Journal of Cleaner Production 176 (2018) 63–75

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Prediction of hourly solar radiation in Abu Musa Island using machine learning algorithms

A. Khosravi*, R.N.N. Koury, L. Machado, J.J.G. Pabon

Postgraduate Program in Mechanical Engineering, Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil

ARTICLE INFO

Article history:
Received 8 October 2017
Received in revised form
19 November 2017
Accepted 9 December 2017
Available online 28 December 2017

Keywords:
Solar radiation
Adaptive neuro-fuzzy inference system
Fuzzy logic
Radial basis function
Support vector regression
Multilayer feed-forward

ABSTRACT

Accurate forecasting of renewable energy sources plays a key role in their integration into the grid. This study proposes machine learning algorithms to predict the hourly solar irradiance. Forecasting models were developed based two types of the input data. The first one uses local time, temperature, pressure, wind speed, and relative humidity as input variables of the models (N_1); the second one is the time-series prediction of solar irradiance (N_2) (forecasting models only use from past time-series solar radiation values to estimate the future values). For this purpose, multilayer feed-forward neural network (MLFFNN), radial basis function neural network (RBFNN), support vector regression (SVR), fuzzy inference system (FIS) and adaptive neuro-fuzzy inference system (ANFIS) are developed. The results demonstrated that for the N_1, SVR and MLFFNN models have the maximum performance to predict the solar irradiance with $R = 0.9999$ and 0.9795, respectively. For the N_2, SVR, MLFFNN and ANFIS models have reported the correlation coefficient more than 0.95 for the testing dataset.

1. Introduction

There is a growing global demand for energy, but also for reduced strain on the environment. Considerable investments were made in alternative energy research worldwide recently. Growing demand for environment-friendly energy makes renewable energy development and production a sound investment for the future (Izadyar et al., 2016). Solar energy is a kind of the renewable energy that is obtained from the sun. The energy of the sun is converted to the thermal or electrical form the energy (Vi et al., 2012). The solar energy is distributed over a wide geographical region and the cost of this energy is stable for the long-term in the future with low maintenance costs. An important feature to design required photovoltaic systems is predicting solar energy, that machine learning algorithms are currently, the most popular for this target (Hossain et al., 2017). Several studies were proposed machine learning algorithms to predict solar irradiance in different worldwide regions (Martín et al., 2010; Alzahrani et al., 2014; Sharma and Kakkar, 2017; Ibrahim and Khatib, 2017).

Ref. (Yaiici and Entchev, 2016) designed an ANFIS model for predicting the performance parameters of a solar thermal energy system. It was considered 10 inputs and 8 outputs to develop the ANFIS model. Input and output membership functions (MFs) were considered as Gaussian and Linear, respectively. The results demonstrated that the ANFIS model can successfully predict the targets with mean relative errors less than 1% (for the stratification temperatures) and 9% (for the solar fractions). Time-series prediction of the solar irradiance after sunrise was done by Ref. (Hirata and Aihara, 2017). They proposed the recently derived infinite-dimensional delay coordinates to predict the solar irradiance after sunrise. It was obtained that the examined short-term time series prediction had effectively predictively skills.

A neuro-fuzzy method was developed by Ref. (Jović et al., 2016) to investigate solar radiation. They used four parameters that were mean sea level, dry-bulb temperature, wet-bulb temperature, and relative humidity in order to estimate the parameters influence on the solar radiation prediction. Their results demonstrated that dry-bulb temperature and relative humidity were the most dominant factors for the solar radiation prediction. Daily global solar radiation using neuro-fuzzy system was predicted by Ref. (Mohammadi et al., 2015). An ANFIS model was developed by using day of the year (n_{day}) as the only input of the model. Long-term measured data for Tabass (Iran) was used to train and test the model. It was reported that the mean absolute percentage error (MAPE), mean absolute bias error (MABE), root mean square error (RMSE) and correlation coefficient (R) to be 3.9569%, 0.6911 MJ/m^2,

* Corresponding author.
E-mail address: Alikhosravii@yahoo.com (A. Khosravi).

https://doi.org/10.1016/j.jclepro.2017.12.065
0959-6526/

0.8917 MJ/m^2 and 0.9908, respectively.

Ref. (Olatomiwa et al., 2015a) used support vector machine-firefly algorithm-based model for global solar irradiance prediction. They used meteorological data (sunshine duration ($\overline{n}$), maximum temperature (T_{max}) and minimum temperature (T_{min})) as inputs of the network in order to estimate solar radiation. The developed model was compared to an artificial neural network (ANN) and genetic programming (GP) and the results illustrated that the developed model works better than ANN and GP models. Ref. (Olatomiwa et al., 2015b) developed an ANFIS model for forecasting solar radiation in Nigeria. This investigation was done by using a series of measured meteorological data that were monthly mean minimum temperature and maximum temperature, and sunshine duration obtained from a meteorological station located in Iseyin, Nigeria. They achieved correlation coefficient for the train and test data respectively with 0.8544 and 0.6567.

Daily solar radiation prediction in a warm sub-humid environment with different soft-computing techniques was done by Ref. (Quej et al., 2017a). They used ANFIS, SVM, and ANN in order to forecast the solar radiation and obtained that SVM with requirements of daily maximum and minimum air temperature, extraterrestrial solar radiation, and rainfall has better performance than the other techniques. A comparison between improved empirical models and ANFIS for forecasting solar radiation was done by Ref. (Zou et al., 2017). In their work, the ANFIS model was compared with Expanded-Improved Bristow-Campbell Model (E-IBCM) and Improved Yang Hybrid Model (IYHM) to forecast daily global solar irradiance in China. Daily sunshine duration, precipitation, relative humidity, air pressure and daily minimum/mean/maximum temperatures were considered as input variables of the network. The results demonstrated that the ANFIS model provided the better performance compared with the empirical models.

