International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:8, No:4, 2014

A New Model for Production Forecasting in ERP

S. F. Wong, W. I. Ho, B. Lin, Q. Huang

***Abstract*—**ERP has been used in many enterprises for management, the accuracy of the production forecasting module is vital to the decision making of the enterprise, and the profit is affected directly. Therefore, enhancing the accuracy of the production forecasting module can also increase the efficiency and profitability. To deal with a lot of data, a suitable, reliable and accurate statistics model is necessary. LSSVM and Grey System are two main models to be studied in this paper, and a case study is used to demonstrate how the combination model is effective to the result of forecasting.

***Keywords*—**ERP, Grey System, LSSVM, production forecasting.

I. Introduction

ALONG with the economic growth, the productions of enterprises are tremendously increased. ERP, Enterprise Resource Planning is used to aid with the management of enterprises. ERP is a cross functional enterprise system driven by an integrated suite of software modules that supports the basic internal business processes of a company, and production forecasting is one of the modules of ERP. Therefore, forecasting the production activities is a critical step to the supply chain management which including determining the amount of inventory to be kept on hand, how much raw material should be purchased or how much of a product should be made. Inaccurate forecasting may lead to costly inventory buildup, which is harmful in a business world.

This paper attempts to apply the combination of Grey system and SVM/LSSBM model to predict the production of automotive industry, all those models are introduced in the next section.

II. Statistics Models

As the calculation processes are done by Mat Lab, the calculation formulas of those models are not introduced in detail, but the meanings and applications do.

A. SVM Regression

Support Vector Machines (SVM) are models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. As shown in the Fig. 1, H_1, H_2 and H_3 are used to separate the black points and the white points. H_1 does not separate the classes; H_2 does but only with a small margin; H_3 separates them with the maximum margin, which is the purpose of SVM models. For some cases, as shown in the Fig. 2, nonlinear situations may involve to obtain a better result. For most of the real situations, linear one may not enough to separate the data well and then nonlinear one is chosen, as this study does.

A version of SVM for regression is called Support Vector Regression (SVR). It depends only on a subset of training data, because the cost function for building the model ignores any training data close to the model prediction within a threshold. In the space, the main purpose of SVM and SVR is to finding the hyperplane, and this is also the main different between SVM and SVR. The hyperplane of SVM is the plane that can separate the data; meanwhile the hyperplane of SVR is the plane that can predict the data. If the training data is represented as (x_1, y_1), …(x_i, y_i), x is the attribute and y is the regression value (target class in SVM). Since f(x) = wx + b, if the difference between $f(x_i)$ and y_i is small enough for every instance x_i, then f(x) can well predict y from x and w is the hyperplane of SVR.

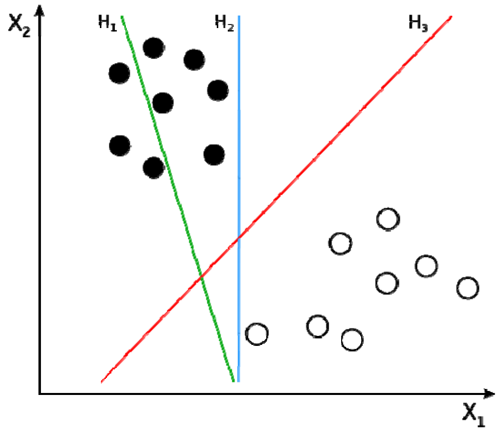

Fig. 1 Separating classes with the maximum margin

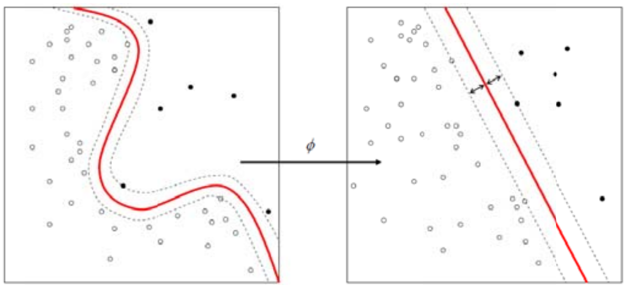

Fig. 2 Nonlinear situation of SVM model

For nonlinear problems, assume sample to be n-dimension vector, then in one certain domain, N samples and the values can be expressed as:

$$(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N) \in R^n \times R \quad (1)$$

S. F. Wong is with the Electromechanical Engineering department, University of Macau, Macau (e-mail: fstsfw@umac.mo).
W. I. Ho is with the Electromechanical Engineering department, University of Macau, Macau (e-mail: edcoho@gmail.com).
B.Lin is with the Electromechanical Engineering Department, University of Macau, Macau (e-mail: ambitionbin@gmail.com).

