,\xi_i,\alpha_i) = \frac{1}{2}\|\omega\|^2 + \frac{1}{2}C\sum_{i=1}^{N}\xi_i^2 - \sum_{i=1}^{N}\alpha_i[\phi(x_i)\omega + b + \xi_i - y_i] \qquad (4)$$

According to the KTT we get

$$\begin{cases} \frac{\partial L}{\partial \omega} = \omega - \sum_{i=1}^{N} \alpha_i \phi(x_i) = 0 \\ \frac{\partial L}{\partial b} = \sum_{i=1}^{N} \alpha_i = 0 \\ \frac{\partial L}{\partial \xi_i} = \alpha_i - C\xi_i = 0 \\ \frac{\partial L}{\partial \alpha_i} = \phi(x_i) + b + \xi_i - y_i = 0 \end{cases} \quad (5)$$

$$\begin{bmatrix} 0 & E^T \\ E & \phi\phi^T + C^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (6)$$

Where E is the matrix whose elements are all 1, I is a $N \times N$ identity matrix.

Inner product of regression in non-linear function can be replaced by kernel function satisfied $Mercer$. Let $\Omega_{ij} = \phi\phi^T$,

then

$$\Omega_{ij} = \phi(x_i)^T \phi(x_j) = K(x_i, x_j) \quad (7)$$

We then have the $LS-SVR$ regression function model

$$f(x) = \sum_{i=1}^{N} \alpha_i K(x_i, x_j) + b \quad (8)$$

3. The Types and Selection of Kernel Function

3.1 The types of Kernel function

In general, the kernel function is commonly used as the linear kernel function, polynomial kernel function, the radial basis kernel function, and sigmoid kernel function. The functions are as follows:

(1) Linear kernel function

$$K(x, x_i) = x * x_i$$

(2) Polynomial kernel function

$$K(x, x_i) = [(x * x_i) + 1]^d$$

Where the d is the order of the polynomial

(3) Radial basis kernel function

$$K(x, x_i) = \exp(-\|x - x_i\|^2 / 2\sigma^2)$$

Where the σ is width of the kernel function

(4) Sigmoid kernel function

$$K(x, x_i) = \tanh(\gamma(x * x_i) + c)$$

Commonly used kernel function can be divided into two categories: one category is the global kernel function, the other is local kernel functions: The linear kernel function, polynomial kernel function, and Sigmoid kernel function is a global common kernel function

3.2. The Performance of Kernel Function

The classification accuracy is an important indicator for measure the prediction model. The meaning of the classification accuracy is that the proportion which is divided into a category in the classification of all sample in the classification results of the model.

The formula can be expressed as follows:

$$classification\ accuracy = \frac{correct\quad samples}{all\quad samples}$$

Table 1. Classification Results of Models based on Different Kernel Functions

The type of kernel function	The classification accuracy of modeling sample (%)	The classification accuracy of training sample (%)
linear kernel function	81.27	76.87
polynomial kernel function	83.09	87.08
radial basis kernel function	89.37	81.33
Sigmoid kernel function	79.13	74.25

Such as: the classification accuracy of RBF kernel function of training samples is 81.33%, it said that the samples 407 samples were correctly classified into the category of samples in the 500 training sample of the daily passenger flow.

Comparison of the kernel functions of classification accuracy of modeling samples and testing samples in Table 1. RBF kernel function and polynomial kernel function has an obvious advantage over the other two kernel function. Among them, the polynomial kernel function is global kernel function. Compared with the RBF kernel function, the learning ability of the polynomial kernel function is relatively weak, but the generalization ability is stronger than the RBF kernel function; RBF kernel function belongs to a local kernel functions, it has the very strong learning ability that is obviously higher than polynomial kernel function. But the generalization ability is weaker than polynomial kernel function,

Estimated time Based on support vector machine model of four kinds of kernel function as shown in Table 2.

Table 2. Computing Time of Models based on Different Kernel Functions

The type of kernel function	Training time (s)	Test time (s)
linear kernel function	0.00433	0.00424
polynomial kernel function	0.00582	0.00455
radial basis kernel function	0.03522	0.01518
Sigmoid kernel function	0.05894	0.02312

The training time is the time of model training that use the grid search method to explore the optimal parameters, and use the optimal parameter penalty coefficient c and kernel function parameter g training the model.

