JESTEC

by Gunawan Dtm

Journal of Engineering Science and Technology
Vol. XX, No. Y (Year) PPP - QQQ

SOLAR IRRADIANCE FORECASTING USING KERNEL EXTREME LEARNING MACHINE : CASE STUDY AT LAMONGAN AND MUARA KARANG REGIONS, INDONESIA

Abstract

Nowadays, the development of renewable energy generation especially photovoltaic (PV) is increasing rapidly since many countries have been triggered to provide a new energy policy that promotes renewable energy applications because of public awareness to degrade global warming and rising in fuel prices. However, PV power output is intermittently affected by weather conditions. This paper proposes a method namely kernel extreme learning machine (K-ELM), as the extension of extreme learning machine in the framework of kernel learning, to forecast solar irradiance where it plays key role in the planning and operation of PV generation system to reduce procurement cost and improve the quality and safety of PV output to supply the load. To examine the efficacy of the proposed method, the forecasting model utilizes meteorological and solar irradiance data collected from Lamongan and Muara Karang areas, Indonesia. The mean absolute error (MAE) criterion is employed to measure the accurateness of the proposed method. From the simulation results, the performance of the proposed method provides better predictive accuracy compared to standard extreme learning machine (ELM), support vector machine (SVM), and least-square support vector machine (LS-SVM).

Keywords: Photovoltaic, ELM, SVM, LS-SVM, K-ELM.

1. Introduction

Indonesia is the largest archipelago in the world consisting of more than 17.000 islands and having a population estimated at 269.54 million in 2019. Moreover, due to the location of Indonesia lies on the equator line, it gives many advantages where almost all areas in this country receive direct solar light along the years. To date, this country is also endowed with rich natural resources including water, solar light, wind, geothermal and other natural resources that can be utilized as renewable energy (RE) sources [1]. The need and necessity of adequate energy harvesting from natural resources for multiple goals including economic, social, and cultural development have been revealed by most of the researchers in the global sector for decades. Along with the merit of solar light possessed by almost all areas in Indonesia, it had encouraged many researchers to develop an advanced RE technology as a solution to energy problems by harvesting energy from solar light, which is naturally replenished.

The advanced technology equipment that can harvest energy from solar light is called photovoltaic (PV) panel, which collects solar light by utilizing solar cells as a component of a PV panel and converts into electric power. However, the electric power outputs of PV panels to satisfy the need of electric power load are highly dependent on solar irradiance. Solar irradiance process is complex issues in nature due to its magnitude is affected by the amount of various climatological and aerial components such as temperature, relative humidity, wind speed, and others [2-3]. Moreover, measurement of solar irradiance in nature is an arduous issue [4] where it was computed by manual and empirical models for decades with several boundaries including case study distinctive behaviors/patterns and diversity in the results due to the high stochasticity is included in actual data [5-6]. In terms of solar irradiance predictions, there are divided into three categories, that are, intra-hour prediction, short-term prediction, and long-term prediction. The intra-hour prediction has an hourly time horizon, while short-term prediction has a daily time horizon, and long-term prediction can have weeks or months or even years of time horizon. Meanwhile, researchers in the energy field are driven to construct a new method that has high predictive accuracy to forecast the solar irradiance notably when reducing the needs of PV panel and its cost installation economically and efficiently are being the main objective.

The solar irradiance prediction using an intelligent approach based on artificial intelligence (AI) model had conducted by researchers since G. Mihalakakou, et. all had successfully proposed artificial neural network (ANN) for the first time utilized for total solar radiation time series prediction and provided satisfactory results [7]. Moreover, the development of numerous models to predict the solar irradiance using AI have investigated such as adaptive neuro-fuzzy inference systems (ANFIS) [8], artificial neural network (ANN) [9], extreme learning machines (ELM) [10], long-short term memory (LSTM) [11], and fuzzy method [12].

