A Novel Hybrid Metaheuristic Algorithm for Short Term Load Forecasting

Zuriani Mustaffa

Faculty of Computer Systems & Software Engineering
Universiti Malaysia Pahang
Gambang, Pahang, Malaysia
zuriani@ump.edu.my

Mohd Herwan Sulaiman

Faculty of Electrical and Electronics Engineering
Universiti Malaysia Pahang
Pekan, Pahang, Malaysia
mherwan@ieee.org

Yuhanis Yusof

School of Computing, Universiti Utara Malaysia, 06010
UUM Sintok, Kedah Malaysia
yuhanis@uum.edu.my

Syafiq Fauzi Kamarulzaman

Faculty of Computer Systems & Software Engineering
Universiti Malaysia Pahang
Gambang, Pahang, Malaysia
syafiq29@ump.edu.my

***Abstract* — Electric load forecasting is undeniably a demanding business due to its complexity and high nonlinearity features. It is regarded as vital in electricity industry and critical for the party of interest as it provides useful support in power system management. Despite the aforementioned situation, a reliable forecasting accuracy is essential for efficient future planning and maximize the profits of stakeholders as well. With respect to that matter, this study presents a hybrid Least Squares Support Vector Machines (LSSVM) with a rather new Swarm Intelligence (SI) algorithm namely Grey Wolf Optimizer (GWO). Act as an optimization tool for LSSVM hyper parameters, the inducing of GWO assists the LSSVM in achieving optimality, hence good generalization in forecasting can be achieved. Later, the efficiency of GWO-LSSVM is compared against three comparable hybrid algorithms namely LSSVM optimized by Artificial Bee Colony (ABC), Differential Evolution (DE) and Firefly Algorithms (FA). Findings of the study revealed that, by producing lower Root Mean Square Percentage Error (RMSPE), the GWO-LSSVM is able to outperform the identified algorithms for the data set of interest.**

Keywords - Grey Wolf Optimizer, Least Squares Support Vector Machines, Load Forecasting, Metaheuristic algorithm, Optimization

I. INTRODUCTION

In today's modern lifestyle, electricity is one of the most influential energy sources that are essential in living and plays a significant role in almost every part of our routines activity. For that matter, forecasting the electricity load can be seen as one of the critical issue for various aspects in power systems and this include planning, operating, growing the security of power system, reducing operational costs for each power generation, transmission and delivery system as well. From industry perspective, an accurate and reliable forecasting would beneficial for decision making process, from the valuation, exploration to the development and production process. This is important not only for their future planning but also to reduce risk. Thus, profits can be maximized [1].

It is well documented that the relationship between load power and factors contributing load power is govern by high non linearity and complex in nature, hence makes this issue complicated to be solved. Previously, numerous methods have been demonstrated in dealing with load forecasting, from parametric models such as Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Moving Average with Exogenous Input (ARMAX) [2, 3] to Computational Intelligence (CI) algorithms [4, 5].

In recent decades, CI based algorithms have been proven to be able to solve many real world practical problems and this includes in electric load forecasting [5-7]. The research on this area can be seen as a hot topic that attracted the attention of various parties and the research is actively carried on [2, 8]. In [8], a kernel based technique namely Support Vector Regression (SVR) has been proposed for electric load forecasting. In the study, the SVR is integrated with Differential Empirical Mode Decomposition (DEMD) to decompose the electric load into several parts. After the decomposition process, the proposed model is realized in electric load data of New South Wales, Australia market and the New York Independent System Operator, United States of America. Based on the findings, the proposed model is able to provide good forecasting accuracy. Focusing on load forecasting in supermarket refrigeration, [2] presents a short term load forecasting based on adaptive linear time series modelling techniques. For input, local observations and weather forecasts are employed. Meanwhile, an extreme learning machines based approach for electricity load forecasting has been demonstrated which was tested on four years of data [7]. Evaluated based on Root Mean Square Error (RMSE), the evaluated approach proved its effectiveness over the other identified models. On the other hand, an advanced Wavelet Neural Network is proposed for very short term load forecasting [1].