For estimating the monthly average daily global solar radiation in Baghdad (Iraq), one model based on ANN was developed by Ref. (Rasheed et al., 2014). The network was developed based mean air temperature, maximum and minimum air temperature, relative humidity and sunshine duration. The results illustrated that the predicted solar radiation data were in good agreement with the measured values and the developed model successfully predicted the monthly mean daily global solar radiation for Baghdad city. Ref. (Quej et al., 2017b) proposed a day of the year-based (DYB) models for predicting the solar radiation in six cities located in the Yucatán Peninsula, Mexico. It was developed a new Gaussian DYB model in order to estimate daily solar radiation. The results presented that the developed model has good performance in all seasons, including in the rainy season.

Ref. (Renno et al., 2016) developed two models based on ANN for predicting the direct normal irradiance (DNI) and the global radiation (GR) for a solar application to a residential building. For each ANN models, a multilayer perceptron (MLP) was trained. To evaluate the performance of the models the MAPE, RMSE and determination coefficient (R^2) were proposed. The best ANN configuration to predict the GR was obtained with MAPE = 4.57%, RMSE = 160.3 Wh/m^2 and R^2 = 0.9918 and this values for the DNI were respectively equal to 5.57%, 17.7 w/m^2 and 0.994.

Ref. (Chiteka and Enweremadu, 2016) implemented ANNs to predict the global solar radiation in Zimbabwe. The geographical data of latitude and longitude and meteorological data of humidity pressure, clearness index and average temperature were used as inputs variables of the model. The best ANN structure was obtained with 10 neurons in the hidden layer and a tansig transfer function for the input and output layers was proposed by them. The network achieved a determination coefficient (R^2) of 99.894%.

Between renewable energy resources, solar energy has a high potential in Iran. This country with 300 clear sunny days for a year and average 2200 kWh/m^2/year solar radiation is especially important to use solar systems (Najafi et al., 2015). In the south of Iran there are several Islands that because of remarkable potential of the solar energy in these regions, it is proposed to use solar energy systems for providing the electrical energy. Prediction of the solar radiation is useful for grid operators in order to make decisions of grid operation, as well as, for electric market operators. In addition energy producers and to negotiate contracts with financial entities or utilities that distribute the produced energy also use of this solar prediction. Hence, authors have been encouraged to use machine learning algorithms in order to predict the solar radiation in Abu Musa Island that is located in the south of Iran.

The main contribution of this study is to developed machine learning algorithms in order to predict the hourly solar radiation in two separate networks. The N_1 uses four meteorological data and local time for estimating the solar radiation. The N_2 is a time-series prediction of solar radiation in which the network uses the past values to estimate the future values. For the N_1, this study provides a comparison between the MLFFNN, SVR, FIS and RBFNN in order to obtain the better one. For the N_2, a comparison between the MLFFNN, SVR, FIS and ANFIS models is reported. For the first time, in the current study, three types of ANFIS model are developed to predict the time-series hourly solar radiation. Although this study is developed to predict the hourly solar radiation for Abu Musa Island, the results of this study can help to select the optimum model of machine learning algorithms in order to estimate the hourly solar radiation for other geographical areas.

2. Methodology

2.1. Case study

Production of electrical energy in most countries is done by the fossil fuel resources (Khosravi et al., 2017). But these resources gradually decrease per year. So, replacing this resource with the renewable energy resources has been taken into consideration (Danish et al., 2017). Despite Iran's vast reserves of hydrocarbons, but this country is potentially one of the best regions for solar energy. The prediction of solar radiation on the collectors is essential for photovoltaic generation plant in enhancing the usage of solar energy for electrical production scheme. The goal of this study is forecasting the solar irradiation for Abu Musa Island (Longitude: 55.03° and Latitude: 25.87°) that is located in the south of Iran. Fig. 1 represents the variation of the solar radiation (a) and temperature (b) for 7 years in the case study zone (the data were provided by NASA (NASA-SSE, 2017)). The amount of the global horizontal irradiance for this zone is 2069 kWh/m^2 for one year. The average daily solar radiation and temperature are approximately 6 kWh/m^2 and 29°C respectively for this region.

2.2. Proposed networks

Two type of networks were designed to predict the solar irradiance in the study zone. The first network (N_1) is to predict the solar irradiance with five inputs that are pressure (hPa), temperature (K), wind speed (m/s), relative humidity (%), and local time (hour) as the inputs of the model. The second network (N_2) is based on the time-series prediction of the solar irradiance. Fig. 2 represents the N_1 and N_2 that have been considered to predict the solar radiation in this study. For evaluation of the predicted data, correlation coefficient (R), root mean square error (RMSE), and mean square error (MSE) were selected. These statistical parameters help to find the best performance for the developed models that are given in Appendix B.

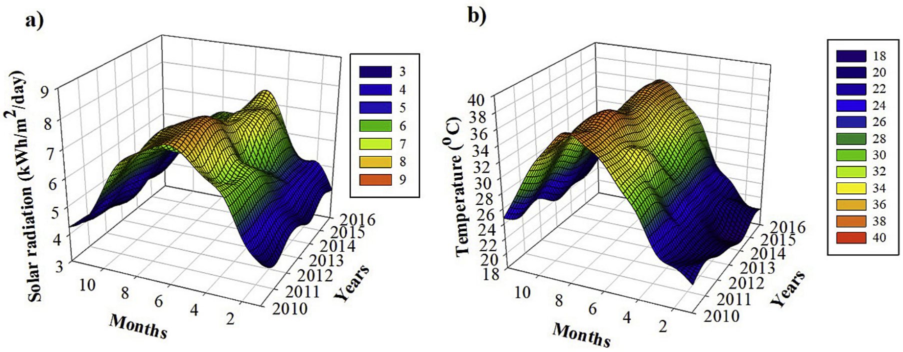

Fig. 1. Solar radiation (a) and temperature (b) for Abu Musa Island between 2010 and 2016.

