http://ansinet.com/itj

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (23): 7854-7858, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.7854.7858


# Regional Logistics Demand Forecasting Based on Lssvm with Improved Particle Swarm Optimization Algorithm

Quan Liang and Li-Yan Geng
School of Economics and Management, Shijiazhuang Tiedao University, Hebei, Shijiazhuang, 050043, China

**Abstract:** Regional logistics demand forecasting plays a crucial role in regional logistics infrastructure construction and regional economic development. Least Squares Support Vector Machines (LSSVM) has been widely applied to forecasting regional logistics demand. One of the main problems in LSSVM forecasting is the randomness and subjectivity of the parameters selection in LSSVM, which leads to the poor forecasting performance on regional logistics demand. To overcome the problem, this paper, integrating an Improved Particle Swarm Optimization (IPSO) algorithm into LSSVM, proposed a regional logistics demand forecasting model based on LSSVM-IPSO in which the IPSO algorithm was adopted to optimize the parameters of LSSVM and the LSSVM with the optimal parameters was used to forecast regional logistics demand. An example on Hebei logistics demand forecasting was taken to verify the effectiveness of the proposed model. The results show that IPSO algorithm effectively enhances the regional logistics demand forecasting accuracy of LSSVM model.

**Key words:** Regional logistics demand, forecasting, least squares support vector machines, improved particle swarm optimization algorithm

## INTRODUCTION

Regional logistics demand is derived from the social economy and affected by lots of factors. The impact of these factors on the regional is of complexity. So the complex nonlinear relationship is existed between the regional logistics demand and its affected factors. The usually used regional logistics demand forecasting methods including the time series method (Fite *et al*., 2002), the regression method (Adrangi *et al*., 2001), the grey prediction model (Huang and Feng, 2009), the production and transportation coefficient (Feng, 2013). These methods have difficulty to describe the nonlinear relationship. With the development of the computer technology, Artificial Neural Network (ANN) is introduced into regional logistics demand forecasting (Yin, 2010). The results show that ANN can describe the complex nonlinear relationship between regional logistics demand and its affected factors well. However, ANN often encounters some problems for itself theory which limits the application of ANN.

Support Vector Machines (SVM) is one of the latest machine learning methods for classification and regression (Vapnik, 1995). Based on statistical learning theory and structural risk minimization, SVM can overcome the problems lied in ANN and gains better regional logistics demand forecasting results relative to ANN (Huang *et al*., 2008; Luo *et al*., 2010). Least Squares Support Vector Machines (LSSVM), suggested by Suykens and Vandewalle (1999), is the reformation to standard SVM. LSSVM holds the advantages of SVM. Furthermore, LSSVM could simplify computation. Selection of parameters is important to generalization ability of LSSVM. However, LSSVM itself can't select the parameters and need to use appropriate method to determine them. Nowadays, there is no uniform rule.

Recent years, Particle Swarm Optimization (PSO) algorithm is widely employed to select the parameters of LSSVM and achieves satisfactory results. Furthermore, Researchers proposed different improved PSO algorithms to enhance the search performance and avoid falling into the local optimization. For example, Hu and Zeng (2007) proposed Two-Order Oscillating Particle Swarm Optimization (TOOPSO) algorithm and Chaturvedi *et al*. (2009) proposed Particle Swarm Optimization with Time Varying Acceleration Coefficients (PSOTVAC) algorithm. These two algorithms have been used to optimize the parameters of LSSVM for forecasting regional logistics demand (Geng *et al*., 2012; Geng and Dong, 2012). Xu and Chen (2013) proposed an Improved Particle Swarm optimization (IPSO) algorithm and applied this new algorithm to optimize the parameters of LSSVM. The

**Corresponding Author:** Quan Liang , School of Economics and Management, Shijiazhuang Tiedao University, Hebei, Shijiazhuang, 050043, China

experimental results showed that the proposed method can enhance fault identification of rolling bearing.

In this study, LSSVM model based on the IPSO algorithm is constructed to forecast regional logistics demand. That is, the parameters of LSSVM are optimized by IPSO algorithm and the optimized LSSVM is used to forecast regional logistics demand.

