TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.5, May 2014, pp. 3483 ~ 3490
DOI: http://dx.doi.org/10.11591/telkomnika.v12i5.3238

Demand Forecasting Model of Port Critical Spare Parts

Zhijie Song[1] , Zan Fu*[1], Han Wang[1], Guibin Hou[2]
[1]Department of Economics and Management, Yanshan University, Hebei, 066004, China
[2]Qinhuangdao Port Co., Ltd., Hebei, 066002, China
*Corresponding author, e-mail: 53115409@qq.com

Abstract
Demand forecasting for port critical spare parts (CSP) is notoriously difficult as it is expensive, lumpy and intermittent with high variability. In this paper, some influential factors which have an effect on CSP consumption were proposed according to port CSP characteristics and historical data. And analytic hierarchy process (AHP) is used to sieve out the more influential factors. Combined with the influential factors, a least squares support vector machines (LS-SVM) model optimized by particle swarm optimization (PSO) was developed to forecast the demand. And the effectiveness of the model is demonstrated through a real case study, which shows that the proposed model can forecast the demand of port CSP more accurately, and effectively reduce inventory backlog.

Keywords: *spare parts, demand forecasting, analytic hierarchy process (AHP), least squares support vector machines (LS-SVM), particle swarm optimization (PSO)*

1. Introduction

Port enterprises play a very important role in economic development, and they also play a significant role in the whole logistic chain. In order to avoid downtime of the equipment due to spare parts shortage, port enterprises tend to store large amounts of spare parts, which take up a lot of inventory capital, but even so, spare parts shortage phenomenon still occurs frequently. Therefore, predicting spare parts demand has become the key to solve these problems [1]. In this paper, we focus on the critical spare parts, which are more important to the equipment and occupied more inventory capital. But CSP have the characteristics of multiple influence factors, non-linearity and high variability, which brings difficulties to predict the demand.

Researchers have developed many forecasting techniques in the last decades, such as time series prediction method, Croston's method, Bootstrapping, neural network [2, 3] and so on. Li and Kuo applied enhanced fuzzy neural network (EFNN) to forecast the demand for automobile spare parts in a central warehouse, using analytic hierarchy process (AHP) method to determine factor's weight. The experimental results show that EFNN outperforms other models in fill rate and stock cost measures [4]. Hua and Zhang proposed an application of support vector machines (SVM) regression method for forecasting spare parts demand, which aiming firstly to forecast the occurrence of nonzero demands, and then to estimate lead-time demand. Their test using real data sets of 30 kinds of spare parts from a petrochemical enterprise in China suggested this method perform better than Croston's, bootstrapping and other methods [5].

Based on the above literatures, there are not many investigations focused on the CSP requirement prediction. Investigations on port industries are even fewer. In general, there is no appropriate forecasting model for predicting the requirement of port CSP. Further more, no methods have been previously proposed that make full use of real data based on port CSP relative factors, as we do in this study. In this article, Least Squares Support Vector Machines (LS-SVM) [6] regression, a semi-parametric modeling technique, is used to predict the port CSP demand. Firstly, we proposed some influential factors which have an effect on CSP consumption after analyzing port CSP characteristics and historical data. And applied analytic hierarchy process (AHP) to sieve out the more influential factors as the inputs of LS-SVM model. Secondly, aiming at the parameter optimization problem in LS-SVM, the particle swarm optimization (PSO) algorithm was adopted to optimize the parameter and improve the learning performance and generalization ability of LS-SVM model. The proposed PSO-LSSVM model

Received May 22, 2013; Revised December 9, 2013; Accepted December 29, 2013

takes into account influential factors of port CSP, and the real case study with real data in Qinhuangdao port illustrates the effectiveness of the model.

2. Research Method

2.1. Analytic Hierarchy Process

Analytic hierarchy process (AHP) was proposed by Thomas L.Saaty in 1980 [7], which is one of widely used multi-criteria decision-making methods. AHP involves the principles of decomposition, pair-wise comparisons, and priority vector generation and synthesis. In this study, the AHP method was used to select influence factors of the port CSP and determine the relative importance.