Vapink [16] had pointed out that, in the Sigmoid kernel function, when the parameters take specific values, matrix of the Sigmoid function is non positive semi definite Only the matrix of kernel function is symmetric and positive semi definite, this function can satisfy the Mercer condition. Therefore, the sigmoid kernel function is a model that the accuracy of classification and test accuracy of classification is lowest. In the selection of kernel function, sigmoid kernel function is the lowest priority.

3.3 Selection of Kernel function

Based on the statistical learning theory, the realization of nonlinear transform is through the appropriate kernel function in the algorithm. No need to know the exact form, as long as the kernel function can satisfy the Mercer condition. According to the condition of kernel function, the result of two kernel functions added is the kernel function too [17].

This paper presents a new form function:

$$K_{mix} = \sqrt{\frac{\lambda}{2}} k_{poly} + (1 - \sqrt{\frac{\lambda}{2}}) K_{RBF} \quad (9)$$

Where, the parameter λ is a constant. It is used for adjusting the degree of combination function of polynomial kernel function and RBF kernel function. When the $\lambda \to 0$, the RBF kernel function is dominant in the combination function; When the $\lambda \to 2$, Polynomial kernel function is dominant in the combination function. In order to ensure the rationality of the combination function without losing the original mapping space, the value of λ must be in the range of $(0, 2)$.

The condition of a new function can be used as a condition of kernel function is:

$$\int_{x*x} K(x, x_i) f(x) f(z) dx dz \geq 0 \quad (10)$$

It requires the definition of matrix over an arbitrary finite point sets is positive semi definite. On the basis of condition (10), we can determine the function of new combination (9) satisfies the Mercer condition, and can be used as the kernel function of support vector machine.

4. The PLS-SVR Model with Mixed Kernel Function

The PLS method has been used in this research to reduce the size of the input data. Next, we will introduce the partial least squares algorithm, and use mixed kernel function and partial least square method to establish the prediction model of mixed kernel partial least squares support vector regression.

4.1. PLS Regression Model

PLS is a reasonably new Multivariate statistical method of data analysis, which is a method for constructing predictive models when the explanatory variables are many and highly collinear. Its main focus is to extract the potential components, which uses data of multiple dependent variables and independent variables for analyzing and modeling.

First, let E_0^T and F_0^T are separately the transposed matrix of E_0 and F_0, then we can obtain the eigenvector w_1 associated with the largest eigenvalue of the matrix $E_0^T F_0 F_0^T E_0$, the component t_1 is:

$$w_1 = \frac{E_0^T F_0}{\left\| E_0^T F_0 \right\|}; t_1 = E_0 w_1; p_1 = \frac{E_0^T t_1}{\left\| t_1 \right\|^2}; E_1 = E_0 - t_1 p_1^T \qquad (11)$$

In the same way, we can obtain the eigenvector w_h associated with the largest eigenvalue of the matrix $E_0^T F_0 F_0^T E_0$, the component t_h is:

$$\begin{cases} w_h = \dfrac{E_{h-1}^T F_{h-1}}{\left\| E_{h-1}^T F_{h-1} \right\|}; \\ t_h = E_{h-1} w_h; \\ p_h = \dfrac{E_{h-1}^T t_h}{\left\| t_h \right\|^2}; \\ E_h = E_{h-1} - t_h p_h^T \end{cases} \qquad (12)$$

If the rank of $X_{n \times p}$ is A, we can use cross validation method for identifying, then,

$$\begin{cases} E_0 = t_1 p_1' + \ldots + t_A p_A' \\ F_0 = t_1 r_1' + \ldots + t_A r_A' + F_A \end{cases} \qquad (13)$$

