VOL. 16, NO. 1, 2017, 1-5
www.fke.utm.my/elektrika
ISSN 0128-4428

ELEKTRIKA
Journal of Electrical Engineering

An Improved Building Load Forecasting Method using a combined Least Square Support Vector Machine and modified Artificial Bee Colony

Mohammad Azhar Mat Daut[1,2], Mohammad Yusri Hassan[1,2*], Hayati Abdullah[1,3], Hasimah Abdul Rahman[1,2], Md Pauzi Abdullah[1,2] and **Faridah Hussin[1,2]**

[1] Centre of Electrical Energy Systems (CEES), Institute of Future Energy, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Johor, Malaysia.
[2]Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Johor, Malaysia.
[3]Faculty of Mechanical Engineering, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Johor, Malaysia

[*]Corresponding author: yusrih@fke.utm.my, Tel: 607-5557001, Fax: 607-5557005

Abstract: This paper presents an improved building load forecasting method using a combined Least Square Support Vector Machine and modified Artificial Bee Colony. The main contribution of the proposed method is the improvement in the exploitation capability of the standard Artificial Bee Colony, in which a different probability selection has been introduced. This was achieved by changing the standard probability selection with the clonal selection algorithm. The results from two other methods were compared with the results from the proposed method to validate the performance of the proposed forecasting method. The accuracy of the proposed method was evaluated using the Mean Absolute Error, Mean Absolute Percentage Error and Root Mean Square Error. It was found that the proposed method had improved the accuracy by more than 50 % compared to the other methods. The results of the study showed that the proposed method has great potential to be used as an accurate forecasting method.

Keywords: load forecasting, least square support vector machine, modified artificial bee colony.

1. INTRODUCTION

The increasing demand for electricity and the need for saving energy have resulted in much interest for research in the area of load forecasting. There are several types of method being used for load forecasting as described in the literature [1].However, the most often method used by current researchers are methods involving artificial intelligence, such as Artificial Neural Network (ANN) [2] and Support Vector Machine (SVM) [3].

The ANN method has become popular due to its capability to deal with nonlinear time series. However, this method faced the over-fitting problem that may affect the performance in terms of accuracy [4]. In order to improve the accuracy in forecasting, the SVM method has been used by a number of researchers. Extended methods based on the SVM have also been used such as the Least Square Support Vector Machine (LSSVM) for nonlinear time series forecasting resulting in better performance [5, 6].

Artificial Bee Colony (ABC) [7] is a recently introduced method under Swarm Intelligence which was inspired from the intelligence of honeybees. This ABC algorithm has been shown to be applicable in solving problems in various forecasting fields [8, 9]. The ABC algorithm has the advantage of being able to conduct both global search and local search in each iteration [10]. In addition, ABC has less parameter to control and uses simple mathematical equations [8] .Although ABC performed better than other optimization techniques, the method can be further improved. The ABC is good for exploration but lacks in exploitation [11]. Thus, some improvement has been made to the exploitation process in this study by introducing the different probability function [11].

In this study, a combination of the LSSVM and an improved ABC method has been proposed in order to improve the load forecasting performance. The LSSVM will train the actual data while the improved ABC method will carry out the global search and local search for the best results to be used in forecasting. The forecasted data is then compared with actual data as well as results from other methods.

This paper is organized as follows. In sections 2 and 3, the theory of LSSVM and ABC will be presented. This is followed by section 4 in which the improvement of ABC will be presented. The methodology of this experiment will be explained in section 5. The empirical results will be discussed in section 6 and followed by the conclusions in section 7.

2. LEAST SQUARE SUPPORT VECTOR MACHINE

In 1995, Vapnik [12] proposed a new machine learning algorithm called the Support Vector Machine (SVM). The SVM is based on the statistical learning theory (SLT) [13]. The SVM is structured in such a way that it combines the advantages of ANN, nonlinear regression and the sparseness of the solution [5, 14, 15].In order to solve the linear equation in an easier manner, Suykens, et al. [16] proposed an upgraded version of SVM known as the Least Square Support Vector Machine (LSSVM) by maintaining the advantages of SVM. The formulation of LSSVM is briefly explained as follows.