Even though abundant of studies have been presented, nonetheless, the importance of this issue is still being debated, making this field is still wide open for improvement. In order to obtain a good forecasting result, basically, several fundamental factors need to be considered. Firstly, the forecast should have a good idea on the contributing factors that will influence the dependent variables. Secondly, there is a sufficient historical data of the contributing factors and also a skill to develop a good model that linking between dependent and independent variables. This criteria will lead to adaptiveness and robustness of forecasting model.

By considering the above factors, in this study, a hybrid Least Squares Support Vector Machines (LSSVM) [9] with a relatively new Swarm Intelligence (SI) algorithm, namely Grey Wolf Optimizer (GWO) [10] is presented for load forecasting. The GWO carries excellent features such as exploration, exploitation and capability in escaping local minima [11]. Meanwhile, the LSSVM is chosen due to its good generalization capability [12]. For the sake of optimality of LSSVM hyper parameters namely regularization parameter, γ and kernel parameter, σ^2 the GWO is utilized as an optimizer for the hyper parameters of interest.

The rest of this paper is organized as follows: Section 2 give a brief review on LSSVM while Section 3 presents theory GWO algorithm. This is followed by Section 4 which presented the hybridization of LSSVM with the GWO algorithm. Later, the implemented methodology is discussed in Section 5. Section 6 presents results and discussion and finally, Section 7 draws the conclusion of the study.

II. Regression Based on Least Suares Support Vector Machines

The LSSVM [9] is a variant of standard SVM which offers better solution strategy. As a modification of conventional SVM, LSSVM uses square errors instead of nonnegative errors in the cost function and applies equality constraint rather inequality constraint of SVM in the problem formulation. As a result, one solves a linear equations instead of Quadratic Programming (QP) solver which in practice is harder to use [9]. The adaptation of QP also raises computational complexity in training which is at least quadratic with respect to the number of training data [13]. With the reformulation of LSSVM, it simplifies complex calculation which led to easier and faster training task. Hence, a simpler optimization problem can be obtained [14]. In addition, LSSVM offers less control parameters, which are γ and σ^2, as compared to three control parameters required in SVM (C, σ^2 and ε)[15]. In addition, in terms of prediction task, LSSVM is proven to be better than SVM [16].

The main concept this approach is the mapping of the non-linear input data to high dimensional feature space, where here, a linear separation can be done. Usually, the training of the LSSVM model involves an optimal selection of regularization parameter, γ and kernel parameter, σ^2. Several kernel functions namely Gaussian Radial Basis Function (RBF) kernel, linear kernel and quadratic kernel are available. In this study, the RBF kernel is used since its suitability in dealing with nonlinear cases [17]. It is expressed as:

$$K(x, x_i) = e^{\frac{\|x - x_i\|^2}{2\sigma^2}} \tag{1}$$

where σ^2 is a tuning parameter which associated with RBF function. By using kernel function, it allows the data which are not linearly separable in input space to become linearly separable in high dimensional feature space. A brief theory of LSSVM is presented here while a detailed description can be found in [9].

III. Optimization Based on Grey Wolf Optimizer

In nature, theoretically, grey wolf [20] is classified as apex predators and placed at the top of the food chain. The grey wolf population constitutes of 4 hierarchies, namely alpha, beta delta and omega. The alpha consist of both male and female grey wolf. In this group, they are responsible in decision making and this includes hunting, sleeping place and others. With such responsibility, they are they resides at the top of the hierarchy. It is worth noting that, the dominant role is measured based on managing wise, not the strength.

Meanwhile, the duty of beta is to assist the alpha in decision making or any other activities occur in the pack. Similarly like alpha, the beta can be male or female as well. If one of the alpha deceased or become old, the beta would be the best suitor for replacement in alpha. Besides, the beta also plays role as advisor for the alpha in managing discipline of the pack.

On the other hand, the delta, have to submit the solution to alpha and beta but they dominate the omega. This group consist of scouts, sentinels, elders, hunters and caretakers. Lastly, the omega, which ranked last in the hierarchy, plays the role as scapegoat. The pseudo code of GWO-LSSVM is as shown in Figure 1 while the mathematical equations for GWO can be referred in [10].