Where $r_1', \ldots, r_A'$ is a row vector of regression coefficient, F_A is error matrix. Concerning the least square regression equation of F_A is

$$\hat{F}_0 = t_1 r_1 + t_2 r_2 + \ldots + t_h r_h \qquad (14)$$

Because $t_1, t_2 \ldots, t_A$ can express as the linear combination of $E_{01}, E_{02} \ldots, E_{0A}$, Hence, according to the property of PLS regression：

$$t_i = E_{i-1} W_i = E_0 W_i^* \qquad (i = 1, 2, \ldots, h) \qquad (15)$$

Where

$$W_i^* = \prod_{k=1}^{i-1} (I - W_k P_k^T) W_i$$

Then equation (13) substitute into equation (12),

$$\begin{aligned} \hat{F}_0 &= r_1 E_0 W_1^* + r_2 E_0 W_2^* + \ldots + r_h E_0 W_h^* \\ &= E_0 (r_1 W_1^* + r_2 W_2^* + \ldots + r_h W_h^*) \end{aligned} \qquad (16)$$

Let $y^* = F_0, x_i^* = E_{0i}, \alpha_i = \sum_{k=1}^{h} r_k W_{ki}^* (i = 1, 2, \ldots, m)$

. Then, equation(14)can be revert to the regression equation of standardized variable as follows

$$\hat{y}^* = \alpha_1 x_1^* + \alpha_2 x_2^* + \ldots + \alpha_m x_m^* \qquad (17)$$

Equation (15) can be written down raw variable y , and its PLS regression equation of estimated value $\hat{y}$ is obtained [18-23].

4.2. PLS-SVR Model

The processing of $PLS-SVR$ is divided into following steps:

Step1: *PLS* for feature extraction of the raw data

From compute the equation (9) and equation (10) we can contain the vector t_i , p_i and w_i .They respectively constitute the score matrix $T_{train} = [t_1,..., t_h]$, load matrix $p = [p_1,..., p_h]$ and correlation coefficient matrix $w = [w_1,..., w_h]$ of training samples.

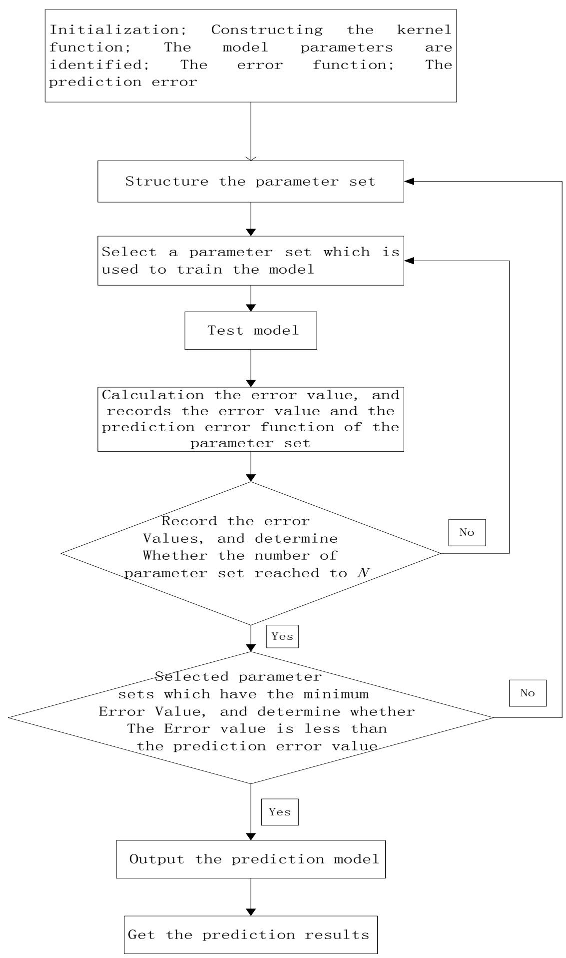

Figure 1. Prediction Process Figure of Improved Support Vector Machine Model

Step2: $LS-SVR$ Modeling

After h dimensions have been extracted, Which can use T_{train} , y_{train} train the $LS-SVR$ model, it contain $Lagrange$ multiplier and bias term b of the optimal parameter. On this basis, the following equation can be written,

$$\begin{bmatrix} 0 & E^T \\ E & \phi\phi^T + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y_{train} \end{bmatrix} \quad (18)$$

Then using equation (16) can result in coefficient b and α .