This paper proposes kernel extreme learning machine (K-ELM), as the extension of standard ELM based on the use of kernel function, to predict the solar irradiance in Lamongan and Muara Karang areas, Indonesia. The meteorological and solar irradiance datasets were collected from those areas due to both areas are a coastal area that has high solar irradiance over the years. The prediction learning algorithm proposed in this study is compared to standard ELM, SVM, and LS-SVM where the performance of those methods is measured by mean absolute error (MAE) obtained on training and testing processes.

2. Kernel Extreme Learning Machine (K-ELM)

Extreme learning machine (ELM) is a new kind of machine learning algorithm (MLA) introduced by G.B. Huang, et. al. in 2006 [13]. The structural model of the ELM learning algorithm is based on a feed-forward artificial neural network with a single hidden layer known as a single hidden layer feed-forward neural network (SLFN) and has the exceptional speed for mapping the correlation between input and output data.

The learning technique of ELM is employed to tackle the drawback of traditional neural network (NN) in terms of learning speed due to learning algorithms in conventional neural network minimize empirical risks by minimizing training error to fit non-linear function based on input and output datasets. Moreover, a conventional neural network commonly requires longer learning time since its network parameter is determined by iteratively and more training samples included to obtain predictive output accurately.

The learning time of ELM is faster than the standard neural network due to there is no iteration required in ELM. However, ELM has a weakness, that is, the hidden neurons of ELM calculated by the trial-and-error method (TEM) and the hidden layer needs more neurons due to the weighting parameter of ELM is selected randomly. The employment of kernel function in ELM is mapping the data from the hidden input layer into higher dimensional feature subspaces, where the non-linear pattern becomes linear and avoids computationally intensive operations. Furthermore, this learning algorithm becomes more flexible and stable due to it does not need randomly chosen nodes parameters of both hidden and input layers.

Given N data samples $\{(x_i, t_i) | x_i \in R^n, t_i \in R^m, i = 1, \ldots, N\}$, the mathematical model of ELM output is defined as follows,

$$y_M(x) = \sum_{i=1}^{M} \beta_i h_i(x) \tag{1}$$

where $\beta = [\beta_1, \beta_2, \ldots, \beta_M]$ is the output weights between the hidden layer of M neuron and output neuron. $h(x) = [h_1(x), h_2(x), \ldots, h_M(x)]$ is the output of hidden layer regarding to input x. The objective of ELM learning algorithm is minimizing the training error and the output weight at the same time as defined in (2).

$$Min : \|H\beta - T\|, \|\beta\|. \tag{2}$$

The Karush-Kuhn-Tucker (KKT) optimality condition is employed to solve the problem of (2), which can be written as follows.

$$\beta = H^T \left(\frac{1}{C} + HH^T \right)^{-1} T \tag{3}$$

where H, C, T are hidden layer output, regularization coefficient, and predicted output, respectively.

After the formulation of β (3) is obtained and substituted to (1). We can obtain the output of ELM as defined in Equation (4).

$$y(x) = h(x) H^T \left(\frac{1}{C} + HH^T \right)^{-1} T \tag{4}$$

Kernel method based on Mercer's condition was suggested by G.B. Huang [14] to overcome if the feature mapping $h(x)$ is unknown. The kernel formulation can be written as follows,

$$O = HH^T : m_{ij} = h(x_i)h(x_j) = \Omega(x_i, x_j) \tag{5}$$

The output function $y(x)$ of kernel extreme learning machine (K-ELM) can be defined as follows,

$$y(x) = [\Omega(x, x_1), \ldots, \Omega(x, x_M))]\left(\frac{1}{C} + O\right)^{-1} T \tag{6}$$

where O= HH^T and $\Omega(x, y)$ is the kernel function of hidden neurons of SLFN. Equation (6) is similar to the support vector machine (SVM) model but provides a closed-form expression for the kernel coefficients. In this work, we use radial basis function (RBF) as kernel to examine the performance of the ELM learning algorithm as defined in Equation (7).

$$\Omega(x, y) = \exp\left(\frac{\|x - y\|^2}{2\sigma^2}\right) \tag{7}$$

where σ is kernel parameter.