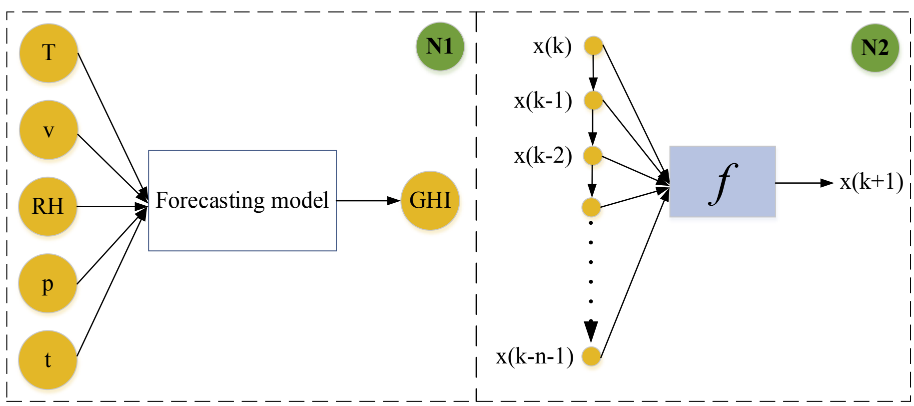

Fig. 2. Two types of proposed networks to predict hourly global horizontal irradiance (GHI).

A time-series prediction (N_2) describes a model that predict the future values of the system only using the past values. For this type of prediction, the data are described by a possible linear or non-linear autoregressive process that is obtained by Eq. (1).

$$x(k+h) = f[x(k),\ x(k-1), x(k-2), \ldots, x(k-n-1)] \qquad (1)$$

In which the past value of the x with the present value is described by the function *f*. Multi-step ahead prediction consists of predicting the h next values of the time-series is described by two ways that are: independent value prediction (training the direct model to forecast $x(k+h)$) and the second way is iterative method consists of repeating one-step-ahead predictions to the desired horizon (this strategy is shown in Fig. 2 (N_2)).

3. Machine learning algorithms

A machine learning algorithm in its most general form can be described as a function *f(x)* which takes an input vector x and generates an output vector y. Fig. 3 illustrates a classification of machine learning that contains two types of techniques (supervised learning and unsupervised learning). Supervised learning divided into two sections that are classification and regression that these two methods were applied to predict the target. Also, unsupervised learning uses clustering technique to forecast the target ("Machine Learning in, 2017).

3.1. MLFFNN model

In this study, a three-layer feed-forward neural network is applied to predict the targets. The first layer consists of a vector with five inputs; the second layer is the hidden layer that contains a nonlinear transfer function; the third layer is output layer contains a linear function. For training the MLFFNN, back-propagation (BP) algorithm was considered. This algorithm uses gradient-descent method for reducing the amount of the error(s) with adjusting the weight(s) and bias(s) (Ceylan et al., 2014). Fig. 4 demonstrates the activation function for the MLFFNN. It shows an ANN with three layers, five inputs, and one output. The MLFFNN was trained by several algorithms that are: Levenberg Marquardt (LM) (*trainlm*), BFGS Quasi-Newton (BFG) (*trainbfg*), Resilient Backpropagation (RP) (*trainrp*), Scaled Conjugate Gradient (SCG) (*trainscg*) and Bayesian Regularization (BR) (*trainbr*). For the hidden layer and output layer Tan-Sigmoid transfer function (*tansig*) and Linear transfer function (*purlin*) were selected. The *tansig* is neural transfer function. Transfer function calculates a layer's output from its net input. This function is described by Eq. (2). Also, the *purlin* that is a linear transfer function is given by Eq. (3).

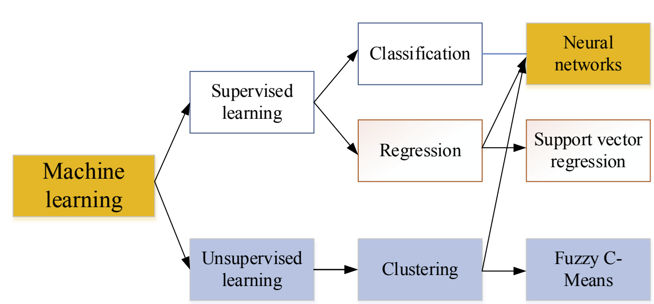

Fig. 3. Classification of the machine learning.

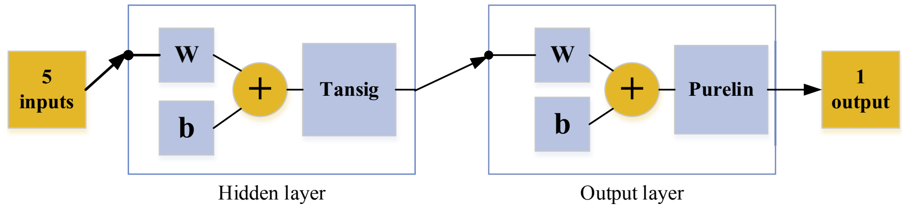

Fig. 4. Activation function of the MLFFNN model.