Step3: *PLS-SVR* model prediction

Calculate the prediction value of test sample data is

$$y_{predict}(t) = \sum_{i=1}^{N} \alpha_i K_{mix}(x_i, x_j) + b \quad (19)$$

$$K_{mix} = \sqrt{\frac{\lambda}{2}} k_{poly} + (1 - \sqrt{\frac{\lambda}{2}}) K_{RBF}$$

The prediction of flow diagram, as shown in Figure 1.

5. The Simulation and the Conclusion

Beijing Metro Line 1 is a busy and important line of rail transit. There are many factors interference the prediction accuracy of the rail transit passenger flow. Therefore, this paper selects the Beijing Metro Line 1, a total of 500 days of traffic data as statistical object, and analyzes its characteristics and changing trend. Then, we use the SVR model and PLS-SVR model which with the kernel function of mixed to forecast the passenger flow after 250 days. Finally, we will analysis and comparison the results.

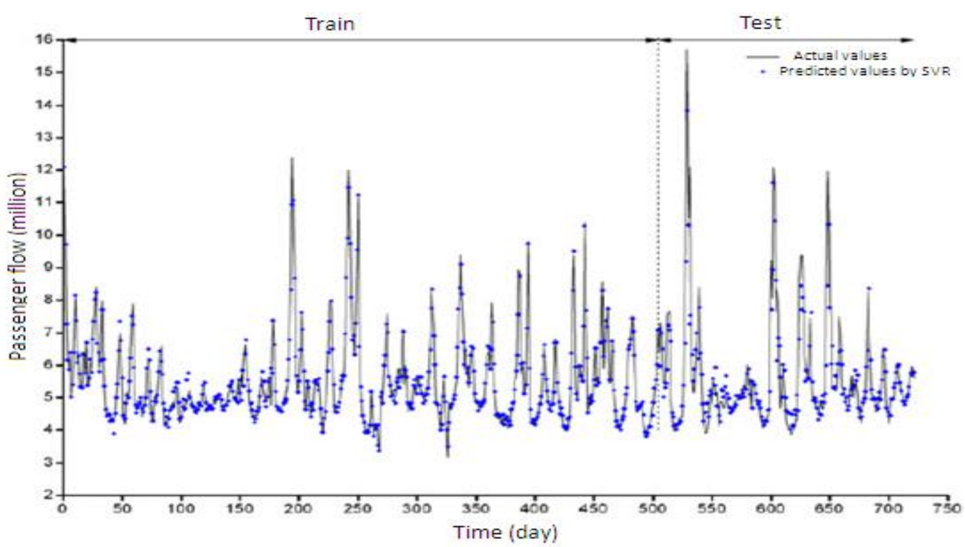

Figure 2. The Value of Passenger Flow Predicted by SVR Model

Correlations between prediction and actual values in the test stage have been determined using R^2 as shown in Figure 2 and Figure 3. Coefficients of determination were 0.696.. This means that the SVM model can predict rail transit passenger flow within an acceptable limit. However, grid search and modeling was found to be very much time consuming. Also, there has been a significant amount of multicollinearity among passenger flow predictors. Multicollinearity can make it difficult to correctly identify the most important contributors to a physical process. Also the large number of the predictors, reduce the learning rate of the prediction process, hence, size reduction methods like PLS, can be useful solution for this predicament.