Initialization
Evaluate hyper-parameters and calculate the fitness value based on training and validation sets using LSSVM
X_α = the best search agent
X_β = the second best search agent
X_δ = the third best search agent
while (t < Max number of iterations)
  for each search agent
    Update the position of the current search agent by (15)
  end for
  update position
  Evaluate hyper-parameters and calculate the fitness value based on training and validation sets using LSSVM
  Update X_α, X_β and X_δ
  t = t + 1
end while
return X_α

Figure 1: Pseudo Code of GWO-LSSVM.

IV. HYBRIDIZATION OF GWO-LSSVM

In GWO-LSSVM algorithm, the GWO is served as an optimization tool for the hyper-parameters of LSSVM. Generally, the LSSVM is hybridized with the GWO algorithm, where here, the forecast results obtained by LSSVM acts as a fitness function evaluation. The optimized value of LSSVM hyper-parameters can be achieved after a maximum number of iteration has been reached. In this study, the objective function is served by Root Mean Square Percentage Error (RMSPE), where the lower the RMSPE, the better the prediction accuracy (see section 6). Before the initialization, the normalized training and validation data is fed into the prediction model.

A. Initialization

Initially, the number of population and the maximum iteration are set. Fig. 5 shows the example of possible solutions which are the parameters of interest in *X*.

The number of population indicates the possible solutions which are the parameters to be optimized, namely the hyper-parameters, γ and σ^2. In this study, the number of possible solutions position is set to 20.

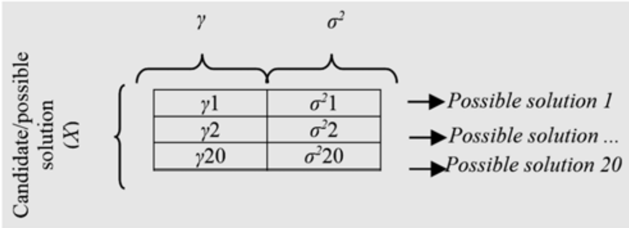

Figure 5. Representation of Search Agent Position as Possible Solutions in *X*.

B. Evaluation

The objective function of the hyper-parameters is evaluated based on training and validation set using LSSVM function which is integrated in GWO algorithm. In this study, the objective function in guided by RMSPE. The goal is to find the ideal combination of hyper-parameters (i.e. possible solutions) that will minimize the RMPSE (see section 6). The best result of RMSPE is kept as alpha's score, the second best result is kept as beta's best score while the third best result is kept as delta's score.

V. METHODOLOGY

This section presents the methodology implemented in this study. The stages incorporated of four major tasks, namely data collection and pre-process, algorithm design, algorithm development and evaluation. The simplified form of the methodology implemented is illustrated in Figure 3.

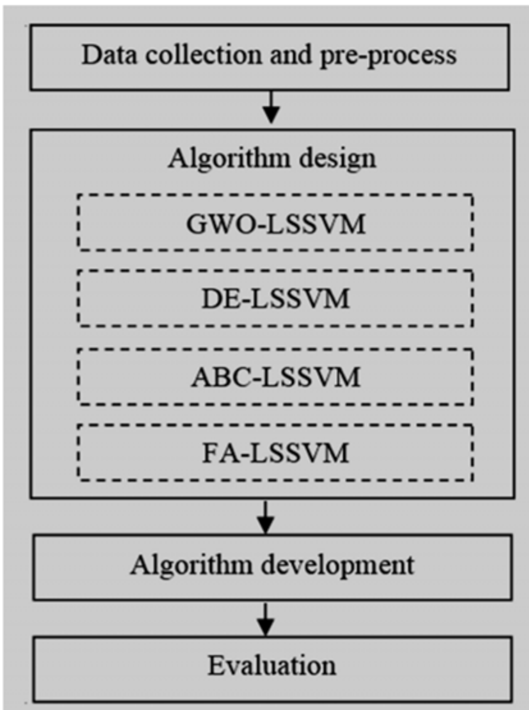

Figure 3. Methodology

A. Data Collection and Pre process

Details on employed data set employed is as tabulated in Table 1. The data set are recorded in daily basis and consist of 175 days. The output is load from day 7 and onwards. The data set covered the area in southern Malaysia.