The performance of K-ELM depends on regularization coefficient C and kernel parameter σ, where these parameters must be selected properly. The K-ELM learning algorithm is more stable compared to standard ELM and it is faster than SVM. The network structure of K-ELM algorithm is illustrated in Fig. 1.

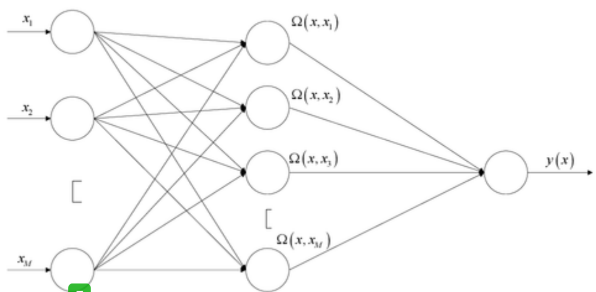


Fig. 1. Network structure of K-ELM algorithm.

3. The Implementation of Proposed Method

The K-ELM model is depicted in Fig. 2, which the features chosen as input data of K-ELM consist of temperature (C), wind speed (m/s), and humidity (%). While the output data of K-ELM as predicted data is solar irradiance (kWh/m^2/day). The procedures in designing solar irradiance forecasting model are classified into two parts, that are, training process and testing phase.

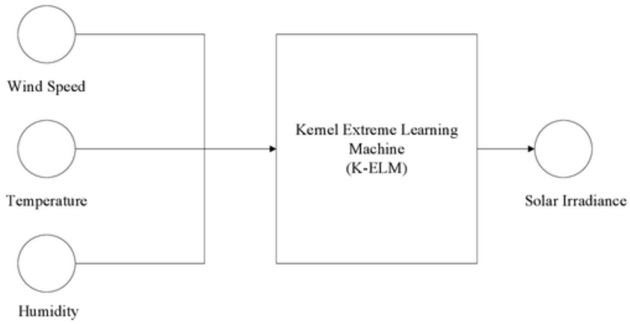


Fig. 2. Proposed K-ELM scheme.

First, the procedures to conducting the training phase on solar irradiance forecasting using K-ELM method can be described as follow:

1) Collect the dataset composed by wind speed, temperature, humidity, and solar irradiance.
2) Utilize wind speed, temperature, and humidity datasets as data input of K-ELM. While solar irradiance is used as the predicted data of K-ELM.
3) The training dataset is normalized to produce the value range [-1,1], where the mathematical model is shown in Equation (8).

   $$X_n = 2(X_p - X_{p,min}) / (X_{p,max} - X_{p,min}) - 1 \qquad (8)$$

   where, X_n is the normalized data, X_p is the original data, $X_{p,max}$, $X_{p,max}$ are the minimum and maximum values of datasets, respectively.
4) Transform the input data of K-ELM into high dimensional features using RBF kernel using Equation (7).
5) Determine the K-ELM parameters consisting of regularization coefficient C and kernel parameter σ .
6) Conduct the training process through Equations (1)-(6) to obtain the forecasted result of K-ELM using Equation (6).
7) The output data generated from ELM trainning process is denormalized using formula defined as follow,

   $$X_d = 0.5(X_n + 1) \text{ x } (X_{p,max} - X_{p,min}) + X_{p,min} \qquad (9)$$

   where, X_d is the data value after denormalization process, X_p, $X_{p,max}$, $X_{p,max}$ are the output data, minimum and maximum values of datasets before denormalization phase, respectively.
8) Compute the accurateness of predicted result. Statistical indicator namely mean absolute error (MAE) is employed for measuring the performance of proposed method as provided by

   $$MAE = \frac{\sum_{i=1}^{N} |\hat{y}_i - y_i|}{N} \qquad (10)$$

   where N is number of datasets, $\hat{y}_i$ is prediction result, and y_i is real data record.