$$y = tansig(n) = \frac{2}{(1 + exp(-2\,n))^{-1}} \tag{2}$$

$$y = purlin(n) = n \tag{3}$$

3.2. RBFNN model

Radial Basis Function (RBF) neural networks operate on a somewhat different principle. Fig. 5 illustrates one RBFNN that contains three layers (Yadav et al., 2017). The first layer is the input layer that contains the input neurons; the second layer is the hidden layer that is composed the RBF neurons; the third layer is the output layer contains linear basis function neurons with one node per category or class of data. The weight c and one extra coefficient for each neuron were defined by the RBF neuron (for the radial basis neuron is not the bias but the width b). The weights of the RBF neuron determine the center for the RBF and actually, it calculates the Euclidean distance between the input vector and the middle of the hypersphere. The neuron's inner potential can be obtained by Eq. (4):

$$g(x) = \frac{\|x - c\|}{b} \tag{4}$$

And the combining with the *Gaussian* activation function is defined by Eq. (5):

$$f(x) = e^{-\frac{\|x-c\|}{b}} \tag{5}$$

Also, Eq. (6) indicates the final activation of the *ith* output:

$$y_i = W_{o,j} + \sum_{i=1}^{n_h} W_{i,j} e^{-\frac{\|x-c_i\|}{b_i}} \tag{6}$$

The RBFNN was designed with the *newrb* function in Matlab software. The network was trained with the 70% of the data and 30% of data were used to test the network. The optimum performance of the network was obtained by setting the various parameters of the network. For this network, the spread of RBF and the maximum number of neurons are the important user-defined parameters.

3.3. SVR model

Support Vector Machines that have been introduced by Ref. (Vapnik, 2013) are classification and regression techniques, which optimize its structure based on the input data. For training data $(x_i, y_i), \ldots, (x_n, y_n)$, where x_i are the vectors with input values and y_i the appropriate output values for x_i. The initial goal of ε-SVR is to develop a function where all errors lie under a predefined value ε but with the best generalization capacity possible (generally related to model flatness). These two conditions are imposed by Eqs. (7) and (8):

$$\min_{\omega,b,\xi,\xi^*} \frac{1}{2}\omega^T\omega + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \tag{7}$$

$$s.t. \begin{cases} y_i - (\omega, x_i + b) \le \varepsilon + \xi_i \\ (\omega, x_i + b) - y_i \le \varepsilon + \xi_i^* \\ \xi, \xi^* \ge 0 \end{cases} \tag{8}$$

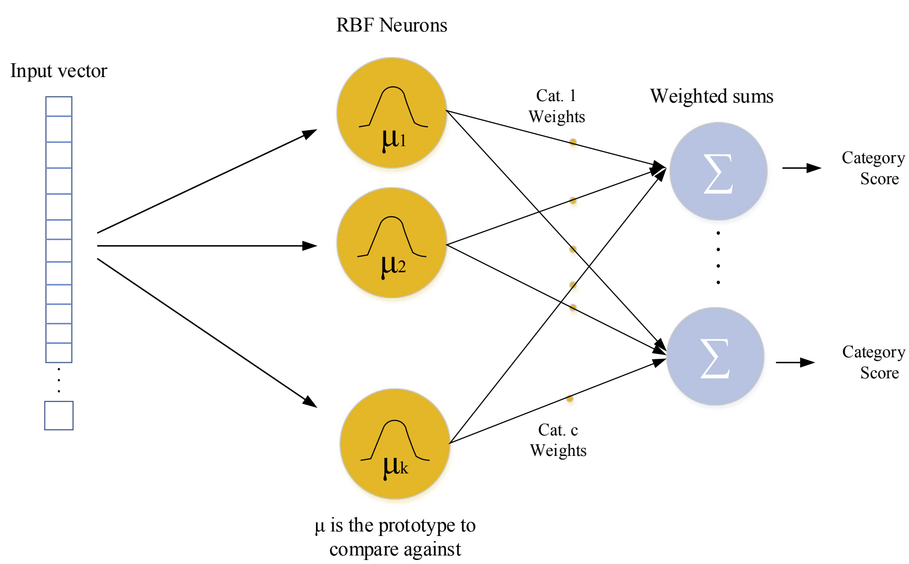

Fig. 5. Structure of the RBF.

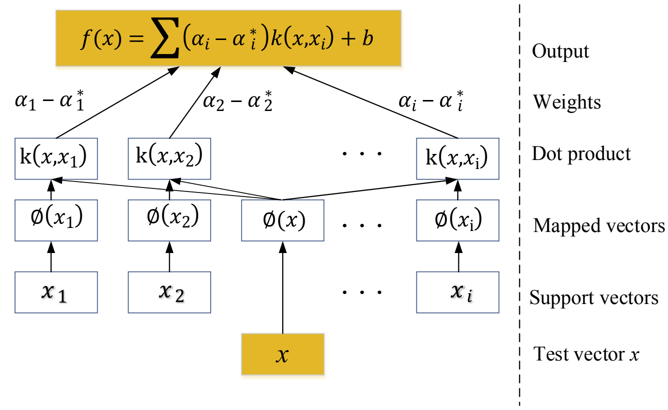

Fig. 6. SVR architecture.

In which, n is the number of samples, ξ_i shows the upper training error, ξ_i^* is the lower training error subject to the ε-insensitive tube $|y_i - (\omega, x_i + b)|$. Also, C is the regularized constant that determines the trade-off between the regularization term and the empirical error, and $C > 0$. Standard dual optimization through Lagrange multipliers is used to solve the optimization problem of Eq. (8). Once the Lagrangian is computed, several transformations are conducted until Eq. (9) is then obtained:

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \kappa(x, x_i) + b \tag{9}$$

And can be found by utilizing the properties Lagrange multipliers, Kernel trick and the optimality constraints. The Lagrange multipliers and Kernel trick are described in Appendix A. For this study RBF was used as a kernel function that its mapping space has an infinite number of dimensions. Fig. 6 shows one SVRNN architecture based on Eq. (9) by considering the Karush-Kuhn-Tucker's conditions for solving a quadratic programming problem. The value of $(\alpha_i - \alpha_i^*)$ that are nonzero are support vectors, which are applied to obtain the decision function. It is important to find the optimum three user-determined parameters that are C, ε, and σ.