## METHODS

**Least squares support vector machines:** Given a set of training sample groups $(x_i, y_i)$ with i = 1, 2, …n. LSSVM maps the original data to a high-dimensional feature space using a selected kernel function. In the high-dimensional feature space, the nonlinear regression problem is transformed into the linear regression problem. The original optimized problem of LSSVM is defined as:

$$\min_{\omega,e} J(\omega, e) = \frac{1}{2}\|\omega\| + \frac{1}{2}\gamma\sum_{i=1}^{n} e_i^2 \quad (1)$$

$$y_i = \omega^T\varphi(x_i) + b + e_i,\ i = 1,\ldots, n. \quad (2)$$

where, $x_i \in R^d$ is the d-dimensional input vector; $y_i \in R$ is the corresponding 1-dimensional output variable; ω, b are the weight vector and bias constant value, respectively; $e_i$ is the error variable and γ is the regularization parameter.

On the basis of the Karush-Kuhn-Tucker (KKT) condition, construct Lagrangian function to transform the primal problem into the following matrix:

$$\begin{bmatrix} 0 & 1^T \\ 1 & O + I/\gamma \end{bmatrix}\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (3)$$

where, $\alpha = [\alpha_1, \alpha_2,\ldots, \alpha_n]^T$, $1 = [1,\ldots,1]^T$, $Y = [y_1, y_2,\ldots, y_n]^T$ and I is an unit matrix of order n. The matrix Ω is the kernel function matrix and the element in Ω is defined by $k(x_i,x) = \varphi(x_i)^T\varphi(x)$ for i = 1, 2, …, n. Here, the radial basis kernel function with the form of $k(x_i,x) = \exp(-\|x_i-x\|^2/\varphi^2)$ is considered, where $\varphi^2$ is the kernel parameter. Finally, the LSSVM model is given as:

$$y_i = \sum_{i=1}^{n} \alpha_i \exp(-\| x_i - x \|^2 / \sigma^2) + b \quad (4)$$

where $\alpha_i(i = 1,2,\ldots,n)$ are the Lagrange multipliers.

**Improved particle swarm algorithm:** There is a particle swarm composed of M particles. Each particle represents a potential solution in a D-dimensional search space of the problem and the jth particle has a velocity vector $V_j = (v_{j1}, v_{j2},\ldots, v_{jD})$ and a position vector $S_j = (s_{j1}, s_{j2},\ldots, s_{jD})$. Given the objective function of problem to be solved, each particle flies following the direction of its best previous position and its global best position to search the optimal solution.

Let *pbest* = $(p_{j1}, p_{j2},\ldots, p_{jD})$ and gbest = $(g_{j1}, g_{j2},\ldots, g_{jD})$ denote the best previous position of the jth particle itself and the best previous position of all particles of the swarm, the particle changes its position by the current velocity and the distance from *pbest* and *gbest*. Each particle updates its position according to the following equations:

$$v_{jD}^{t+1} = w \cdot v_{jD}^t + c_1 r_1 (p_{jD}^t - s_{jD}^t) + c_2 r_2 (g_{jD}^t - s_{jD}^t) \quad (5)$$

$$s_{jD}^{t+1} = s_{jD}^t + v_{jD}^{t+1} \quad (6)$$

where, $v_{jD}^t$ and $s_{jD}^t$ are the current velocity and position of the jth particle. Generally, the velocity and position of D-th dimension are restricted to $V_{max}$ and $S_{max}$. During the process of iteration, if the velocity and position exceed the limits, they will be set as $v_{jD}= V_{max}$, $s_{jD}=S_{max}$. And $c_1$, $c_2$ are cognitive and social coefficients; $r_1$, $r_2$ are two distinct random values in the range of [0,1].

And, w is inertia weight which is a key parameter for enhancing the search ability of PSO algorithm. By modifying the inertia weight w, the IPSO algorithm can enhance the search accuracy effectively and get a steady improvement of solution quality. The inertia weight w in IPSO algorithm is nonlinear diminishing with the iteration process:

$$w = (1 - \frac{t}{t_{max}})^2 (w_{max} - w_{min}) + w_{min} \quad (7)$$

where, $w_{max}$, $w_{min}$ are the maximum and minimum inertia weight respectively; t and $t_{max}$ are the number of the current and maximum iteration, respectively.