TABLE 1: DATA SET SET FOR LOAD FORECASTING

Period	Input
1st April – 30 Sep 2014	Load Mean temperature Mean dew point Mean humidity

B. Algorithm Design

In this phase, the algorithm of GWO-LSSVM is designed accordingly which is aligned with the objectives of the study; to automatically optimize the hyper parameters of LSSVM.

For the purpose of parameter tuning, GWO is employed to automatically tune the hyper-parameters of LSSVM. To describe how this hybrid algorithm functions, LSSVM is embedded in the GWO algorithm. Here, the LSSVM acts as a fitness function evaluation. With that, the optimized value of hyper-parameters can be achieved after a stopping criterion is met. In this study, stopping criterion is maximum number of iteration. Meanwhile, the objective function is served by Root Mean Square Percentage Error (RMSPE). The lower the value of RMSPE, the better the result.

This section describes experimental setup of the conducted experiment, employed performance evaluation criteria and benchmarking algorithms.

C. Evaluation

1) Experimental setup: This phase consist of arrangement of input and output, data proportion for training, validation and testing, and properties setting for the identified forecasting algorithms. In this study, experiments have been conducted using GWO-LSSVM which is realized in load forecasting using the described data set (see sub section 5.1).

2) Input and Output Variables: In this study, the utilized inputs are a set of load time series data, mean temperature, mean dew point and mean humidity. Meanwhile the output is load from day 7 onwards.

3) Training, Validation and Testing: For simulation purposes, the data set is divided into three independent subsets namely training, validation and testing. The proportion for each subset is 70:15:15 respectively. The division is made by using experimental approach

4) Performance Evaluation Metrics: Performance Evaluation Criteria: For evaluation purposes, two statistical criteria were selected namely RMSPE, and Theil's U. These criteria interpret the learning and generalization capabilities of forecasting model. The formula for the above stated criteria are shown in (2) to (3).

$$RMSPE = \sqrt{\left(\frac{\sum_{n=1}^{N} \left(\frac{y_n - y(x_n)}{y_n} \right)^2}{N} \right)} \quad (2)$$

$$Theil'sU = \frac{\sqrt{\frac{1}{N} \Sigma_{n=1}^{N} (y_n - y(x_n))^2}}{\sqrt{\frac{1}{N} \Sigma_{n=1}^{N} (y_n)^2} + \sqrt{\frac{1}{N} \Sigma_{n=1}^{N} (y(x_n))^2}} \quad (3)$$

where $n = 1, 2, ..., N$; y_n = actual values; $y(x_n)$ = forecast values and N =Number of test data.

VI. RESULTS AND DISCUSSION

The empirical results on performance comparison of GWO-LSSVM with popular hybrid algorithms namely LSSVM optimized by three optimization algorithms; Differential Evolution (DE-LSSVM), Artificial Bee Colony (ABC-LSSVM) and Firefly Algorithm (FA-LSSVM) on two metrics, namely RMSPE and Theil's U are reported in Table 2.

Based on the recorded results in the table, for GWO-LSSVM, the optimal LSSVM hyper-parameters namely γ and σ^2 are 1 and 16.8739 respectively. With the recorded values, the RMSPE achieved are 0.0919 while for Theil's U, is is recorded 0.0498. Meanwhile, quite similar results were obtained by ABC-LSSVM and FA-LSSVM when the RMSPE recorded by both algorithms are around 0.11 while 0.06 was achieved based on Theil's U. Based on the table, DE-LSSVM ranked last when it recorded the highest RMSPE which is 0.1342. The result obtained from Theil's U seems to be agreed. Bases on the reported results, it is clear that the GWO-LSSVM showed higher accuracy as compared to the identified algorithms, in terms of lower metrics utilized.