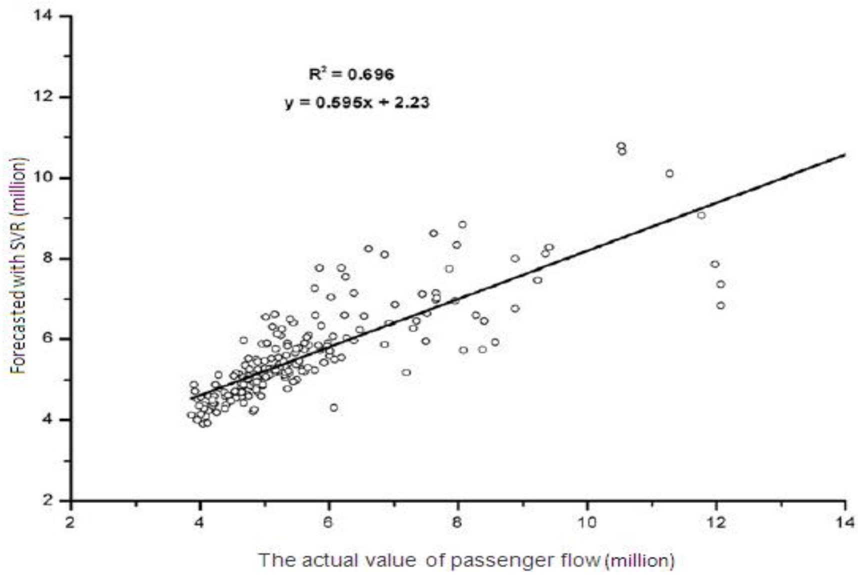

Figure 3. Correlation between Passenger Flow and SVR Predictions Predicted by SVR Model during Testing Stage

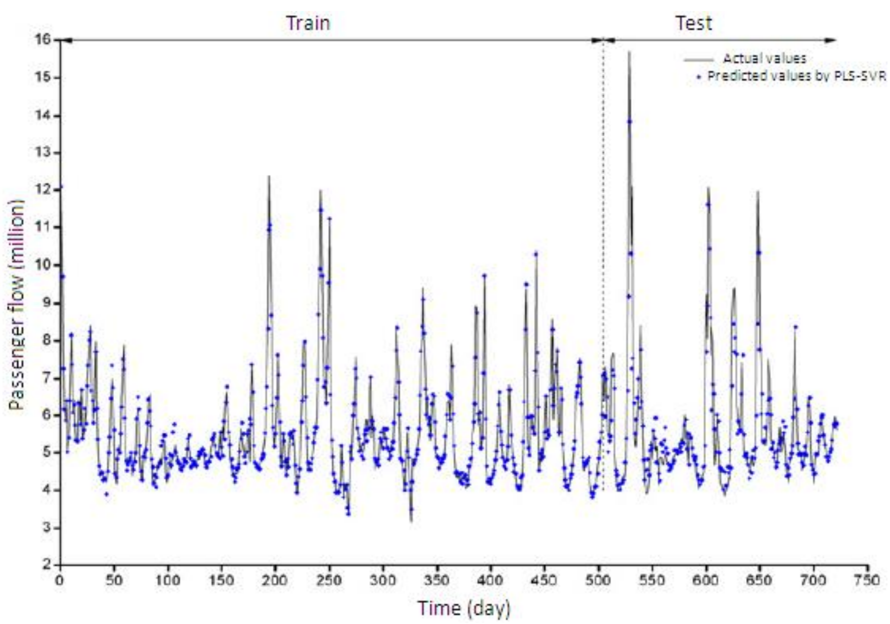

Figure 4. The Value of Passenger Flow Predicted by PLS-SVR Model

The PLS method has been used in this research to reduce the size of the input data. So we use this method to predict the passenger flow as below.

Similar to the SVM model, correlations between prediction and actual values in the test stage have been determined for PLS-SVR model as shown in Figure 4 and Figure 5. Coefficients of determination for daily prediction were found to be 0815.

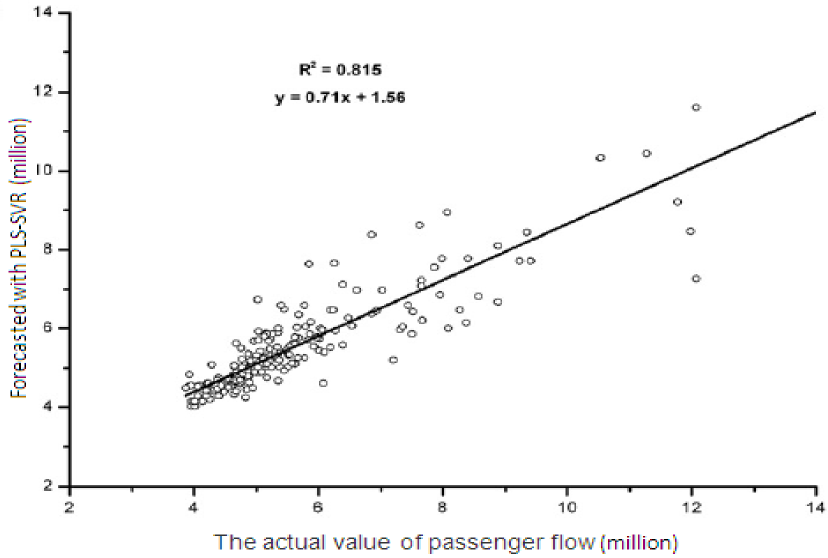

Figure 5. Correlation between Passenger Flow and PLS-SVR Predictions Predicted by SVR Model during Testing Stage