**Parameters of LSSVM optimized by IPSO:** The two parameters: γ and $\sigma^2$ are important for forecasting ability of LSSVM. This paper uses IPSO to optimize these two parameters. The steps of the parameters of LSSVM optimized by IPSO algorithm (LSSVM-IPSO) for forecasting are given as follows:

**Step 1:** Preprocessing of the data. The whole data groups are normalized through the mean and standard deviation of each variable

**Step 2:** Initialization of the particles swarm. Set the swarm scale M and the other parameters encompassing acceleration coefficients $c_1$ and $c_2$, the maximum and minimum inertia weight $w_{max}$ and $w_{min}$, the number of maximum iteration $t_{max}$

**Step 3:** Definition of the Fitness function. The Fitness function is defined as the mean squared error of LSSVM-IPSO model:

$$F = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (8)$$

where $y_i$, $\hat{y}_i$ are the actual and forecasted values from the training set, respectively. And i is the number of the training samples in the training set.

**Step 4:** Searching the optimal position. Calculate the fitness function value of each particle on the basis of (8). Update the optimal particle position as the *pbest* position. And update the global optimal particle position as the *gbest* position. The inertia weight w is updated according to (7)

**Step 5:** Stopping condition. If the stopping condition reaches, the evolutionary process is terminated. The global optimum particle position corresponds to the optimal parameters ($\gamma^*$,$\sigma^{2*}$), otherwise, t = t+1, go back to step 3

**Step 6:** LSSVM model is constructed using the obtained optimal parameters ($\gamma^*$, $\sigma^{2*}$) for forecasting regional logistics demand. Finally, the forecasted values need to be transformed into the original logistics demand forecasts

## EXPERIMENT RESULTS

**Index selection:** Some different indicators for measuring the regional logistics demand were used in different literatures. In this study, the freight traffic volume ($x_0$) is the choice to measure the regional logistics demand, quantitatively.

Generally, the factors affecting the regional logistics demand are divided into four categories, economic factors, industry factors, environment factors and other factors. As a derived demand of social economy, the change of the regional logistics demand is mainly affected by various regional economic indices. So the paper forecasts the regional logistics demand using the regional economic factors. The regional economic factors affecting the regional logistics demand including the following nine indices: The regional gross domestic product ($x_1$),the regional total output of the primary industry ($x_2$), the regional total output of the second industry ($x_3$), the regional total output of tertiary industry ($x_4$), the regional total investment in fixed assets ($x_5$), the total value of the regional import and export by customs ($x_6$), the total retail sales of the regional consumer goods ($x_7$), the regional household consumption ($x_8$), the regional retail price index ($x_9$).

The data used in the experiment is yearly logistics data provided by Heibei Bureau of Statistics encompassing the period from 1980 to 2008. The whole data set is split into a training set and a test set, in which the training set contains the period from 1980 to 2002 for training the model and the test set contains the period from 2003 to 2008 for verifying the performance of the model.

**Experiment process:** The parameters of IPSO algorithm itself are given as follows: Swarm scale, M = 10; acceleration coefficients, $c_1$=2 and $c_2$ = 2; maximum inertia weight factors, $w_{max}$ = 0.9 and $w_{min}$ = 0.1; the number of the maximum iteration $t_{max}$=30.

Meanwhile, in order to verify the effectiveness of LSSVM-IPSO model, the three particle swarm optimization algorithms: TOOPSO algorithm, PSOTVAC algorithm and standard PSO algorithm, are utilized to select the optimal parameters ($\gamma$, $\sigma^2$) in LSSVM and the established models are denoted by LSSVM-TOOPSO model, LSSVM-PSOTVAC model and LSSVM-PSO model, respectively. In LSSVM-TOOPSO model, the parameters of the TOOPSO algorithm itself are set as M = 10; $c_1$ = 0.2, $c_2$=1.8; $w_{max}$= 0.9, $w_{min}$ = 0.1; $t_{max}$=30. In LSSVM-PSOTVAC model, the parameters of the PSOTVAC algorithm itself are set as M = 10; $w_{max}$ = 0.9, $w_{min}$ = 0.1; $c_{1,ini}$ = 2.5, $c_{1,fin}$ = 0.5, $c_{2,ini}$ = 0.5, $c_{2,fin}$ = 2.5; $t_{max}$=30. In LSSVM-PSO model, the parameters of the PSO algorithm itself are set as the same as those of IPSO algorithm in LSSVM-IPSO model. The regional logistics demand is forecasted using these three models and the forecasting results are compared with those of LSSVM-IPSO model.

**Experiment results analysis:** Table 1 lists the forecasting results of the four models and Fig. 1 gives comparison of the forecasts curves of the four models. It is clear in Table 1 that from 2003 to 2007, LSSVM-IPSO model has smaller relative forecasting error comparing with the other three models. In 2008, LSSVM-PSO has the smallest relative forecasting error among the four models. In other words, as a whole, the LSSVM-IPSO model provides higher forecasting accuracy for the regional logistics demand than the other three models.