Second, the parameters of K-ELM that obtained from training phase is utilized for testing process. The procedures of testing phase are described as follow:

1. Prepare the datasets for testing phase including wind speed, temperature, and humidity, and solar irradiance.
2. Use wind speed, temperature, and humidity datasets as data input of K-ELM. While solar irradiance is used as the forecasted data of K-ELM.
3. Set the regularisation coefficient C and kernel parameter σ of K-ELM that obtained from training process.
4. Normalize the data input using Equation (8).
5. Transform the data input of K-ELM to high-dimensional feature using kernel function in Equation (7).
6. Compute the output result using Equation (6).
7. The output result is denormalized using Equation (9).
8. Compute the degree of accuracy for K-ELM output using Equation (10).

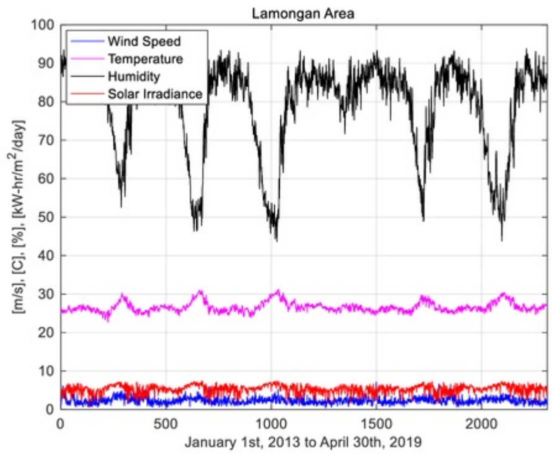
4. Results and Discussion

4.1. Forecasting data model and statistics for evaluation

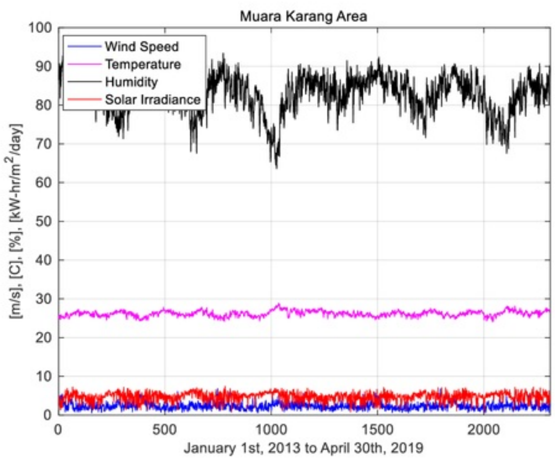
To develop the proposed forecasting method, we utilize the real meteorological and solar irradiance datasets from Lamongan and Muara Karang areas, Indonesia, which downloaded from the website of NASA prediction of worldwide energy resource [15]. The real datasets measurements were recorded from January 1st, 2013 to April 30th, 2019 (5 years 4 months). All simulations were carried out using AMD Ryzen 7 with 3.20 GHz, 16 GB of RAM Memory. The proposed prediction method is implemented using MATLAB software [16]. In this work, we divide the dataset model into two parts for training set and testing set. The meteorological and solar irradiance datasets aforementioned in section 3 are employed for training and testing processes. The datasets included for training process were recorded from January 1st, 2013 to December 31th, 2018 (2191 days). While the meteorological and solar irradiance data collected from January 1st, 2019 to April 30th, 2019 (120 days) were included as testing datasets. The patterns of meteorogical and solar irradiance datasets of Lamongan and Muara Karang regions for 5 years 4 months are shown in Fig. 3, where their maximum and minimum values are listed in Tables 1-2. The level of wind speed, temperature, humidity, and solar irradiance in Lamongan is higher than Muara Karang.

Table 1. Maximum and minimum values of datasets in Lamongan.

	Wind speed (m/s)	Temperature (C)	Humidity (%)	Solar Irradiance (kWh/m^2/day)
max	7.12	31.19	95.32	7.47
min	0.43	22.69	43.64	0.58



(a) Lamongan.



(b) Muara Karang.

Fig. 3. Meteorological and solar irradiance datasets.

Table 2. Maximum and minimum values of datasets in Muara Karang.