The three main steps of AHP are shown as follows:

Step 1: Construction of hierarchical structure.

Step 2: Calculation of weights between factors at each hierarchical level.

Step 3: Calculation of the overall hierarchical weights.

Ask evaluators to make pair-wise comparisons of the relative importance of variables using the scale. Based on the results of the questionnaire, a pair-wise comparison matrix is constructed to calculate the characteristic values and the characteristic vectors, thereby examining the consistency of the matrix to derive a consistency index ($C.I$). For each alternative, the consistency ratio ($C.R$) is measured by the ratio of the consistency index to the random index (RI). The equations are as follows:

$$C.I = \frac{\lambda_{max} - n}{n - 1}, C.R = \frac{CI}{RI}. \tag{1}$$

Generally, the value of $C.R$ should be less than 0.1 to guarantee consistency. If consistency does not comply with the requirement, it means that judgments made are inconsistent. And the researcher shall explain the problem of every pair-wise comparison. After calculating weights of every factor, we can obtain the more influential factors as the inputs of LS-SVM model.

2.2. Least Squares Support Vector Machines

Least squares support vector machines (LS-SVM) is a modification of the standard support vector machine (SVM) and was develop by Suykens [6]. LS-SVM is used for the optimal control of non-linear systems for classification as well as regression.

Given the sample of $D = \left\{\left(x_k, y_k\right)\right\}, k = 1,2,\dots,N$, with input vectors $x_k \in R^n$ and output values $y_k \in R$. The goal is to estimate a model of the form:

$$y(x) = \omega^T \cdot \phi(x) + b \tag{2}$$

Where $\phi(\cdot)$ is the mapping to a high dimensional feature space. Combine the functional complexity and fitting error, the optimization problem of LS-SVM is given as:

$$\begin{aligned} &\min_{\omega,b,e} Q\left(\omega,b,e\right) = \frac{1}{2}\left\|\omega\right\|^2 + \frac{\gamma}{2}\sum_{k=1}^{N} e_k^2 \quad \gamma > 0 \\ &\text{s.t. } y_k = \omega^T\phi(x_k) + b + e_k \quad k = 1,2,\dots,N \end{aligned} \tag{3}$$

This formulation consists of equality instead of inequality constraints. Constructing the Lagrangian:

$$L\left(\omega,b,e,a\right) = Q\left(\omega,b,e\right) - \sum_{k=1}^{N} a_k\left[\omega^T\phi(x_k) + b + e_k - y_k\right] \tag{4}$$

Where $a_k \in R$ are the Langrange multipliers. Setting the partial derivatives of $L(\omega, b, e, a)$ respect to ω, b, e and a to be 0, we get:

$$\begin{pmatrix} 0 & 1_N^T \\ 1_N & \Omega + \gamma^{-1} I_N \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \tag{5}$$

With $1_N = [1, \ldots, 1]$, $\alpha = [\alpha_1, \alpha_2, \ldots \alpha_N]$, $y = [y_1, y_2, \ldots y_N]$, and Mercer's condition is applied within the Ω matrix:

$$\Omega_{ij} = \phi(x_i)^T \phi(x_j) \quad i, j = 1, 2, \ldots, N \tag{6}$$

The output of LS-SVM regression is:

$$y(x) = \sum_{k=1}^{N} a_k K(x_i, x_j) + b \tag{7}$$

Where a_k and b are the solutions to the linear system. Note that in the case of RBF Kernel, one has only two additional tuning parameter which is γ and σ, γ stands for the weight at which the testing errors will be treated in relation to the separation margin, and σ stands for the width of the kernel function. The RBF Kernel is defined as:

$$K(x_i, x_j) = \exp\left(-\left\|x_i - x_j\right\|^2 / \sigma^2\right) \tag{8}$$