The results of learning and testing have been assessed on the basis of the Relative Mean Errors (RME), Root Mean Squared Errors (RMSE), Mean Absolute Relative Error (MARE) and coefficient of determination (R^2). These parameters have been defined as below:

$$RME = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - Y_i^*\right| \tag{20}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|Y_i - Y_i^*\right|^2} \tag{21}$$

$$MARE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - Y_i^*}{Y_i}\right| \tag{22}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - Y_i^*)^2}{\sum_{i=1}^{n}(Y_i - \overline{Y_i})^2} \tag{23}$$

Where Y_i is the actual value, $\bar{Y}$ is the average value, and Y_i^* is the predicted value.

Table 3. Comparison of SVR and PLS-SVR Models by Statistics Estimators

Estimator	PLS-SVR	SVR
RME	0.455	0.577
RMSE	0.743	0.925
MARE	0.045	0.084
R^2	0.884	0.702

Results of the SVR and hybrid PLS-SVR models have been compared and discussed statistically as show in table3 .It can be seen that the hybrid PLS-SVR model has higher coefficient of determination and fewer errors than the SVR model. Thus, the hybrid PLS-SVR model produced more accurate results and size reduction had positive effect on the model performance. Moreover, the hybrid PLS-SVR model achieves faster training speed and grid search. As it was observed, the total calculation time obviously is saved.

This paper has presented the prediction method of the daily Rail Transit Passenger Flow by applying the Support Vector Regression (SVR) and Partial Least Square (PLS). It can be concluded that such models provide a more promising alternative to time series forecasting. However, the important point of this approach is the data size reduction by PLS. The proposed hybrid PLS-SVM model provides a considerable improvement in the forecasting of Rail Transit Passenger Flow over the SVR model based on the same set of input variables.

Acknowledgements

This paper was supported by the State Key Laboratory of Rail Traffic Control and Safety (Contract No. RCS2011K011), Beijing Jiaotong University, the National High Technology Research and Development Program of China (863 Program) (No. 2012AA112401) and Beijing City Committee of Science and Technology Project (PXM2013_014212_000031).

References

[1] C. Chen, J. Chen and J. Barry, "Diurnal Pattern of Transit Ridership: a Case study of the New York city subway system", Journal of transport geography, vol. 17, no. 3, **(2009)**, pp. 176-186.

[2] X. Zhang and A. R. John, "Short-term travel time prediction using a time-varying coefficient linear model", Transportation Research Part C, vol. 11, no. 3, **(2003)**, pp. 187-210.

[3] E. I. Vlahogianni, J. C. Golia and M. G. Karlaftis, "Short-term traffic forecasting: Overview of objectives and methods", Transport reviews, vol. 24, no. 5, pp. 53-557.

[4] A. S. Alfa, "A Review of models for the temporal distribution of peak traffic demand", Transportion research, vol. 20, no. 6, **(1986)**, pp. 491-499.

[5] M. A. Abdel and R. Pemmanaboina, "Calibrating a real-time traffic crash prediction model using archived weather and its traffic data", IEEE transactions on intelligent Transportation systems, vol. 2, no. 7, **(2006)**, pp. 167-174.

[6] A. Simroth and Zhleh, "Travel time prediction using floating car data applied to logistics planning", IEEE transaction on intelligent transportation systems, vol. 12, no. 1, **(2011)**, pp. 243-353.

[7] C. H. Wu, J. M. Ho and D. T. Lee, "Travel-time prediction with support vector regression", IEEE transaction on intelligent transportation, vol. 5, no. 4, **(2004)**, pp. 276-281.