	Wind Speed (m/s)	Temperature (C)	Humidity (%)	Solar Irradiance (kWh/m^2/day)
max	6.86	28.88	93.44	7.46
min	0.52	24.03	63.58	0.81

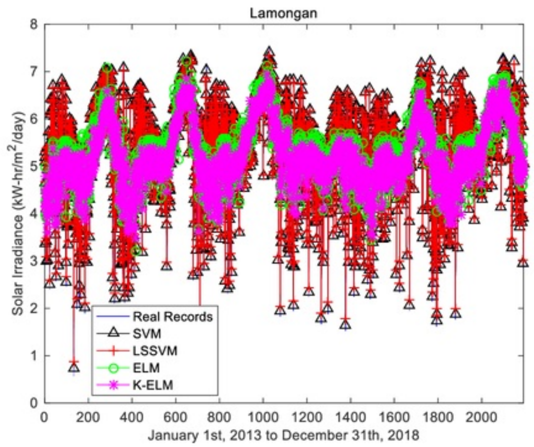
4.2. Simulation Results and discussions

Four prediction models based on machine learning including support vector machine (SVM), least-square support vector machine (LS-SVM), extreme learning machine (ELM), and proposed method called kernel extreme learning machine (K-ELM) are applied to the sample datasets to compare their performances in term of MAE and speed of computation process. The parameters of four prediction models are illustrated in Table 3.

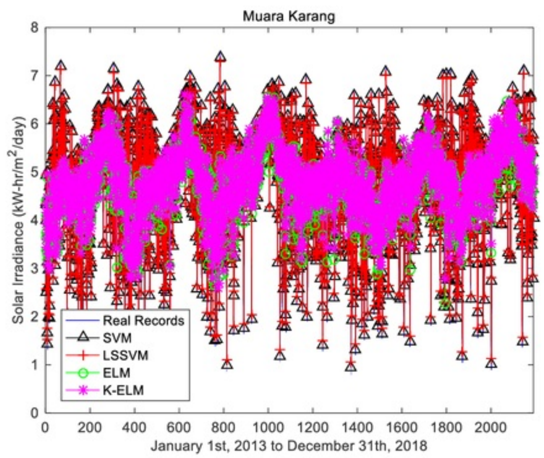
Table 3. Parameters of SVM, LS-SVM, ELM, and K-ELM.

SVM	LS-SVM	ELM	K-ELM
C = 0.0009	C = 0.0009	Hidden Neuron = 50	C = 10
σ = 15	σ = 15		σ = 50

We use radial basis function (RBF) as activation function for ELM learning algorithm, while kernel type employed for SVM, LS-SVM, and KELM are radial basis function (RBF). The real record of solar irradiance dataset and the training results of four prediction models are provided in Fig. 4. The highest error in training process as depicted in Fig. 5 is produced by K-ELM method. While the lowest error and average error values are yielded by LS-SVM approach. This condition occurs due to the meteorological and solar irradiance variations in Lamongan and Muara Karang areas have non-linear characteristics and greatly affect the forecasting process.



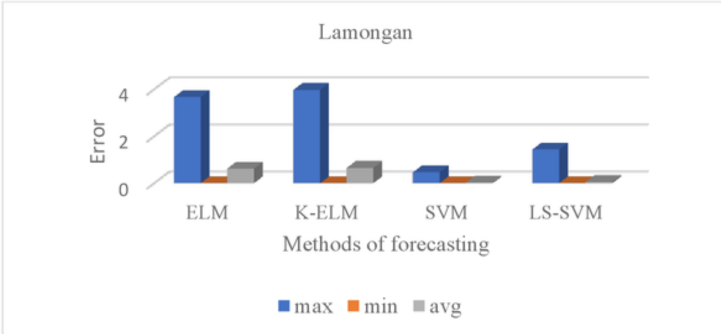
(a) Lamongan.