3.4. FIS model

Fuzzy inference system (FIS) is represented as the process of mapping a set of input data sets into a set output data, using an approach based on fuzzy logic (an architecture of FIS is shown in Fig. 7). The process contains the following main parts: membership functions (MFs), fuzzy logic operators, and if-then rules. The main steps of a FIS are: the first part is called fuzzification that is defined as: compare the input(s) with the MFs on the previous part to obtain the membership values; the second part is defined as: combine the membership values on the premise part to get firing strength (degree of fulfillment) of each rule; the third part: generate the qualified consequents or each rule depending on the firing strength; and the forth part is called the defuzzification and aggregate the qualified consequents to produce a crisp output (Katambara and Ndiritu, 2009).

Two methods have been introduced for the FIS that are Mamdani (is the most commonly seen FIS method) and Sugeno or Takagi–Sugeno–Kang. This two types of FIS are different in the consequent of fuzzy rules. Fig. 8 shows the FIS structure that is proposed to predict the solar radiation in this study (N_1). A model with the five inputs, and one output; also Mamdani method is selected for FIS structure.

3.5. ANFIS model

Fuzzy rules will be obtained from the human expert in the most FIS systems, hence ANN was incorporated into a fuzzy system to obtain the knowledge of human expert by applying the learning algorithms. This method was used for automatic fuzzy if-then rules generation. This connection (an ANN into fuzzy system) is called neuro-fuzzy system. The most frequently used ANN in neuro-fuzzy system is RBFNN in which each node has radial basis function such as Gaussian and Ellipsoidal. There are many developed neuro-fuzzy algorithms that adaptive neuro inference system is one of them. This algorithm uses RBFNN to determine the parameters of the fuzzy system. In this model Takagi–Sugeno–Kang models are involved in framework of adaptive system. For each ANFIS, two basic learning algorithm is required that one of them is applied to find the suitable fuzzy logic rules and is called structure learning algorithm. The second one is the parameter learning algorithm to adjust the membership functions and other parameters according to desired performance from the system. In this study to obtain the fuzzy logic parameters, gradient-descent training algorithms are used.

Fig. 9 shows the ANFIS structure that has been used from two fuzzy if-then rules under Takagi–Sugeno–Kang model that are given by Eqs. (10) and (11):

$$\text{Rule 1: If } (\text{x is } A_1) \text{ and } (\text{y is } B_1) \text{ then } f_1 = p_1 x + q_1 y + r_1 \tag{10}$$

$$\text{Rule 2 : If } (\text{x is } A_2) \text{ and } (\text{y is } B_2) \text{ then } f_2 = p_2 x + q_2 y + r_2 \tag{11}$$

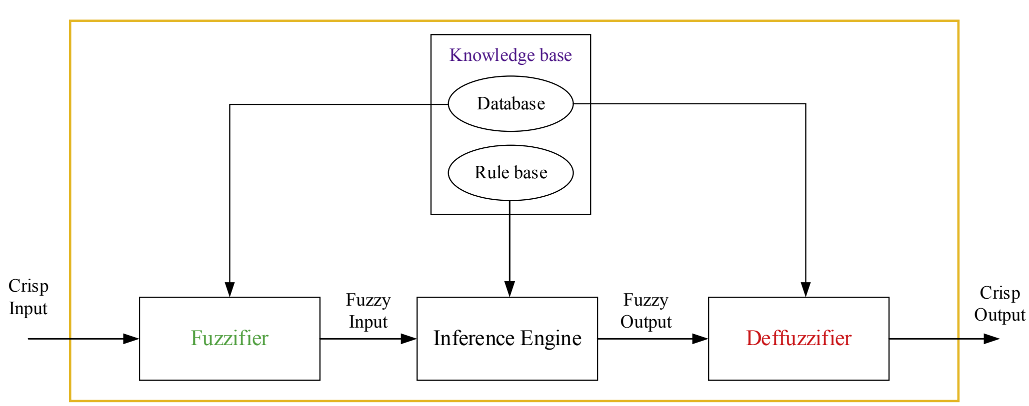

Fig. 7. The architecture of fuzzy inference system.

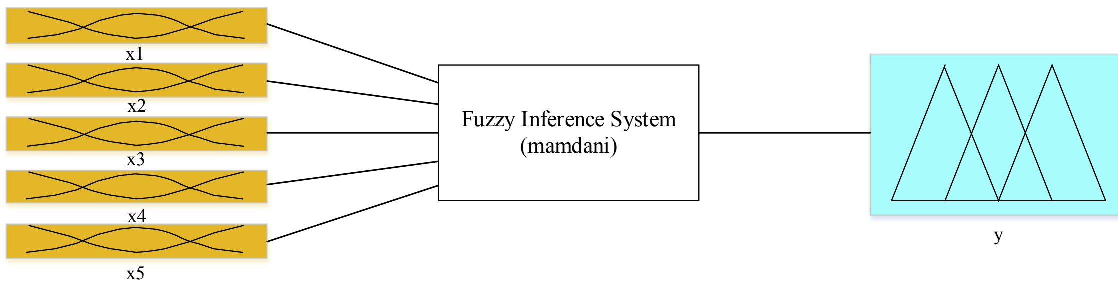

Fig. 8. FIS structure.