2.3. Particle Swarm Optimization

As shown in (6) and (9), γ and σ are important parameters in LS-SVM model which determine the accuracy and generalization ability of LS-SVM model. However, these two parameters are given arbitrarily in general LS-SVM model. In this section, a new evolutionary computation called particle swarm optimization (PSO) [8] is applied to obtain optimal parameters of LS-SVM. PSO is an evolutionary computation technique based on swarm intelligence. It has many advantages over other heuristic techniques. PSO algorithm can exploit the distributed and parallel computing capabilities, to escape local optima and quick convergence.

In PSO, individuals are called particles and the population is called a swarm. In a n-dimensional complex search space, the ith particle updates its position and speed with the rules as given below:

$$V_{id}^{k+1} = \omega_i V_{id}^{k} + c_1 r_1 (P_{id}^{k} - X_{id}^{k}) + c_2 r_2 (P_{gd}^{k} - X_{id}^{k}) \tag{9}$$

$$X_{id}^{k+1} = X_{id}^{k} + V_{id}^{k+1}, i = 1, 2, \ldots n, d = 1, 2, \ldots D \tag{10}$$

Where $V_i = (V_{i1}, V_{i2}, \ldots V_{iD})^T$ is the speed vector of the ith particle, $X_i = (x_{i1}, x_{i2}, \ldots x_{iD})^T$ is the position vector of the ith particle, ω_i is the inertia weight, c_1 and c_2 are learning parameters, r_1 and r_2 are random value between 0 and 1.

3. Modeling Based on AHP-PSO-LSSVM for Port CSP Demand Forcasting

3.1. Analysis of Port CSP Influential Factors Based on AHP

As the port enterprises locate in coastal areas, the surroundings of the equipment are very bad, and the equipment must undergo the bad environment, such as high-low temperature, vibration and so on. Besides, port takes long-term continuous operation with a higher productivity, and the management level of the equipment maintenance personnel also plays an important part. Therefore, the reasons for CSP replacement are influenced by tasks, equipment, environment, human, accidents and many other complicated factors.

In order to quantify these factors for the model's input, we analyzed the historical data and referred to the opinions of the experts. Besides, the authors have discussed with the experts and the questionnaire based on AHP was distributed to 30 managers and staffs, 28 effective questionnaires were collected. After questionnaire investigation, data analysis and weight calculation according to AHP method, the descriptions and weights of each influential factor are listed as Table 1. In order to eliminate irrelevant noise and derive the forecasting result more accurate, the author will first assume five factors of the front as the input variable. They are equipment working hours, equipment handling volume, failure time, maintenance time and failure rate. In the next experiment section, the real data of four factors, five factors and six factors will be tested to see which dataset has the higher accuracy, and then the number of input variables is determined.

Table 1. The Weights of each Influential Factors

The influential factors	Weights
Equipment working hours	0.1525
Equipment handling volume	0.1311
Failure time	0.1249
Maintenance time	0.1130
Failure rate	0.1022
CSP lifetime	0.0925
The historical requirement at the same month	0.0879
Quantity of CSP in one equipment	0.0757
Environmental factors	0.0720
Accidental factors	0.0482

3.2. The PSO-LSSVM Model for Port CSP

In order to build the LS-SVM model for Port CSP, we make the influential factors proposed above as the input vectors of LS-SVM, and make CSP demand as the output values [9]. The LS-SVM model for port CSP demand is shown in Figure 1.