Consider a set of training data, $\{x_i, y_i\}^N$ where represents the input value and represents the output value. According to the LSSVM theory, the unknown nonlinear function can be estimated by;

$$y(x) = w^T \varphi(x^i) + b + e_i \tag{1}$$

where $\varphi(x_i)$ is a nonlinear function, w^T is the weight, b is a bias and e_i is the error between the actual and predicted outputs. The input value x_i and output $y(x)$ are explained in section 4. By using the equation below, the coefficient vector of and can be obtained [16].

$$\min_{w,b,e} J(w,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{i=1}^{N} e_i^2 \tag{2}$$

Subject to the equality constraints

$$y_i = w^T \varphi(x_i) + b + e_i, i = 1,2,....,N$$

By applying the Langrage multiplier into (2):

$$\begin{aligned} L(w,b,e;\alpha) &= J(w,e) \\ &- \sum_{i=1}^{N} \alpha_i \left\{ w^T \varphi(x_i) + b + e_i - y_i \right\} \end{aligned} \tag{3}$$

where $\alpha_i (i = 1,...,n)$ is Langrage multiplier and γ is the penalty parameter which balances the complexity of the LSSVM model. Using the Karush-Kuhn-Tucker (KKT) conditions, the following equations can be obtained.

$$\begin{aligned} &\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{N} \alpha_i \varphi(x_i) \\ &\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{N} \alpha_i = 0 \\ &\frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \lambda e_i \\ &\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow w^T \varphi(x_i) + b + e_i - y_{i=0} \\ &i = 1,2,..., N \end{aligned} \tag{4}$$

The above equations can be simplified to the following set of linear equations by eliminating the w and e_i :

$$\begin{bmatrix} \alpha \\ b \end{bmatrix} \begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + \dfrac{I}{\gamma} \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{5}$$

And equation (1) will become:

$$y(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i) + b \tag{6}$$

where α and b are from equation (5) and $K(x, x_i)$ is defined as the kernel function. There are two types of typical kernel namely the polynomial kernel and radius basis function (RBF) kernel:

$$\text{Polynomial: } K(x, x_i) = (x_i^T x + 1)^d \; (d = 1,..., n)$$

$$\text{RBF: } K(x, x_i) = \exp(-\|x - x_i\|^2 / 2\sigma^2$$

3. INCORPORATION OF ARTIFICIAL BEE COLONY ALGORITHM

Inspired from the behaviour of honeybees searching for food, the Artificial Bee Colony (ABC) algorithm was developed by Karaboga and Basturk [12]. In the ABC algorithm, the main objective of the artificial bees is to find the largest amount of nectar (fitness) by finding the position of the food sources (solutions).

The ABC can be categorized into three groups, namely, employed bees (EB), onlooker bees (OB) and scout bees (SB). Each group of bees has a different task to do. The employed bees will search for the food sources position in their memory and will transfer the information to the onlooker bees group. In the dance area, the onlooker bees will make a decision to select the best food sources provided by the employed bees. The scout bees group consists of a few employed bees, which left their food source to seek new ones. The main steps of the ABC algorithm can be explained as follows.

3.1 Initialization Level

In this stage, equation (7) will randomly generate the population of food sources within the range of the variable boundaries.

$$x_{ij} = x_j^{\min} + rand(0,1)\left(x_j^{\max} - x_j^{\min}\right) \tag{7}$$

Where i represents the number of food source and j represents the variable associated with the food source. Meanwhile, $x_j^{\min}$ and $x_j^{\max}$ will represent the lower and upper boundaries of the parameter of interest respectively. Then, the fitness of the food source will be calculated after the initialization of the population is defined as follows [7]:

$$fit_i = \frac{1}{(1 + obj.Fun_i)} \tag{8}$$

where $obj.Fun$ is the objective function.