[8] S. Z. Zhao, "A new approach to the prediction of passenger flow in a transit system", Computers and mathematic with application, vol. 61, no. 8, **(2010)**, pp. 1968-1974.

[9] J. A. Suykens, J. D. Brabanter and L. Lukas, "Weighted least squares support vector machines: Robustness and sparse approximation", Neurocomputing, vol. 48, no. 1, **(2002)**, pp. 85-105.

[10] K. Nikolopoulos, P. Goodwin, A. Patelis and V. Assimakopoulos, "Forecasting with cue information: a comparison of multiple regression with alternativeforecasting approaches", European Journal of Operational Research, vol. 180, **(2007)**, pp. 354-368.

[11] T. Miyano, S. Kimoto, H. Shibuta, K. Nakashima, Y. Ikenaga and K. Aihara, "Time series analysis and prediction on complex dynamical behavior observed ina blast furnace", Physica D, vol. 135, **(2000)**, pp. 305-330.

[12] Y. Lu and S. M. AbouRizk, "Automated Box–Jenkins forecasting modeling", Automation in Construction, vol. 18, **(2009)**, pp. 547-558.

[13] Md. R. Hassan, "A combination of hidden Markov model and fuzzy model for stock market forecasting", Neurocomputing, vol. 72, **(2009)**, pp. 3439-3446.

[14] C.-T. Lin and S.-Y. Yang, "Forecast of the output value of Taiwan's opto-electronics industry using the Grey forecasting model", TechnologicalForecasting and Social Change, vol. 70, **(2003)**, pp. 177-186.

[15] P.-S. Yu, S.-T. Chen, C.-J. Chen and T.-C. Yang, "The potential of fuzzy multi-objective model for rainfall forecasting from typhoons", Natural Hazards, vol. 34, **(2005)**, pp. 131-150.

[16] V. Vapnik, "T he nature of statistica 1 learning theory", New York: Springer- Verlag, **(1995)**, pp. 25-37.

[17] K. Siwek, S. Osowski and R. Szupiluk, "Ensemble Neural Network Approach For AccurateLoad Forecasting In A Power System", Int. J. Appl. Math. Comput. Sci., vol. 19, no. 2, **(2009)**, pp. 303-315.
[18] K. E. Train, "Discrete Choice Methods with Simulation", London: Cambridge University Press, **(2001)**, pp. 56-70.
[19] A. H. David and H. G. William, "The mixed Logit model: the state of practice and warning for the unwary", Proceedings of Institute of Transportation Studies of Sydney University, Sydney: Sydney University Press, **(2001)**, pp. 12-14.
[20] S. Lv and X. C. Guo, "Passenger flow forecasting of railway", Systems Engineering-Theory and Practice, vol. 21, no. 8, **(2001)**, pp. 35-40.
[21] X. Q. Niu, W. Wang and Z. W. Yin, "Urban passenger trafficsharing forecasting method research", Road Traffic Technology, vol. 3, no. 3, **(2004)**, pp. 31-35.
[22] C. Q. Liu, "Planning of Modern Transport System", Beijing: China Communication Press, **(2001)**, pp. 20-30.
[23] Z. Q. Wang, "Time Value Research of Urban Traffic", Tongji University, **(2007)**.

Authors

Huijuan Zhou. She received her M.Sc. in Cartography & Geographic Information System (2002) from Beijing Normal University and PhD in System Analysis & Integration (2011) from Beijing Jiaotong University. Now she is assistant professor at Beijing Key Laboratory of Urban Intelligent Traffic Control Technology, North China University of Technology. Her current research interests include intelligent transportation system and traffic safety.

Yong Qin. He received his M.Sc. in Transport Automation and Conrol (1996) from Tongji University and PhD in Traffic Information and Control Engineering(1999) from China Academy of Railway Sciences. Now he is professor at State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. His current research interests include traffic information engineer and security.

Yinghong Li, Professor. Graduated from University of Science and Technology Beijing, received Ph.D in control science and Engineering. Now works in the Beijing Key Laboratory of Urban Intelligent Traffic Control Technology, North China University of Technology. The main research direction is the traffic information detection and its prediction technology.