Firstly, a nonlinear mapping *Ψ()*is used to map samples from former space R^n to feature space

$$\Psi(x) = \left(\Phi(x_1), \Phi(x_2), \cdots \Phi(x_N)\right) \tag{2}$$

Then, in this high-dimension feature space, optimal decision function

$$f(x) = \omega\phi(x) + b \tag{3}$$

is established. In this function, ω is weighed value vector and b is threshold value. In this way, nonlinear prediction function is transformed to linear prediction function in high-dimension feature space [1].

The optimization problem is given [2], [3]

$$\min \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{4}$$
$$\text{s.t.} \begin{cases} y_i - \omega\phi(x_i) - b \le \varepsilon + \xi_i \\ \omega\phi(x_i) + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i \xi_i^* \ge 0 \end{cases}$$

where the ε-insensitive loss function

$$|y - f(x,\omega)|_\varepsilon = \begin{cases} 0, \text{if}|y - f(x,\omega)| \le \varepsilon \\ |y - f(x,\omega)| - \varepsilon, \text{otherwise} \end{cases} \tag{5}$$

where ε is the approximation accuracy that can be violated by means of the slack variables $\xi_i, \xi_i^*$for the non-feasible case. Constant C > 0 determines trade-off between flatness of f and the amount up to which deviations larger than ε are tolerated. A smaller value of C tolerated a larger deviation.

B. LSSVM Regression

Although SVM models scale to high dimensional input spaces very well, the major drawback is higher computational burden for the constrained optimization programming. Thus, another SVM version Least Squares Support Vector Machines (LSSVM) appeared. Compared to SVM, LS-SVM is a reformulation of principle of SVM, which involves equality instead of inequality constraints. Furthermore, LS-SVM uses the least squares loss function instead of the ε-insensitive loss function. Therefore, it is easier to optimize and the computing time is short [4]-[8].

$$\min \quad J(\omega, e) = \frac{1}{2}\|\omega\|^2 + \frac{1}{2}\gamma\sum_{i=1}^{N} e_i^2 \tag{6}$$
$$\text{s.t.} y_i = \omega^T x_i + b + e_i, i = 1, \cdots, N$$

where $e_i \in R$ are error variables; $\gamma \ge 0$ is a regularization constant. Smaller γ can avoid overfitting in case of noisy data [4].

The Lagrangian is given by

$$L_{LS-SVM} = \frac{1}{2}\|\omega\|^2 + \frac{1}{2}\gamma\sum_{i=1}^{N} e_i - \sum_{i=1}^{N}\alpha_i\{\omega^T\phi(x_i) + b + e_i - y_i\} \tag{7}$$

with Lagrange multipliers $\alpha_i \in$. The conditions for optimality are given by

$$\begin{cases} \frac{\partial L_{LS-SVM}}{\partial b} = 0 \to \omega = \sum_{i=1}^{N}\alpha\phi(x_i) \\ \frac{\partial L_{LS-SVM}}{\partial b} = 0 \to \sum_{i=1}^{N}\alpha_i = 0 \\ \frac{\partial L_{LS-SVM}}{\partial \varepsilon_i} = 0 \to \alpha_i = \gamma\varepsilon_i, \ (k = 1, \cdots, N) \\ \frac{\partial L_{LS-SVM}}{\partial \alpha_i} = 0 \to \omega^T\phi(x_i) + b + e_i + y_i = 0 \end{cases} \tag{8}$$

After elimination of ω, e one obtains the following linear equations

$$\begin{bmatrix} 0 & 1_V^T \\ 1_V & \Omega + \frac{1}{\Upsilon} \end{bmatrix}\begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{9}$$

where $y = [y_1; \cdots; y_N], 1_v = [1; \cdots; 1], \alpha = [\alpha_1; \cdots; \alpha_N]$ and the Mercer condition has been applied again

$$\Omega, = \langle\Phi(x_i)\Phi(x_l)\rangle = K(x_i, x_l) \quad i, l = 1, \cdots, N \tag{10}$$

The resulting LSSVM model for function estimation becomes

$$f(x) = \sum_{i=1}^{N}\alpha_i K(x, x_i) + b \tag{11}$$

C. Grey System

The key of grey system theory is to generate function and grey differential equation from grey data which is not clear. The concepts of grey block and differential similarity are natures to establish grey forecasting model, which is described as GM(N,M), where N is the rank of differential equation, and M is the number of variables. The tradition of grey system, GM(1,1) is demonstrated as follow [9].