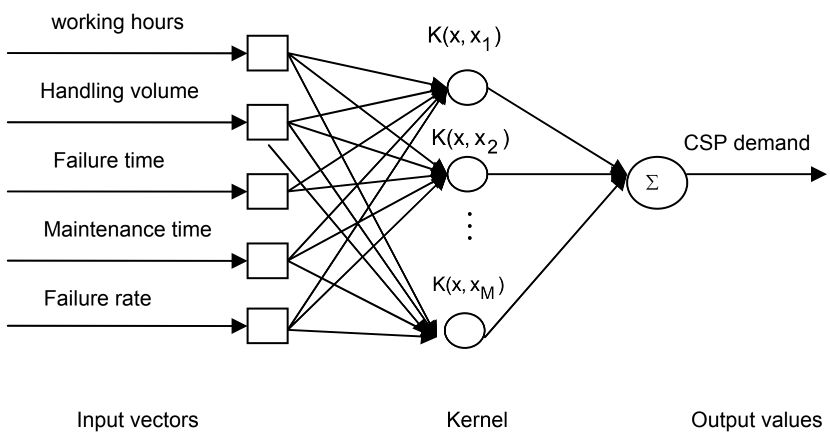

Figure 1. LS-SVM Model for Port CSP Demand

The parameters γ and σ are optimized by PSO [10] with flowchart shown in Figure 2.

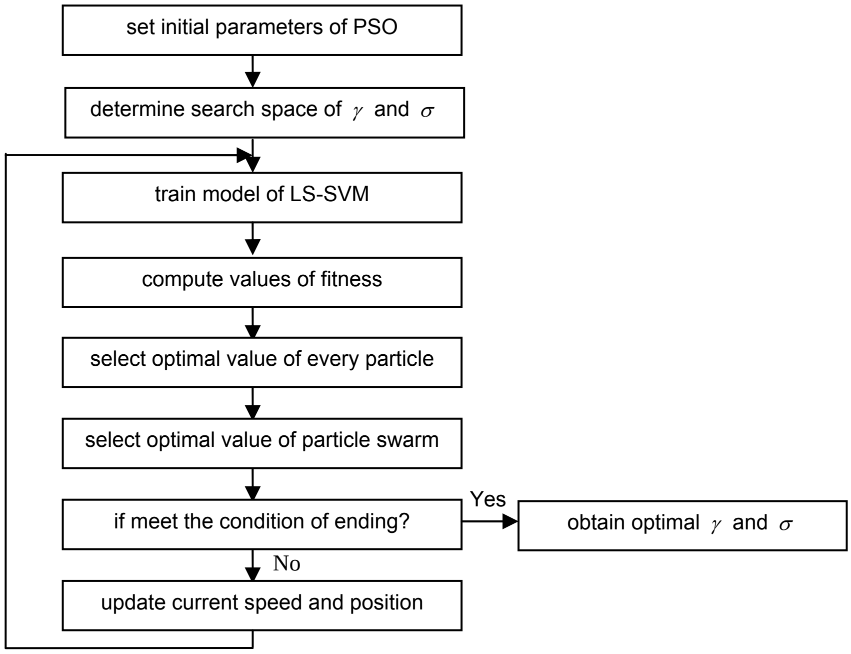

Figure 2. PSO Optimization Flowchart of LS-SVM

4. Experiments and Discussion

4.1. Data Preparation

In this study, the time series data of port CSP consumption along with the five influential factors covered from January, 2008 to December, 2012, which is obtained from Qinhuangdao Port. From the dataset, there are 60 samples as each month is a sample, and the first 80% is used for training while the balance of 20% is for testing. The model is simulated with the LS-SVMlab1.8 toolbox in MATLAB environment. Prior to training, all input and output were normalized using function scaleforSVM. The objective is to independently normalize each feature component to the specified range $[0,1]$, as normalization technique may improve the prediction accuracy and data mining algorithm [11].

4.2. Parameters Optimization

The original values of the parameters in PSO are shown in Table 2. And evaluate the desired optimization fitness function for each particle as the Mean Square Error (MSE) over the data set. Through the optimization steps in Figure 1, the optimal parameters are obtained as $\gamma = 26.8024$, $\sigma = 1.7686$.