3.2 Employed Bees Level

The number of food source (SN) is equal to the number of colony bee. EB will be assigned for each of the food source's position. New food sources are obtained according to equation (9) below:

$$v_{ij} = x_{ij} + \varphi_{ij}(x_{ij} - x_{kj}) \tag{9}$$

where $i \in \{1,2,...,SN\}$, $j \in \{1,2,...,n\}$, φ_{ij} is a randomly generalized real number within the range and x is a randomly selected index number in the colony and k has a different value from i

After the value, v_{ij} is obtained, it will be evaluated and compared to the previous, x_{ij} . This comparison will determine whether the new solution is better than the previous solution. The bees will memorize the new solution if the new solution is better; otherwise it will memorize the previous solution.

3.3 Onlooker Bees Level

In this phase, the OB will execute the information taken from EB. The OB will select the food source which relies on the probability value associated with the food source, P_i . This P_i will be calculated using the following equation:

$$P_i = \frac{fit_i}{\sum_{j=1}^{SN} fit_j} \tag{10}$$

where fit_i is the fitness value of the solution i . As in the EB level, it will calculate a new solution from its food source. The EB and OB use the same method to evaluate the food source. However, in the EB level, each solution will be updated but in the OB level, only the selected one will go through this step.

3.4 Scout Bees Level

At this level, if the food source, cannot be improved for a given number of trials (limit), it is abandoned and the corresponding EB will transform to SB. The scout produces a food source randomly as follows:

$$x_{id} = x_d^{\min} + rand(0,1)(x_d^{\max} - x_d^{\min}) \tag{11}$$

where $d = 1,2,...,n.$

3.5 Modified ABC

A modification of the ABC algorithm is explained in this section to enhance the process of the ABC algorithm. The fitness of the food source has been chosen by OB. By using equation (10), the probability value of the food source can be determined. This probability value is being used to find the exploitation rate. As highlighted before, some modification has been done in order to improve the exploitation mechanism of OB. Thus, a new probability calculation has been introduced [17]:

$$P_i = \exp(-1/\rho * fit) \tag{12}$$

where ρ represent a new control parameter in the ABC algorithm and fit is the fitness value. This modified ABC algorithm has been called the ABCclo [11].

4. METHODOLOGY

Description of the data, evaluation criteria and parameter setting for building load forecasting will be discussed in this section.

4.1. Data Description

In this study, there are four sets of data used and tested which includes the dry bulb, dew point, holidays and actual hourly load. The hourly load data used are from January 1, 2004 to December 31, 2008. From the data set, 60% is being used in the training set while the remaining is used for testing.

4.2. Evaluation Criteria

The performance of this proposed method has been evaluated by using three different types of evaluation criteria namely the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). All these parameters are important in determining the forecasting capability of the model. They are expressed as follows:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|F_t - A_t\right| \tag{13}$$

$$MAPE = \frac{1}{N}\left[\sum_{t=1}^{N}\left|\frac{A_t - F_t}{A_t}\right|\right]x100 \tag{14}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(A_t - F_t\right)^2} \tag{15}$$

4.3. Parameter Setting

In order to evaluate the performance of the proposed method, the parameter setting of ABC needs to be established. The colony size, dimension (D) and limit of the ABC algorithm are summarized in Table 1 below.

Table 1. ABC Parameter Setting

ABC Parameter	
Colony Size	4
SN	2
D	6000
Parameter Range	[-1,100]
Limit	10

5. EMPIRICAL RESULT

As discussed earlier, the computational intelligence-based modelling approach had been used in this study. The proposed method has been compared with two other methods; standard ABC-LSSVM and LSSVM to evaluate the accuracy of the proposed method. All the performance results of ABCclo-LSSVM, standard ABC-LSSVM and LSSVM are shown in Table 2. The best results are shown in Table 2 for 15 independent runs.

Table 2 shows that the proposed method has a better performance than the standard ABC-LSSVM and LSSVM for most ρ values. But for some ρ values, it shows that the ABCclo-LSSVM did not perform as well as the other methods. This is due to the selection of the probability value problem encountered in ABCclo-LSSVM when the value is 1000 and 1250 which is indicated by the small value of the probability that affects the exploitation process. This finally led the algorithm to forecast a bit further from the actual data.