As shown in Fig. 1, the forecasts curves of the four models present the obviously similar changing tendency and the freight traffic volume increases over time. But the forecasts provided by LSSVM-IPSO are closer to the actual values than the other three models. And the forecasts provided by LSSVM-TOOPSO model and LSSVM-PSOTVAC model are very close to each other which is also verified in Table 1.

The demonstration of the performance of the four models is performed using following four statistical

Table 1: Regional logistics demand forecasts and relative forecasting errors by four models

| | | LSSVM-IPSO | | LSSVM-TOOPSO | | LSSVM-PSOTVAC | | LSSVM-PSO | |
|---|---|---|---|---|---|---|---|---|---|
| Year | Actual values /10kt | Forecasts /10kt | Relative error/% | Forecasts /10kt | Relative error/% | Forecasts /10kt | Relative error/% | Forecasts /10kt | Relative error/% |
| 2003 | 80551 | 76121 | -5.50 | 75573 | 6.18 | 75664 | 6.07 | 75367 | 6.44 |
| 2004 | 87265 | 82898 | -5.00 | 82118 | 5.90 | 82249 | 5.75 | 81825 | -6.23 |
| 2005 | 91330 | 88348 | -3.27 | 87393 | 4.31 | 87554 | 4.13 | 87035 | -4.70 |
| 2006 | 96784 | 94223 | -2.65 | 93079 | 3.83 | 93272 | 3.63 | 92649 | -4.27 |
| 2007 | 104188 | 104307 | 0.11 | 102836 | 1.30 | 103085 | 1.06 | 102281 | -1.83 |
| 2008 | 111383 | 118514 | 6.40 | 116608 | 4.69 | 116934 | 4.98 | 115882 | 4.04 |

[kt]Denotes kilo-tons. R.E. denotes relative forecasting error. LSSVM-IPSO, LSSVM-TOOPSO, LSSVM-PSOTVAC and LSSVM-PSO denote LSSVM model optimized by IPSO algorithm, TOOPSO algorithm, PSOTVAC algorithm and PSO algorithm, respectively

Table 2: NMSE, NMAE, MPE, and THEIL of regional logistics demand forecasts by four models

| Models | NMSE | NMAE | MPE | THEIL |
|---|---|---|---|---|
| LSSVM-IPSO | 0.7037 | 0.6241 | 0.0382 | 0.0219 |
| LSSVM-TOOPSO | 0.7189 | 0.7037 | 0.0437 | 0.0225 |
| LSSVM-PSOTVAC | 0.7125 | 0.6892 | 0.0427 | 0.0223 |
| LSSVM-PSO | 0.7392 | 0.7359 | 0.0459 | 0.0232 |

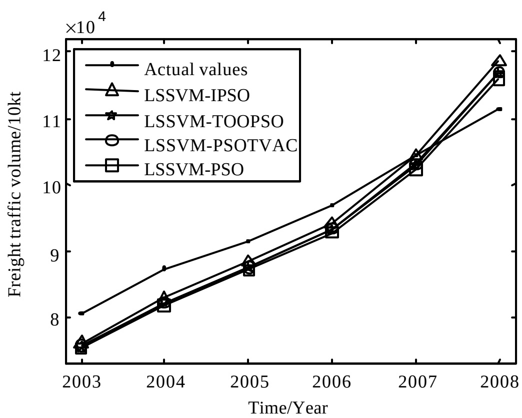


Fig. 1: Comparison of regional logistics demand forecasts by four models

criteria: The Normalized Mean Squared Error (NMSE), the Normalized Mean Absolute Error (NMAE), the Mean Percentage Error (MPE) and Theil statistic (THEIL). These four criteria are expressed as:

$$\text{NMSE}=\left(\sum_{k=1}^{K}(\hat{y}_k - y_k)^2 \Big/ \sum_{k=1}^{K}(y_{k-1} - y_k)^2\right)^{1/2} \tag{9}$$

$$\text{NMAE}=\sum_{k=1}^{K}\left|\hat{y}_k - y_k\right| \Big/ \sum_{k=1}^{K}\left|y_{k-1} - y_k\right| \tag{10}$$

$$\text{MPE}=K^{-1}\sum_{k=1}^{K}\left|(\hat{y}_k - y_k)/y_k\right| \tag{11}$$

$$\text{THEIL} = \frac{(K^{-1}\sum_{k=1}^{K}(\hat{y}_k - y_k)^2)^{1/2}}{(K^{-1}\sum_{k=1}^{K}(\hat{y}_k)^2)^{1/2} + (K^{-1}\sum_{k=1}^{P}(y_k)^2)^{1/2}} \tag{12}$$

where $y_k$ and $y_k$ are the actual freight traffic volume and the forecasted freight traffic volume of different models, respectively. And K is the number of the forecasted freight traffic volume. Table 2 contains the results.