Given the original dataset

$$x^{(0)}(k) = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\} \tag{12}$$
$$x^{(0)}(k) \ge 0, k = 1,2, \cdots n$$

For increasing sequence smoothness, to $x^{(0)}(k)$do 1-AGO:

$$x^{(1)}(k) = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\} \tag{13}$$
$$x^{(1)}(k) = \sum_{n=1}^{k} x^{(0)}(n), k = 1,2,3 \cdots n$$

Then on $x^{(1)}$set below differential equation:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \tag{14}$$

Solution of differential equation:

$$\mathrm{x}^{(1)}(\mathrm{t}) = \left(\mathrm{x}^{(0)}(1) - \frac{\mathrm{b}}{\mathrm{a}}\right)\mathrm{e}^{-\mathrm{ak}} + \frac{\mathrm{b}}{\mathrm{a}} \tag{15}$$

The model parameters can be determined by the least square regression.

$$\hat{\mathrm{a}} = (\mathrm{a}, \mathrm{b})^{\mathrm{T}} = (\mathrm{B}^{\mathrm{T}}\mathrm{B})^{-1}\mathrm{B}^{\mathrm{T}}\mathrm{Y}_{\mathrm{N}} \tag{16}$$

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:8, No:4, 2014

where

$$Y_N = \left[x^{(0)}(2), x^{(0)}(3), x^{(0)}(4), \cdots, x^{(0)}(n)\right]^T \tag{17}$$

$$B = \begin{bmatrix} -\frac{1}{2}\left[x^{(1)}(1) + x^{(1)}(2)\right] & 1 \\ -\frac{1}{2}\left[x^{(1)}(2) + x^{(1)}(3)\right] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k+1)\right] & 1 \end{bmatrix}$$

The original sequence of prediction:

$$\hat{x}^{(0)}(k) = \left[\left(\hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)\right)\right]$$
$$= (1 - e^{a})\left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} \tag{18}$$

III. Methodology

The SVM and LSSVM are both mainly aimed at solving the large scale samples and multiple-dimension series to get a high accuracy. The only difference between SVM and LSSVM is that, the LSSVM solves linear equations instead of quadratic programming problem which lead to high computational efficiency. After pruning both sparseness and performance of LSSVM are comparable with those of SVM. However, the Grey System targets the single time series, and can perform even in several samples, but sometimes the results are not satisfied very well. Thus, to combine the advantages of these methods to get a more satisfied accuracy is the main purpose of this study, so a combination model of SVM/LSSVM and Grey System is established and tested.

Using the data of automotive industry which is collected from USA (as Fig. 3), Germany, China, Japan, Korea, and India (from 2004 to 2012), to build three models: Grey System Model, LSSVM Model and Grey-LSSVM Model.

2004

Country	Type	Jan.	Feb.	Mar.	Apr.	May	Jun.	First Half (Jan.-Jun.)
USA	Cars	476,946	559,156	657,654	609,680	703,624	634,956	3,642,016
	Light Trucks	625,588	693,280	811,802	774,027	884,426	776,129	4,565,252
	Medium Trucks	15,872	17,172	20,834	19,840	19,438	19,922	113,078
	Heavy Trucks	11,275	12,035	16,767	16,416	16,161	17,192	89,846
	Total	1,129,681	1,281,643	1,507,057	1,419,963	1,623,649	1,448,199	8,410,192

Country	Type	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Second Half (Jul.-Dec.)	Adjusted value	Total
USA	Cars	648,061	607,825	579,894	556,185	494,830	624,749	3,511,544	336,514	7,490,074
	Light Trucks	869,416	786,755	824,758	745,446	675,587	882,238	4,784,200	39,222	9,388,674
	Medium Trucks	18,895	18,985	17,821	18,743	17,239	23,695	115,378	0	228,456
	Heavy Trucks	17,338	18,233	18,642	19,523	17,793	21,822	113,351	0	203,197
	Total	1,553,710	1,431,798	1,441,115	1,339,897	1,205,449	1,552,504	8,524,473	375,736	17,310,401

Fig. 3 Automotive productivity of USA in 2004

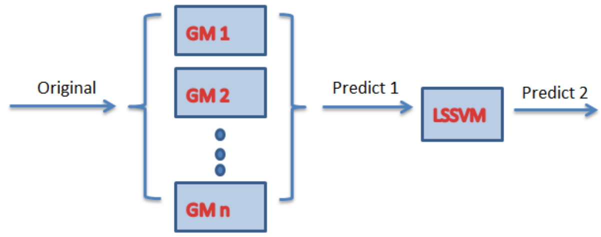

Fig. 4 Series connection

Fig. 5 Parallel connection

Fig. 6 Residual connection

To combine the Grey System and the LSSVM models, there are three connection methods: Series (as Fig. 4), Parallel (as Fig. 5), and Residual (as Fig. 6).