TABLE 2: COMPARISON OF LOAD FORECASTING: GWO-LSSVM VS. ABC-LSSVM VS. DE-LSSVM

	GWO-LSSVM	DE-LSSVM	ABC-LSSVMB	FA-LSSVM
γ	1	52.5112	1	636.1848
σ^2	16.8739	824.1093	490.4455	12.9566
RMSPE	0.0919	0.1342	0.1157	0.117
THEIL'S U	0.0498	0.0735	0.0633	0.062

Comparison between the GWO-LSSVM and three other identified algorithms is plotted in Figure 3. From the figure, actual value of load is represented by straight solid line, GWO-LSSVM is represented by straight solid line with cross mark while dash-dot line is representing DE-LSSVM. Meanwhile, dash and dot marks are representing ABC-LSSVM and FA-LSSVM. Based on the figure, it is illustrated that the forecasting values produced by GWO-LSSVM are closer to the actual values compared to the identified algorithms. Meawnhile, the incapability of DE-LSSVM in this situation is clearly showed, especially on day 165. On that day, we can see that the actual value (target) is moving up, however, the produced result by DE-LSSVM showed the opposite way.

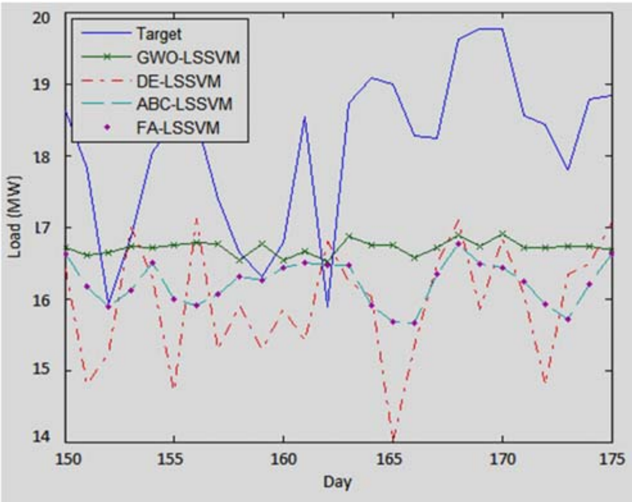

Figure 3 : GWO-LSSVM vs. DE-LSSVM vs. ABC-LSSVM vs. FA-LSSVM

On the other hand, data tabulated in Table 5 shows day to day comparison for the first 10 days of testing phase. The bold figures indicate the best value recorded for the respective day.

TABLE 3 :TARGET VS. GWO-LSSVM VS. DE-LSSVM VS. ABC-LSSVM VS. FA-LSSVM

Date	Target	GWO-LSSVM	DE-LSSVM	ABC-LSSVM	FA-LSSVM
9/4/2014	18.6146	**16.7035**	16.4081	16.6241	16.4323
9/5/2014	17.8076	**16.6087**	14.7906	16.1727	15.4005
9/6/2014	15.9215	16.6381	15.2443	**15.8867**	15.6363
9/7/2014	16.8689	16.7359	**16.9990**	16.1143	15.3454
9/8/2014	18.0396	**16.7084**	16.2970	16.4966	15.3650
9/9/2014	18.4514	**16.7460**	14.7182	15.9993	15.7342
9/10/2014	18.4764	16.7785	**17.1169**	15.9063	15.7360
9/11/2014	17.3868	**16.7659**	15.2955	16.0554	15.8809
9/12/2014	16.6479	**16.5391**	15.9115	16.3103	15.8256
9/13/2014	16.3111	16.7585	15.2881	**16.2485**	18.3413

VII. CONCLUSION

Undeniably, electricity is one of the fundamental necessity in community. In order to supply the electricity with minimal costs, reliable load forecasting is one of the critical issue. For that matter, feeding the algorithm with the sufficient and appropriate independents variables are vital. In this study, we can see that the GWO algorithm that inspired by the behaviour of grey wolf is proved to be able to optimize the hyper-parameters value of LSSVM. Comparing with another three comparable hybrid algorithms namely DE-LSSVM, ABC-LSSVM and FA-LSSVM, the GWO-LSSVM showed a lower error rate, which reflects better accuracy in prediction. For future works, the GWO-LSSVM can be improved in order to avoid the algorithm from falling to the minimum or maximum values of the LSSVM hyper parameters. With that, over fitting and under fitting problem can be avoided. Hence, better generalization of LSSVM in prediction can be achieved.