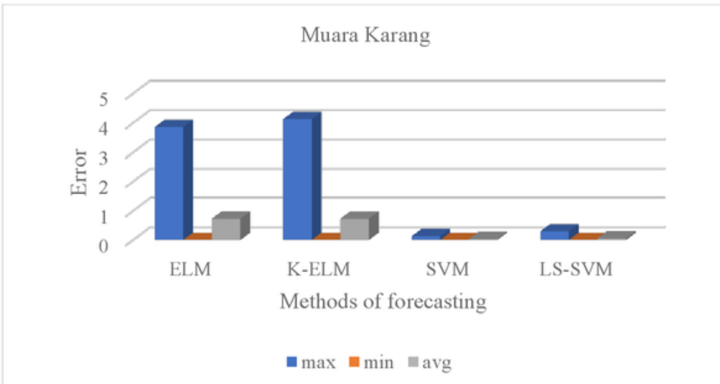
(b) Muara Karang.

Fig. 4. Training process results.

The main objective of training process is to obtain the parameter of proposed method, then, its paramater is utilized for testing phase to examine the effectiveness of the proposed method. The testing results of forecaster model are depicted in Fig. 6. The predictive accuracy criterion of prediction model for testing phase is shown in Fig. 7 where the proposed method is better than SVM, LS-SVM, and ELM due to it has the lowest MAE where the MAE value is 0.6669 and 0.791 for Lamongan and Muara Karang, respectively.



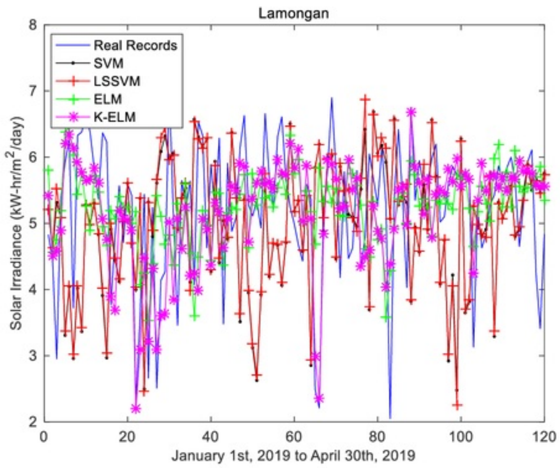
(a) Lamongan.



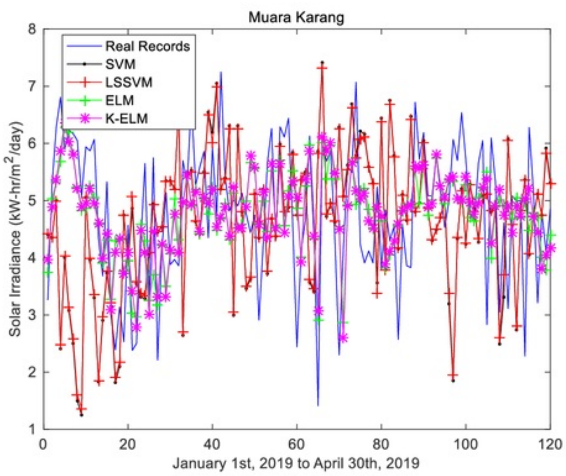
(b) Muara Karang.

Fig. 5. Maximum, minimum, and average values of error.

The computation time process of four prediction models for training and testing phase is described in Fig. 8. The proposed method is faster than SVM, LS-SVM, and ELM as shown in Fig. 8, where the computational burden time of proposed method is 0.0055 and 0.0057seconds during testing process for Lamongan and Muara Karang, respectively. In this work, the feature selection plays key role for solar irradiation forecasting. When the features selections are unrelated to the predicted response, the understanding of the relationship between the features are disturbed and encourages long-term observation of the selection of features in the future. Simulation results have shown that the selection of features with little correlation will lead to a significant decrease in prediction accuracy. For example, learning algorithm model employed in this work, which has a good predicted result for training phase, produces unsatisfactory forecasted result in testing process.