Series connection, using the different single dimension GM models to build a multiple dimensions LSSVM model. Since the GM targets the single dimension series, it can perform even in several samples (4 at least), and LSSVM is good at dealing the multiple dimensions series and large samples.

Parallel connection, applying the Grey System and LSSVM models at the same time.

Residual connection, applying the Grey System first, and then using the output to be the input of the LSSVM model next.

Finally, some indexes are used to measure the performances of the results and are listed as below:

A. The Squared Correlation Coefficient

$$R^2 = \frac{(n\sum x_i y_i - \sum x_i \sum y_i)^2}{[n\sum x_i^2 - (\sum x_i)^2][n\sum y_i^2 - (\sum y_i)^2]} \tag{19}$$

R^2 is a measure of the linear correlation (dependence) between two variables X and Y, giving a value between +1 and -1 inclusive. It is widely used in the sciences as a measure of the strength of linear dependence between two variables. R is used to determine whether the relationship is positive or negative based upon the sign of R.

R^2<30% are considered to have no correlation and behavior is explained by chance.

R^2 of 30% to 49.99% are considered to be a mild relationship.

R^2 of 50% to 69.99% are considered to be a moderate relationship.

R^2 of 70% to 100% are considered to be a strong relationship.

B. Root Mean Squared Error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_i)^2} \tag{20}$$

C. Mean Absolute Error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i| \tag{21}$$

D. Average Absolute Relative Error

$$AARE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - x_i}{y_i}\right| \tag{22}$$

where n is the number of data, $x_i$and$\bar{x}_i$ are actual data and the mean of the actual data respectively, and y_i is the predict results.

IV. Results

Using the MatLab, LSSVM and Grey System are the built in models and can be used direcly. After all the data is input into the systems, the results are obtained as below. Figs. 7 to 10 are the results of LSSVM model and Figs. 11 to 14 are the results of Grey System model. After that, the results of Grey System model is input to the LSSVM model and the results of the Grey-LSSVM model are obtained. Finally, three groups of results are compared in Table I.