Table 2. PSO Parameters

Parameter	value
Swarm size	30
Evolution generations	300
Learning parameters c_1	1.5
Learning parameters c_2	1.7
numeric area of γ	[0.01,1000]
numeric area of σ	[0.01,100]

4.3. Evaluation Metrics

For the purpose of evaluating the proposed technique, two quantitative evaluation metrics are utilized, namely Mean Absolute Percentage Error (MAPE), and Mean Square Error (MSE), which are defined as follows:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_i}\times 100 \quad (11)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \quad (12)$$

Where y_i stands for actual values, $\hat{y}_i$ stands for predicted values, and n stands for the number of samples. It is generally considered that the smaller the value of MAPE and MSE, the better accuracy of the prediction.

4.4. Results and Discussions

As shown in Figure 3, the proposed model PSO-LSSVM can obtain relatively accurate predictions for CSP demand, as 75% of the samples are predicted absolutely correct and the error of the rest is at most 2. By using the same test samples, compared with LS-SVM model using cross-validation optimization method. Computing the MAPE and MSE shown in Table 3, it is obvious that approximate accuracy of PSO-LSSVM is much better than LS-SVM with cross-validation optimization.

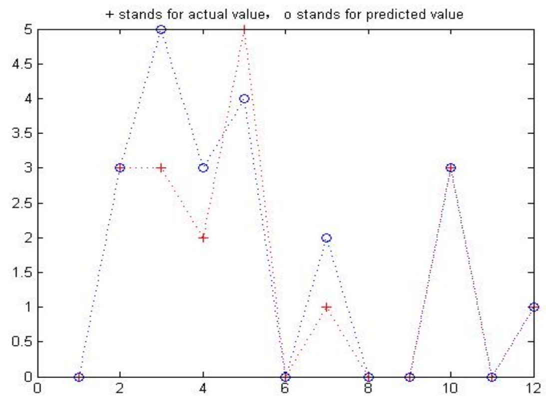

Figure 3. The Comparison Chart of Predicted Value and Actual Value of PSO-LSSVM Model

Table 3. Table of Error Comparison

model	MSE	MAPE
PSO-LSSVM	0.4500	5.2944
LS-SVM	0.8913	8.6571

In the above experiment, five influential factors are taken as the input variables. In order to see if it is the best choice, we will compare with the models of four factors and six factors of the front in the result of AHP method. The comparison charts of predicted value and actual value of four factors model and six factors model are shown in Figure 4 and Figure 5. And the MSE comparison of the three model is shown in Table 4. From these fiures and tables, three prominent findings can be observed.

First, it is obvious that LS-SVM model did a great job in port CSP demand forecasting. And the performance of PSO is better than normal parameter optimization method.

Second, the model of six input variables has the best MSE, but Figure 3 and 5 illustrate that five and six factors models have the same prediction result. It means that the sixth factor CSP lifetime does not play an important role in the model. So, the demand forecasting model still choose the first five factors in the result of AHP method.

Last, but not least. Using the proposed model, we can predict the mean demand of the spare parts next month, and set the range of safety inventory, which will effectively reduce inventory backlog.

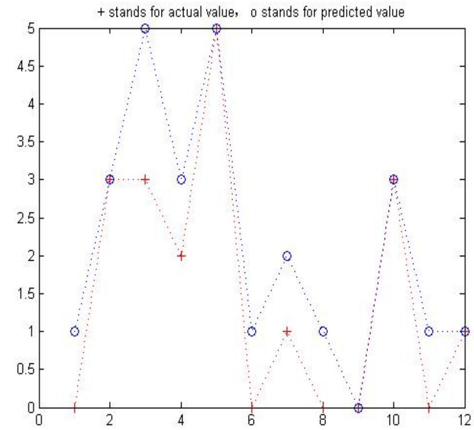

Figure 4. The Comparison Chart of Predicted Value and Actual Value of Four Factors Model