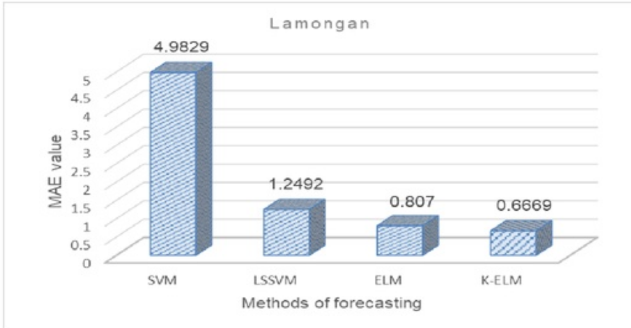


(a) Lamongan.

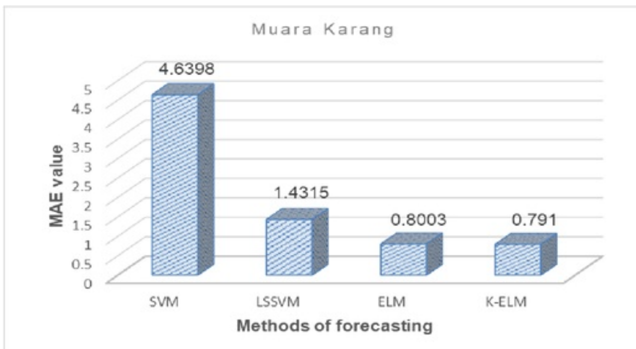


(b) Muara Karang.

Fig. 6. Testing phase results.

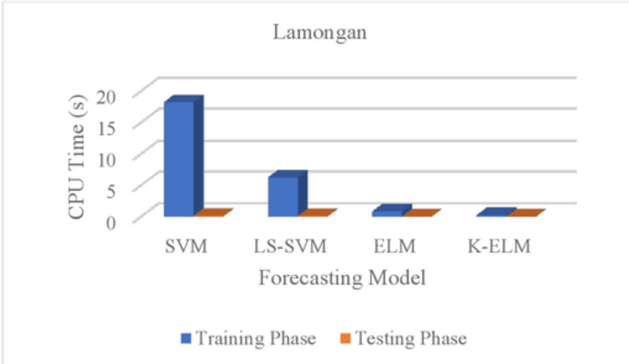


(a) Lamongan.

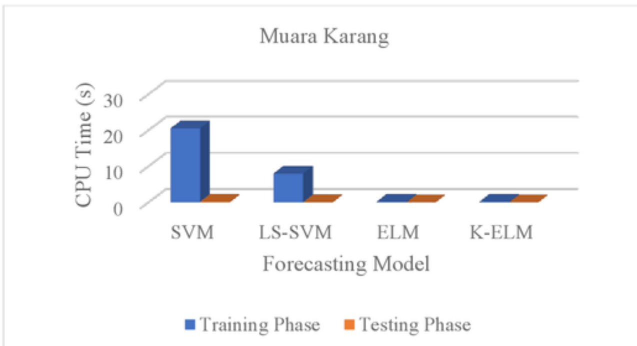


(b) Muara Karang.

Fig. 7. MAE values.



(a) Lamongan.



(b) Muara Karang.

Fig. 8. CPU times.

5. Conclusions

In this research work, K-ELM has been proposed to predict the solar irradiance in Lamongan and Muara Karang Areas, Indonesia. As shown from the simulation results, the proposed approach in terms of MAE value provides better predictive accuracy compared to standard ELM, SVM, and LS-SVM for testing phase. In viewpoint of speed of learning algorithm, the proposed algorithm is faster than other learning algorithms tested in this paper. The proposed algorithm can achieve better generalization performance and has more stable ability than ELM, SVM, and LS-SVM algorithms in term of learning speed. Since the present proposed learning algorithm is not optimized, the application of meta-heuristic approach to improve the performance of learning algorithm is necessary in the futures.

Acknowledgement

The research work described in this paper was sponsored by international research grants in collaboration between Universitas Pertamina (UP), Indonesia, and Universiti Teknologi Petronas (UTP), Malaysia.