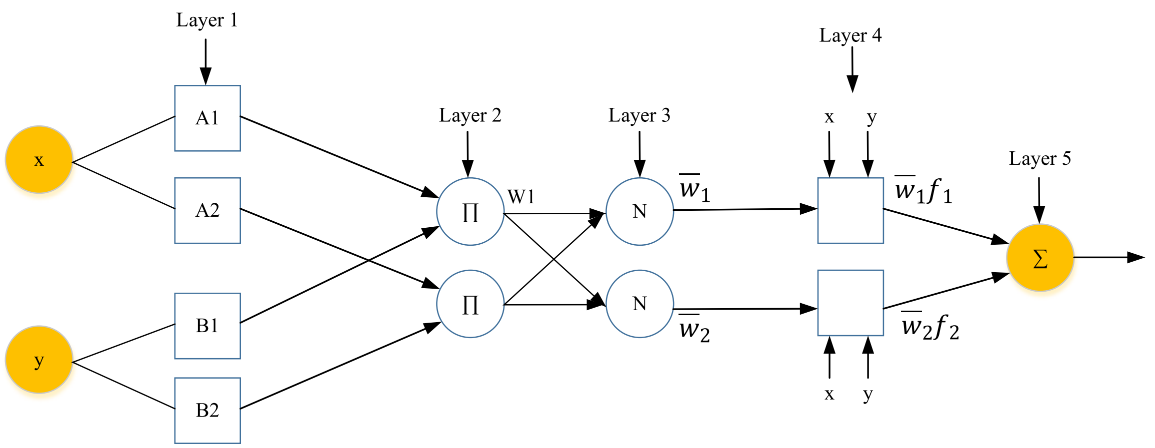

Fig. 9. ANFIS structure.

In which r_i, p_i, and q_i are the design parameters that will be obtained during the period of training phase. This structure has two adaptive layers (layers 1 and 4); layer 1 has three adjustable parameters related to input MFs (a_i, b_i and c_i) (define the center of sigma, slope, and bell type MF, respectively) that are pioneer parameters. Also, layer 4 has three adjustable parameters (r_i, p_i and q_i) that are called result parameters and are related to first-degree polynomial. The goal of using learning or training algorithms in ANFIS is to change all the adjustable parameters to compare ANFIS output with train data. Each data training process is defined by two steps: the first one the result parameters are adjusted with Least-squares method and the second step is the pioneer parameters are adjusted with gradient descent (back propagation) method which if these parameters are fixed the ANFIS output is given by Eq. (12) (Petkovic et al., 2016):

$$z = \frac{w_1}{w_1 + w_2} z_1 + \frac{w_2}{w_1 + w_2} z_2 = \overline{w}_1 (p_1 x + q_1 y + r_1) + \overline{w}_2 (p_2 x + q_2 y + r_2) \quad (12)$$

Three types of ANFIS are applied to predict the time-series data of solar radiation in this study (N_2). The first one generates FIS structure from data using grid partition (this method generates one output Sugeno-type FIS using a grid partition on the data) (ANFIS-GP). The second one generates FIS structure using subtractive clustering (this method uses subtractive clustering to create a Sugeno-type FIS structure and requires separate sets of input and output data as input arguments) (ANFIS-SC). The third one produces FIS structure using FCM clustering (this method by using fuzzy c-means (FCM) clustering generates a FIS) (ANFIS-FCM).

4. Result and discuss

There are a number of traditional methods for prediction of solar radiation that most of them are involve the statistical analysis of the data (Ayodele et al., 2016). In addition, development of an empirical model using mathematical methods, such as least squares curve fitting or nonlinear regression. In dynamic systems, these models cannot predict the targets with high accuracy. Hence, authors proposed machine learning algorithms to predict hourly solar irradiance.

For N_1:

In this network, pressure, temperature, relative humidity, wind speed and local time are used as inputs of the network and the MLFFNN, RBFNN, SVR, and FIS models are applied to predict the solar radiation for the study zone. The networks use data that have been provided by NASA (NASA-SSE, 2017). The proposed models are trained with the hourly data collections. During data training, a smaller error between the outputs and targets can be found. When the new data are interred into the neural network, it can be achieved high error in outputs that this situation is called overfitting. In this study, some methods are applied to avoid overfitting problem. These methods are:

1. Divided the data into two sections: the train data and test data. In this method, the error between the outputs and targets is evaluated for two sets of data.
2. The second method to avoid overfitting is using automated regularization. The basic idea of the Bayesian framework procedure is first to create random values and then to add them to weights and biases. In this way, Bayesian regularization provides a variance that is distributed between random values. This concept is implemented in *trainbr* algorithm in Matlab software.

3. MSE can be modified to avoid overfitting problem. This statistical parameter relies on changing a certain factor to the MSE that contains a sum square error of the weight and biases.
4. It is noteworthy that SVR model has a regularization parameter, which makes the user think about avoiding overfitting problem.

For N_2:

In this network, proposed models use the past values of hourly solar radiation to estimate the future values. The MLFFNN, SVR, FIS, ANFIS-FCM, ANFIS-GP and ANFIS-SC are designed in order to predict the targets.

4.1. MLFFNN model

A three layer feed-forward neural network was used to predict the targets in this study. The network was trained with six different data training algorithms that are LM, BFG, RP, SCG, CGP and BR (are shown in Table 1). Several tests using a different number of neurons in the hidden layer were performed. Finding the number of neurons in the hidden layer and type of data training algorithm can help to obtain the maximum performance for the developed model. The best performance was obtained with 150 neurons in the hidden layer. Also, the results show that LM has the best performance for predicting the targets with the highest correlation coefficient as 0.9887, lowest RMSE as 41.0876 (Wh/m^2) and lowest MSE as1688.188 (Wh/m^2) for the test data *(these performance evaluation measures are described in* Appendix B). As can be seen in Table 1, BR, SCG and RP illustrate a satisfactory output with a correlation coefficient close to 0.98. Also, BFG has the maximum computation time and number of iterations with the low performance of this network (N_1).