Table 2. Performance Results

Algorithm	**MAE (kWh)**	**MAPE (%)**	**RMSE**
LSSVM	193.995	1.047	11.940
Standard ABC-LSSVM	146.382	0.79	9.01
ABCclo -LSSVM $\rho = 250$	43.811	0.24	2.696
ABCclo -LSSVM $\rho = 500$	56.020	0.30	3.448
ABCclo -LSSVM $\rho = 750$	8.663	0.05	0.533
ABCclo -LSSVM $\rho = 1000$	200.753	1.08	12.369
ABCclo -LSSVM $\rho = 1250$	238.970	1.31	14.708
ABCclo -LSSVM $\rho = 1500$	129.066	0.70	7.943

The actual and forecasted data from the three methods are shown in Figure 1. It can be seen that the ABCclo-LSSVM (ρ =750) resulted in the smallest difference between the actual and forecasted data and thus, improved the accuracy of the forecasted load. The other ρ values (250, 500, 1250 and 1500) also performed better than the standard ABC-LSSVM and LSSVM model. The numerical results for the actual and forecasted data are given in Table 3.

Table 3. Numerical results for actual and forecasted data

Method	Actual Data (MW)	Forecasted Data (MW)
LSSVM	18.435	18.242
ABC-LSSVM	18.435	18.581
ABCclo -LSSVM $\rho = 250$	18.435	18.479
ABCclo -LSSVM $\rho = 500$	18.435	18.491
ABCclo -LSSVM $\rho = 750$	18.435	18.426
ABCclo -LSSVM $\rho = 1000$	18.435	18.636
ABCclo -LSSVM $\rho = 1250$	18.435	18.196
ABCclo -LSSVM $\rho = 1500$	18.435	18.564

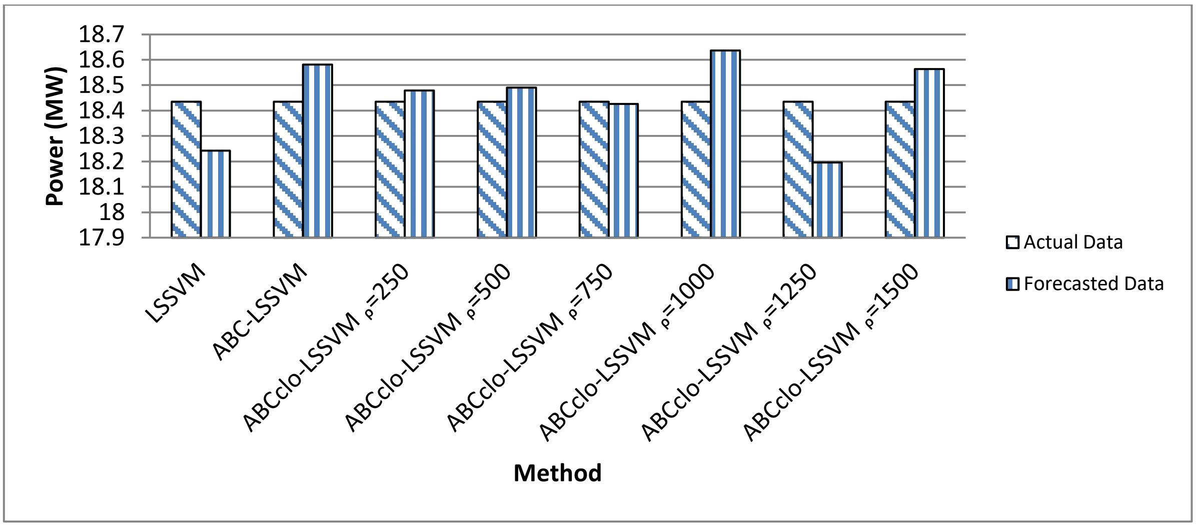

Figure 1. Comparison of actual and forecasted data

7. CONCLUSION

In this paper, a new method which is the ABCclo-LSSVM method has been proposed for building load forecasting. This new method uses the new probability function to improve accuracy performance of building load forecasting. The ABCclo-LSSVM method was tested with six different values and the results were compared with two other methods. The results showed that the proposed method has better performance than the standard ABC-LSSVM and LSSVM methods. For future work, this ABCclo – LSSVM will be applied to other methods to enhance the forecasting accuracy in terms of the search mechanism.