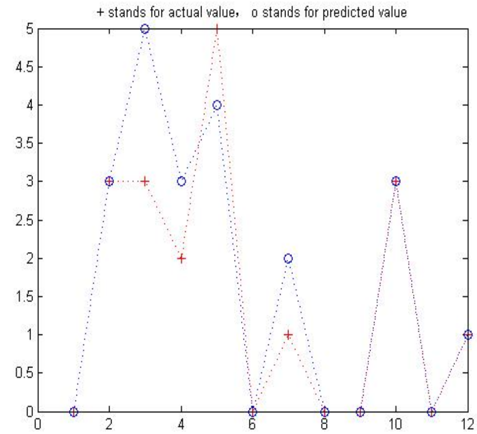

Figure 5. The Comparison Chart of Predicted Value and Actual Value of Six Factors Model

Table 4. Table of MSE Comparison

Model with different number of factors	Four facrors	Five facrors	Six factors
MSE	0.6528	0.4500	0.4405

5. Conclusion

Spare parts management has always been a very important part in factories, especially in port industry. Excessive spare parts will cause backlog of the inventory and insufficiency will cause termination of equipment operation, leading to loss. In this paper, a PSO-LSSVM model for port CSP is proposed. Unlike previous spare parts demand research, some influential factors are included as inputs in the model, which makes the model more comfort to the characteristics of the port CSP. In addition, a new PSO algorithm is introduced to optimize the parameters of LS-SVM model, and the experiments with real data in Qinhuangdao port show that the forecast accuracy of PSO-LSSVM method is better than normal LS-SVM method.

The forecasting model of this paper can be provided as a reference of critical spare parts management in port enterprises to make planning and reduce risks and costs. And to other type of companies which also have spare parts problems, change some of the influential factors in the model may also adjust to their situation. It is noted that the paper gives a new forecasting method for spare parts demand, and the method is closer to the enterprise practice.

Acknowledgements

This work was financially supported by the Educational Commission of Hebei Province of China (ZH2012021) and The Natural Science Foundation of Hebei Province Youth Fund (G2011203195).

References

[1] FeiLong Chen, YunChin Chen, JunYuan Kuo. Applying moving back-propagation neural network and moving fuzzy neuron network to predict the requirement of critical spare parts. *Expert Systems with Applications*. 2010; (37): 4358–4367.

[2] Somnath Mukhopadhyay, Adriano O.Solis, Rafae S Gutierrez. The Accuracy of Non-traditional versus Traditional Methods of Forecasting Lumpy Demand. *Journal of Forecasting.* 2012; (31): 721-735.

[3] Andrea Bacchettin, Nicola Saccani. Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice. *Omega.* 2012; (40): 722–737.

[4] SG Li, X Kuo. The inventory management system for automobile spare parts in a central warehouse. *Expert Systems with Applications.* 2008; (34): 1144–1153.

[5] Hua H, Zhang B. A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts. *Applied Mathematics and Computation.* 2006; (181): 1035–48.

[6] Suykens JAK, Vandewalle J. Least Squares Support Vector Machine Classifiers. Neural Processing Letters. 1999; 9(3): 293-300.
[7] Saaty TL. The analytic hierarchy process. New York: McGraw-Hill. 1980.
[8] Kennedy J, Eberhart R. *Particle swarm optimization.* Proceedings of IEEE International Conference on Neural Networks. Perth: IEEE Press. 1995: 1942-1948.
[9] Peng Bai, Xibin Zhang, Bin Zhang et al. *Support Vector Machine and its Application in Mixed Gas Infrared Spectrum Analysis*. Xi'an Electronic Sience &Technology University Press. 2008: 69-76.
[10] Qianwen Xiang, Yukun Sun, Xiaofu Ji. *Modeling Inductance for Bearingless Switched Reluctance Motor based on PSO-LSSVM*. Chinese Control and Decision Conference (CCDC), 2011: 800-803.
[11] LA Shalabi, Z Shaaban, B Kasasbeh. Data mining: a preprocessing engine. *J Comput Sci.* 2006; 2: 735-739.