Table 2 shows the time-series prediction of the solar irradiance (N_2) for the MLFFNN in the case study zone. The best performance was obtained with 70 neurons in the hidden layer. Six data training algorithms were applied to estimate the targets. LM and BR have the best performance respectively with R = 0.9526 and 0.9570 for the test data.

4.2. RBF model

Another type of neural network that has been applied to predict the solar radiation is RBF. This network adds neurons to the hidden layer until it achieves the goal (specific MSE). The main user-determined parameters in this network are spread of radial basis functions and the maximum number of neurons. The results of the proposed RBF model is shown in Table 3. The maximum number of neurons was determined 400 neurons and the amount of the R, RMSE and MSE for different spread of the model are given in Table 3. The large spread shows the smoother the function approximation. Too small a spread illustrates that the network needs many neurons to fit a smooth function and the network might not generalize well. Too large a spread means that the network needs a lot of neurons to fit a fast-changing function. In this investigation, the network was trained with 70% of the data and 30% of data was used to test the network. The results show that the network with the spread to be 6 has the best performance that for test data R = 0.88011 and RMSE = 150.1416 (Wh/m^2).

4.3. SVR model

The SVR model was proposed to predict the targets in this study. For this model, there are three user-determined parameters to obtain the best performance of the model. Parameters ε (is the error defined by the user), σ (is a predefined value which controls the width of the Gaussian function), and *C* (cost parameter *C* handles the trade-off between errors in the predictions (first term) and complexity (second term)) account for a significant effect on the SVR performance. As a result, in choosing user-defined parameters, a large number of trials were carried out with different combinations of *C* and σ. Fig. 10 illustrates different user-defined parameters for predicting the solar radiation in which the best performance

Table 3
RBFNN with different spread parameter for finding the best performance.

Spread	RMSE (Wh/m^2)		R		MSE (Wh/m^2)	
	train	test	train	test	train	test
3	57.715	205.438	0.9856	0.7179	3331.086	42204.912
4	60.037	163.747	0.9844	0.8195	3604.504	26813.365
5	63.857	161.446	0.9824	0.8432	4077.755	26065.027
6	67.189	150.141	0.9805	0.8801	4514.375	22542.507
7	69.378	171.432	0.9792	0.8733	4813.315	29389.247
8	71.321	159.967	0.9780	0.8699	5086.758	25589.632

Table 1
Comparison of the different data training algorithms in MLFFNN (N_1).

	Algorithm	RMSE (Wh/m^2)		R		Computation time	Epoch	MSE (Wh/m^2)	
		train	test	train	test			train	test
LM	trainlm	20.2989	41.0876	0.9968	0.9887	0:00:01	19	412.0455	1688.188
BFG	trainbfg	104.4125	129.8327	0.9122	0.8696	0:01:34	2000	10901.964	16856.518
RP	trainrp	48.7892	59.5264	0.9829	0.9652	0:00:01	182	2380.385	3543.394
SCG	trainscg	39.0337	69.3104	0.9878	0.9670	0:00:03	227	1523.6318	4803.928
CGP	traincgp	59.2158	76.1573	0.97299	0.9586	0:00:04	147	3506.513	5799.929
BR	trainbr	41.453	46.6413	0.98711	0.9819	0:00:09	33	1718.352	2175.407

Table 2
Different data training algorithms in MLFFNN (N_2) to predict solar irradiance.

	Algorithm	RMSE (Wh/m^2)		R		Computation time	Epoch	MSE (Wh/m^2)	
		train	test	train	test			train	test
LM	trainlm	52.2695	80.4823	0.9787	0.9526	0:00:01	18	2732.097	6477.3995
BFG	trainbfg	69.1641	87.1539	0.9632	0.9431	0:01:27	2000	4783.67	87.1539
RP	trainrp	64.7522	85.0304	0.9677	0.9429	0:00:00	99	4192.850	7230.174
SCG	trainscg	65.0755	80.9183	0.9674	0.9485	0:00:00	65	4234.825	6547.776
CGP	traincgp	78.9906	96.1576	0.9525	0.9302	0:00:00	41	6239.518	9246.273
BR	trainbr	67.4387	73.8315	0.9649	0.9570	0:00:02	41	4547.978	5451.097

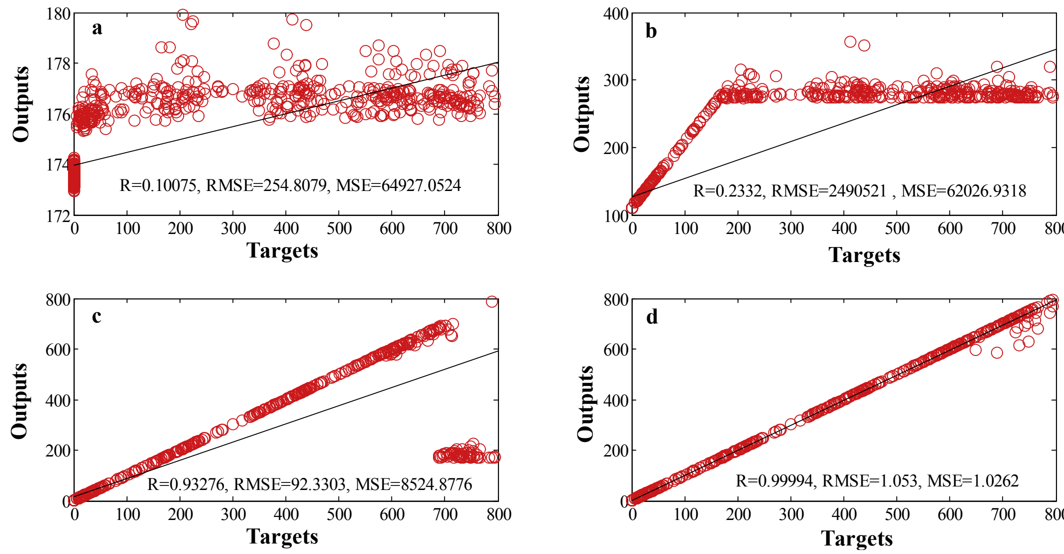

Fig. 10. The SVR model for predicting the target with $\varepsilon = 1$, $\sigma = 1$, $C = 1$ (a), $\varepsilon = 1$, $\sigma = 1$, $C = 100$ (b), $\varepsilon = 1$, $\sigma = 0.5$, $C = 500$ (c), and $\varepsilon = 0.5$, $\sigma = 1$, $C = 500$ (d).