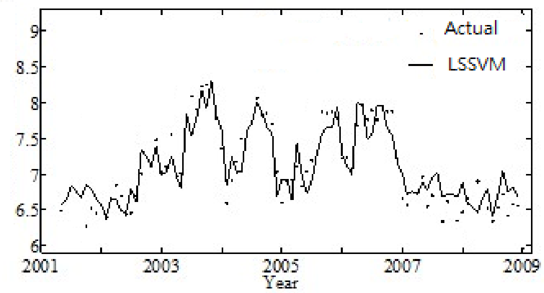

Fig. 7 Actual and predictive values of the training data set in LSSVM model

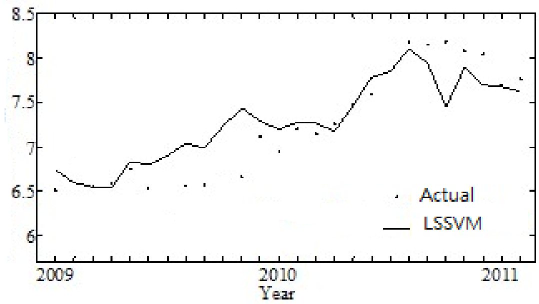

Fig. 8 Actual and predictive values of the testing data set in LSSVM model

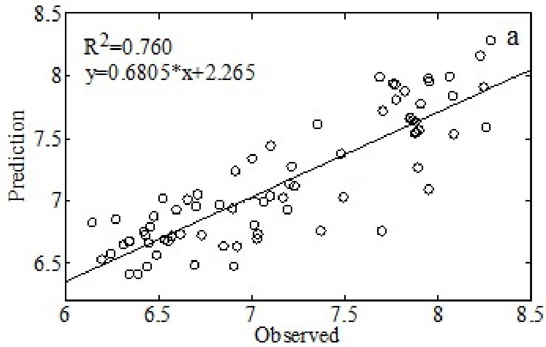

Fig. 9 LSSVM result of the training data set

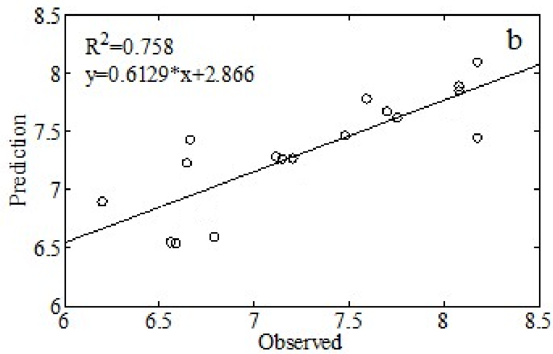

Fig. 10 LSSVM result of the testing data set

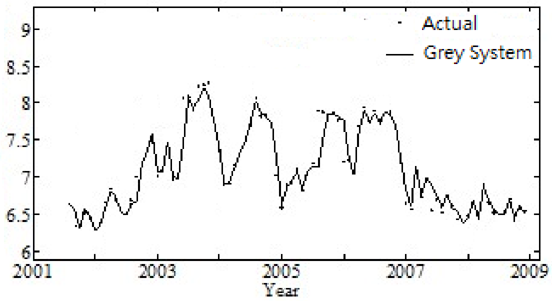

Fig. 11 Actual and predictive values of the training data set in Grey System

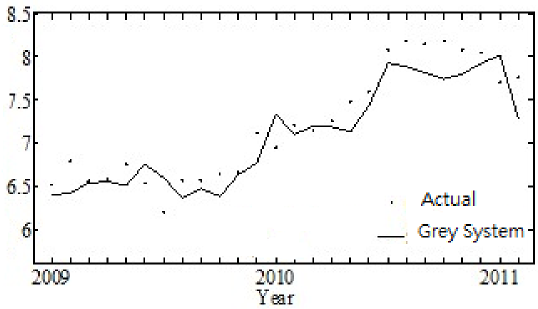

Fig. 12 Actual and predictive values of the testing data set in Grey System

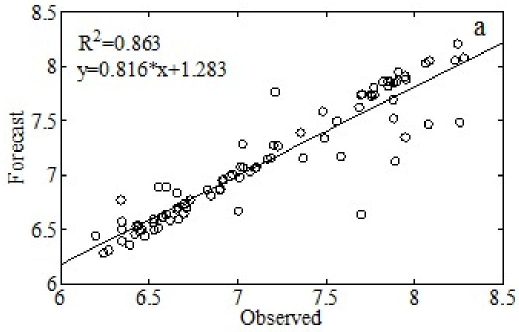

Fig. 13 Grey System result of the training data set

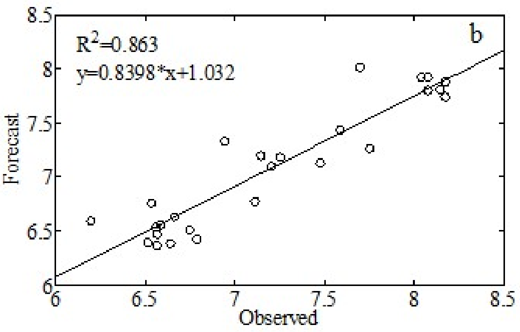

Fig. 14 Grey System result of the testing data set

V. Discussion

In both of the LSSVM and Grey System models, the input data is also separated into two parts: training data set and test data set. The separation is randomly, and then the training data set will be processed by the models. As the forecasting results are obtained, then the results will be compared with the test data set and checked the accuracy.

If the R^2 value is more close to 1, then the forecasting is more accurate; and the smaller the RMSE, MAE and AARE values are, the better accuracy is achieved.

A. LSSVM

For the LSSVM model, one prominent point of this statistic model is that it is suitable for the forecasting of any categories of data. As the data which is used to calculate the parameters in the model are selected randomly, the values of the parameters are changed every time. As a result, the accuracy will not be so high, but also not so low for any situation. For ensuring to have a better result, a “choosing loop” is added into the original program. Since the parameters and the results are changed every time, once one calculation is finished and then the second time of the calculation is started again until the calculation is done for 100 times, finally the best result will be chosen and the loop is finished. In the case study, the R^2 values of LSSVM are 0.760 and 0.758; both of the values are larger than 0.7, which means the prediction accuracy is acceptable.