ACKNOWLEDGMENT

Use the singular heading even if you have many acknowledgments. This work was supported by the Ministry of Education Malaysia and Universiti Teknologi Malaysia through the Research University Grant (GUP) vot 07H57.

REFERENCES

[1] S. M. Islam, S. M. Al-Alawi, and K. A. Ellithy, "Forecasting monthly electric load and energy for a fast growing utility using an artificial neural network," Electric Power Systems Research, vol. 34, pp. 1-9, 7// 1995.

[2] I. S. Markham and T. R. Rakes, "The effect of sample size and variability of data on the comparative performance of artificial neural networks and regression," Computers & Operations Research, vol. 25, pp. 251-263, 4// 1998.

[3] B. Dong, C. Cao, and S. E. Lee, "Applying support vector machines to predict building energy consumption in tropical region," Energy and Buildings, vol. 37, pp. 545-553, 5// 2005.

[4] A. Foucquier, S. Robert, F. Suard, L. Stéphan, and A. Jay, "State of the art in building modelling and energy performances prediction: A review," Renewable and Sustainable Energy Reviews, vol. 23, pp. 272-288, 7// 2013.

[5] C.-M. Vong, P.-K. Wong, and Y.-P. Li, "Prediction of automotive engine power and torque using least squares support vector machines and Bayesian inference," Engineering Applications of Artificial Intelligence, vol. 19, pp. 277-287, 4// 2006.

[6] X. Youshen, H. Leung, and H. Chan, "A prediction fusion method for reconstructing spatial temporal dynamics using support vector machines," Circuits and Systems II: Express Briefs, IEEE Transactions on, vol. 53, pp. 62-66, 2006.

[7] D. Karaboga, "An idea based on Honey Bee Swarm for Numerical Optimization," 2005.

[8] Z. Mustaffa, Y. Yusof, and S. S. Kamaruddin, "Application of LSSVM by ABC in energy commodity price forecasting," in Power Engineering and Optimization Conference (PEOCO), 2014 IEEE 8th International, 2014, pp. 94-98.

[9] Osman Hegazy, Omar S. Soliman, and M. A. Salam, "LSSVM-ABC Algorithm for Stock Price prediction," International Journal of Computer Trends and Technology (IJCTT)2014., p. 12, 2014.

[10] W.-C. Hong, "Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm," Energy, vol. 36, pp. 5568-5578, 9// 2011.

[11] B. Babayigit and R. Ozdemir, "A modified artificial bee colony algorithm for numerical function optimization," in Computers and Communications (ISCC), 2012 IEEE Symposium on, 2012, pp. 000245-000249.

[12] V. N. Vapnik, The nature of statistical learning theory: Springer-Verlag New York, Inc., 1995.

[13] C. Cortes and V. Vapnik, "Support-Vector Networks," Machine Learning, vol. 20, pp. 273-297, 1995/09/01 1995.

[14] H. Fazeli, R. Soleimani, M.-A. Ahmadi, R. Badrnezhad, and A. H. Mohammadi, "Experimental Study and Modeling of Ultrafiltration of Refinery Effluents Using a Hybrid Intelligent Approach," Energy & Fuels, vol. 27, pp. 3523-3537, 2013/06/20 2013.

[15] M. A. Ahmadi, M. Ebadi, and S. M. Hosseini, "Prediction breakthrough time of water coning in the fractured reservoirs by implementing low parameter support vector machine approach," Fuel, vol. 117, Part A, pp. 579-589, 1/30/ 2014.

[16] J. A. K. Suykens, T. V. Gestel, J. D. Brabanter, B. D. Moor, and J. Vandewalle, LS-SVMlab Toolbox User's Guide, 2011.

[17] L. N. de Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection principle," Evolutionary Computation, IEEE Transactions on, vol. 6, pp. 239-251, 2002.