It is observed from Table 2 that these four statistical criteria values of LSSVM-IPSO are all smaller than those of LSSVM-TOOPSO model, LSSVM-PSOTVAC model and LSSVM-PSO model indicating that contrasting with the other three models, LSSVM-IPSO model has superior performance in terms of regional logistics demand forecasting. Additionally, it can also been seen from the four statistical criteria values that LSSVM-PSOTVAC model has better forecasting performance than LSSVM-TOOPSO model. The forecasting performance of LSSVM-PSO model is the worst among the four models.

## CONCLUSION

In this study, IPSO algorithm is used to optimize the parameters of LSSVM and a LSSVM-IPSO model is built to forecast the regional logistics demand. An example on forecasting Hebei logistics demand is conducted to verify the built model. From the experimental results, it can be shown that LSSVM-IPSO model has been successfully used to regional logistics demand forecasting. It gains good forecasting performance and provides higher regional logistics demand forecasting accuracy contrasting with LSSVM-TOOPSO model, LSSVM-PSOTVAC model and LSSVM-PSO model. Therefore, LSSVM-IPSO model is an alternative approach for forecasting regional logistics demand. Further works could focus on employing other version of improved PSO algorithms to optimize the parameters of LSSVM and improve the forecasting performance of LSSVM further.

## REFERENCES

Adrangi, B., A. Chatrath and K. Raffiee, 2001. The demand for US air transport service: A chaos and nonlinearity investigation. Transp. Res. E, 37: 337-353.

Chaturvedi, K.T., M. Pandit and L. Srivastava, 2009. Particle swarm optimization with time varying acceleration coefficients for non-convex economic power dispatch. Int. J. Electr. Power Energy Syst., 31: 249-257.

Feng, S.Z., 2013. Forecasting of logistics demand based on production and transportation coefficient. Logist. Technol., 32: 91-93.

Fite, J.T., G.D. Taylor, J.S. Usher, J.R. English and J.N. Roberts, 2002. Forecasting freight demand using economic indices. Int. J. Phys. Distrib. Logistics Manage., 34: 299-308.

Geng, L.Y. and Q.T. Dong, 2012. Forecast of regional logistics demand using KPCA-based LSSVMs optimized by PSOTVAC. Adv. Inform. Sci. Serv. Sci., 4: 313-319.

Geng, L.Y., P. Zhao and Z.F. Zhang, 2012. Logistics demand forecasting based on LSSVM optimized by two-order oscillating PSO. Appl. Res. Comput., 29: 2558-2560.

Hu, J.X. and J.C. Zeng, 2007. A two-order particle swarm optimization model. J. Comput. Res. Dev., 44: 1825-1831.

Huang, H., G.F. Jiang, Y.S. Yan, B.S. Liao and G.E. Xia, 2008. Model of regional logistics demand forecasts based on support vector regional and its application. Appl. Res. Comput., 25: 2738-2740.

Huang, M.Z. and Y.B. Feng, 2009. The grey prediction model applies to regional logistics demand forecasting. Logist. Technol., 32: 17-20.

Luo, S.G., S. Ye and R. Hu, 2010. A research of forecasting the logistics amount based on Multi-output support vector regression. J. East China Jiaotong Univ., 27: 67-71.

Suykens, J.A.K. and J. Vandewalle, 1999. Least squares support vector machine classifiers. Neural Process. Lett., 9: 293-300.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. 1st Edn., Springer-Verlag, New York, USA.

Xu, H.B. and G.H. Chen, 2013. An intelligent fault identification method of rolling bearings based on LSSVM optimized by improved PSO. Mech. Syst. Signal Process, 35: 167-175.

Yin, Y.L., 2010. Logistics demand forecast based on adaptive neural network. J. Henan Polytechnic Univ. (Nat. Sci.), 29: 700-704.