B. Grey System

For the Grey System model, according to the mathematical process of this model, the result of the forecasting will be changed if the arrangement of the input data is changed. And also the data processing methods before inputting the data into the model affect the results directly, as the values of the data are quite large, all the values are extracted a root for obtaining the smaller values in the case study. Normally, the result is more stable as the values of the parameters will be changed only if the arrangement of the input data is changed, but this model may not suitable for every category of data. Although it is not suitable for all situations, the accuracy of this model is quite high and stable. In the case study, the R^2 values of Grey System are 0.863, which is also larger than 0.7 and the accuracy is even better than the LSSVM model.

C. Grey-LSSVM Model

As the advantage of LSSVM model is wide suitability and Grey System is high accuracy, if a model can combine those of the advantages, then the forecasting model will be more reliable. From the Table I, the R^2 values of the combination model that is purposed in this study are 0.981 and 0.965, which are very close to 1 and it indicates that the accuracy is very high.

On the other hand, according to the RMSE, MAE and AARE values, the combination model also has the best performance, and then the next one is Grey System.

VI. Conclusion

From the comparison of the results that applying three different methods to the single time series of the production forecasting, the combination model of Grey System and LSSVM has the best performance. From another perspective of the R^2, the results of Grey System have the enough accuracy, as the difference between the combination model and the single Grey System model is not significant. It proves that the Grey System model has the more powerful forecasting ability for the small scale samples than single SVM or LSSVM.

TABLE I
Performance Indexes of Three Prediction Models

Performance index	Accuracy performance (Training set) LSSVM	Grey System	Combination
R^2	0.760	0.863	0.981
RMSE	0.307	0.229	0.147
MAE	0.238	0.127	0.097
AARE	0.040	0.018	0.018
Performance index	**Generalization performance (Testing set) LSSVM**	**Grey System**	**Combination**
R^2	0.758	0.863	0.965
RMSE	0.314	0.264	0.151
MAE	0.256	0.226	0.104
AARE	0.047	0.032	0.020

Acknowledgment

The authors acknowledge the funding support by University of Macau under research grant number MYRG082 (Y1-L2)-FST12-WSF.

REFERENCES

[1] Z.W. LIU, X.Y. WANG, "Research on Water Bloom Prediction Based on Least Squares Support Vector Machine", WRI World Congress on Computer Science and Information Engineering, Vol. 5, pp. 764-768, 2009.
[2] V. Vapnik, "The nature of statistical learning theory", New York: Springer-Verlag, 1995.
[3] A.J. Smola, B. Scholkopf, "A Tutorial on Support Vector Regression", Statistics and Computing, Vol. 14, No. 3, pp. 199-222, 2003.
[4] H.F. Wang, E.J. Hu, "Comparison of SVM and LS-SVM for regression", International Conference on Neural Networks and Brain (ICNN&B), Vol. 1, pp.279-283, 2005.
[5] J.A.K. Suykens, "Nonlinear Modeling and Support Vector Machines", IEEE Instrumentation and Measurement Technology Conference, Budapest, 2001.
[6] J.A.K. Suykens, L. Lukas, J. Vandewalle, "Sparse approximation using least squares support vector machines", IEEE International Symposium on Circuits and Systems, Vol. 2, pp. 757-760, 2000.
[7] J.A.K. Suykens, J. D. Brabanter, L. Lukas, J. Vandewalle, "Weighted least squares support vector machines: robustness and sparse approximation", Neurocomputing, Vol. 48, pp. 85-105, 2002.
[8] B. Ustun, "A comparison of support vector machines and partial least squares regression on spectral data", Magisterial dissertation, University of Nijmegen, 2003.
[9] L.F. Wu, S.F. Liu, L.G. Yao, S.L. Yan, D.L. Liu, "Grey system model with the fractional order accumulation", Communications in Nonlinear Science and Numerical Simulation, Vol. 18, no. 7, pp. 1775-1785, 2013.

S. F. Wong, is an assistant professor of Electromechanical Engineering department in University of Macau, in charge of the Industrial Engineering Laboratory and the major research areas are knowledge management, RFID and human factors engineering.

She is a PHD student and teaching assistant of Electromechanical Engineering department in University of Macau, working at Industrial Engineering Laboratory and the major research area is Radio Frequency IDentification (RFID).

He is a master student of Electromechanical Engineering department in University of Macau and working at Industrial Engineering Laboratory.

She is a master student of Electromechanical Engineering department in University of Macau and working at Industrial Engineering Laboratory.