ACKNOWLEDGEMENT

This work was supported by UMP Grant (Grant No. RDU 150393)

REFERENCES

[1] M. Rana and I. Koprinska, "Forecasting Electricity Load with Advanced Wavelet Neural Networks," *Neurocomputing,* 2016.
[2] L. B. Rasmussen, P. Bacher, H. Madsen, H. A. Nielsen, C. Heerup, and T. Green, "Load Forecasting of Supermarket Refrigeration," *Applied Energy,* vol. 163, pp. 32-40, 2016.
[3] M. Y. Cho, J. C. Hwang, and C. S. Chen, "Customer Short Term Load Forecasting by using ARIMA transfer function model," in *Proceedings of the Energy Management and Power Delivery (EMPD)*, 1995.
[4] L. Xiao, W. Shao, T. Liang, and C. Wang, "A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting," *Applied Energy,* vol. 167, pp. 135-153, 4/1/ 2016.
[5] Y. Chen, Y. Yang, C. Liu, C. Li, and L. Li, "A Hybrid Application Algorithms based on the Support Vector Machine and Artificial Intelligence: An Example of Electric Load Forecasting," *Applied Mathematical Modelling,* vol. 39, pp. 2617-2632, 2015.
[6] H. Li, S. Guo, H. Zhao, C. Su, and B. Wang, "Annual Electric Load Forecasting by a Least Squares Support Vector Machines with a Fruit Fly Optimization Algorithm," *Energies,* vol. 2012, pp. 4430-4445, 2012.
[7] O. F. Ertugrul, "Forecasting Electricity Load By a Novel Recurrent Extreme Learning Machines Approach," *Electrical Power and Energy Systems,* vol. 78, pp. 429-435, 2016.
[8] G.-F. Fan, L.-L. Peng, W.-C. Hong, and F. Sun, "Electric Load Forecasting by the SVR Model with Differential Empirical Mode Decomposition and Auto Regression," *Neurocomputing,* vol. 173, pp. 958-970, 2016.
[9] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*. Leuven, Belgium: World Scientific Publishing Co. Pte. Ltd., 2002.
[10] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Anvances in Engineering Software,* vol. 69, pp. 46-61, 2014.
[11] N. Muangkote, K. Sunat, and S. Chiewchanwattana, "An Improved Grey Wolf Optimizer for Train q-Gaussian Radial Basis Fuction-link Nets," in *International Computer Science and Engineering Conference (ICSEC)*, 2014.
[12] G. Cheng, R. Guo, and W. Wu, "Petroleum Lithology Discrimination Based on PSO-LSSVM Classification Model," in *Proceedings of the*

Second International Conference on Computer Modeling and Simulation (ICCMS) 2010, pp. 365-368.

[13] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics,* vol. 42, 2012.

[14] Q. Chen, Y. Wu, and X. Chen, "Research on Customers Demand Forecasting for E-business Web Site Based on LS-SVM," in *Proceedings of the International Symposium on Electronic Commerce and Security*, 2008, pp. 66-70.

[15] P. Ou and H. Wang, "Prediction of Stock Market Index Movement by Ten Data Mining Techniques," *Modern Applied Science,* vol. 3, pp. 28-42, 2009.

[16] M. Tarhouni, K. Laabidi, S. Zidi, and M. Ksouri-Lahmari, "A Nonlinear MIMO Systems Identification Based on Improved Multi-Kernel Least Squares Support Vector Machines (Improved Multi-Kernel LSSVM)," in *Proceedings of the 8th International Multi-Conference on Systems, Signals and Devices*, 2011.

[17] S. A. Bessedik and H. Hadi, "Prediction of Flashover Voltage of Insulators using Least Squares Support Vector Machine with Particle Swarm Optimization," *Electric Power Systems Research,* vol. 104, pp. 87-92, 2013.