References

1. Mardlijah; Zhai G.; Adzkiya D.; Mardianto L.; Ikhwan, M. (2019). Modified T2FSMC approach for solar panel systems. *Systems Science & Control Engineering*, 7(2), 189-197.
2. Deo R.C.; Wen X.; Qi, F. (2016). A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. *Applied Energy*, 168, 568–593.
3. Salcedo-Sanz S.; Deo R.C.; Cornejo-Bueno L.; Camacho-Gómez C.; Ghimire S. (2018). An efficient neuro-evolutionary hybrid modelling mechanism for the estimation of daily global solar radiation in the Sunshine State of Australia. *Applied Energy*, 209, 79-94.
4. Al-Shamisi M.H.; Assi A.H.; Hejase. H.A.N. (2013). Artificial neural networks for predicting global solar radiation in Al Ain City-UAE *International Journal of Green Energy*, 10(5), 443-456.
5. Chukwujindu, N.S. (2017). A comprehensive review of empirical models for estimating global solar radiation in Africa. *Renewable and Sustainable Energy Reviews*, 78, 955-995.
6. Zhang J.; Zhao L.; Deng S.; Xu W.; Zhang Y. (2017). A critical review of the models used to estimate solar radiation. *Renewable and Sustainable Energy Reviews,* 70, 314–329.
7. Mihalakakou G.; Santamouris M.; Asimakopoulos D.N. (2000). The total solar radiation time series simulation in Athens, using neural networks. *Theoretical and Applied Climatology,* 66(3-4), 185–197.
8. Zou L.; Wang L.; Xia L.; Lin A.; Hu B.; Zhu H. (2017). Prediction and comparison of solar radiation using improved empirical models and Adaptive Neuro-Fuzzy Inference Systems. *Renewable Energy*, 106, 343-353.
9. Kamadinata J. O.; Ken T.L.; Suwa T. (2019). Sky image-based solar irradiance prediction methodologies using artificial neural networks. *Renewable Energy* 134, 837-845.
10. Walch A.; Castello R.; Mohajeri N.; Guignard F.; Kanevski M.; Scartezzini, J.L. Spatio-temporal modelling and uncertainty estimation of hourly global solar irradiance using Extreme Learning Machines. (2019). *Energy Procedia*, 158, 6378-6383.
11. Qing X.; Niu Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, 148, 461-468.
12. Gomez V.; Casanovas A. (2002). Fuzzy logic and meteorological variables: a case study of solar irradianc. *Fuzzy Sets and Systems*, 126(1), 121-128.
13. Huang G.B.; Zhu Q.Y.; Siew C.K. (2006). Extreme learning machine: theory and applications. *Neurocomputing,* 70(1-3), 489-501.
14. Huang, G.B. (2014). An insight into extreme learning machines: random neurons, random features and kernels. *Cognitive Computation* 6(3), 376-390.

15. https://power.larc.nasa.gov/data-access-viewer/
16. https://www.mathworks.com/products/matlab.html

JESTEC

ORIGINALITY REPORT

15%	10%	12%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Universiti Teknologi MARA Student Paper	2%
2	www.tandfonline.com Internet Source	2%
3	res.mdpi.com Internet Source	2%
4	Proceedings in Adaptation Learning and Optimization, 2016. Publication	2%
5	"Proceedings of ELM-2014 Volume 1", Springer Science and Business Media LLC, 2015 Publication	2%
6	jestec.taylors.edu.my Internet Source	1%
7	file.scirp.org Internet Source	1%
8	Submitted to University of Nottingham Student Paper	1%

9	Submitted to Universiti Tenaga Nasional Student Paper	1%
10	pdfs.semanticscholar.org Internet Source	1%
11	Lecture Notes in Computer Science, 2015. Publication	1%
12	icitacee.undip.ac.id Internet Source	1%
13	Proceedings in Adaptation Learning and Optimization, 2015. Publication	1%

Exclude quotes	On	Exclude matches	< 17 words
Exclude bibliography	On		