was obtained in $\varepsilon = 0.5$, $\sigma = 1$, and $c = 500$ with $R = 0.9999$, $RMSE = 1.054$ (Wh/m^2) and $MSE = 1.0262$ (Wh/m^2).

Also, this model is applied to predict the time-series data of hourly solar radiation (N_2). To obtain the best performance of the model several iterations are done. Four models of the SVR technique are shown in Table 4 in which Model-4 (with $R = 0.9999$, $RMSE = 3.026$ (Wh/m^2) and $MSE = 9.1363$ (Wh/m^2)) indicates the highest accuracy of the model. In addition, these model are shown in Fig. 11.

Table 4
Different aspect of designer-determined parameters for the proposed SVR.

Model		RMSE (Wh/m^2)	R	MSE (Wh/m^2)
1	$\varepsilon = 0.5$, $\sigma = 0.5$, $C = 100$	276.185	0.6477	76278.28
2	$\varepsilon = 1$, $\sigma = 1$, $C = 100$	166.554	0.8119	27740.23
3	$\varepsilon = 0.5$, $\sigma = 1$, $C = 500$	14.1872	0.9987	201.277
4	$\varepsilon = 0.5$, $\sigma = 1$, $C = 1000$	3.026	0.9999	9.1363

4.4. FIS model

Forecasting model based the FIS was proposed to predict the solar irradiance. The data were divided into two sections as 70% for the network training and 30% for the test data. At first, the number of fuzzy rules in order to obtain the maximum performance was determined. Fig. 12 shows the best performance of the model by considering the four MFs as inputs and target $[\underbrace{4\ 4\ 4\ 4\ 4}_{Inputs}\quad \underbrace{4}_{Targets}]$. For this model, MFs were determined as *gaussmf* for the inputs and output and the Mamdani model as FIS structure. Gaussian MF is popular method for specifying fuzzy sets. Selecting the type of MF and its number is mostly dependent on the application and problem. The number of MFs are obtained through a trial and error process. In addition, for N_2 the train and test data are shown in Fig. 13 in which R values for the N_2 (train: 0.96207 and test: 0.94683) are greater than N_1 (train: 0.92955 and test: 0.8787) and this model for the time-series prediction has the better performance.

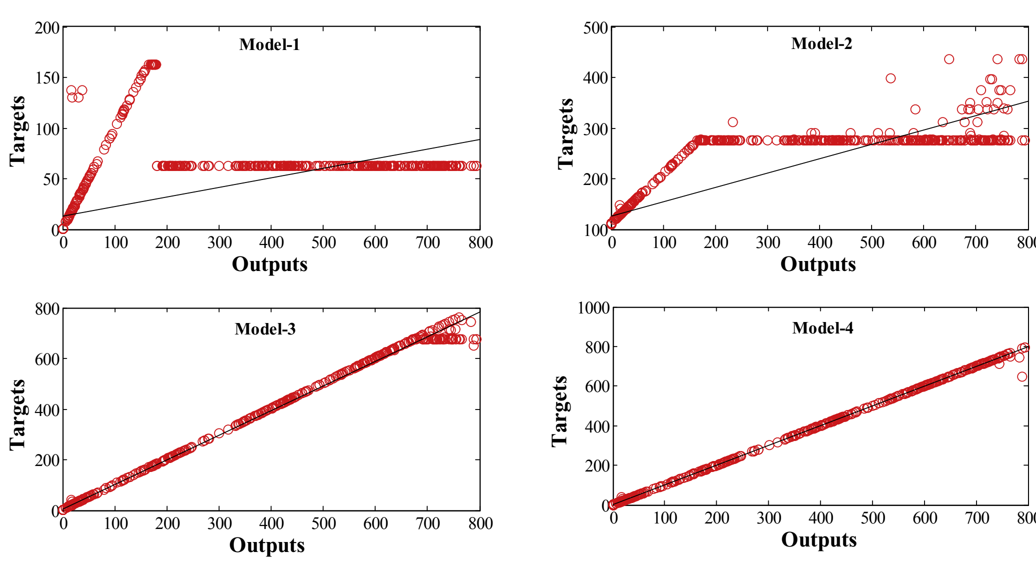

Fig. 11. The SVR model to predict the time-series hourly solar radiation (N_2) (based on Table 4).

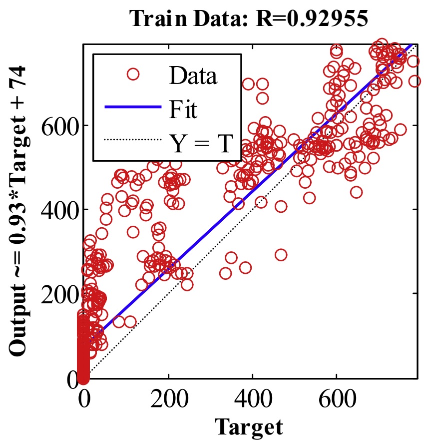

Fig. 12. Train and test data for the FIS in N_1.

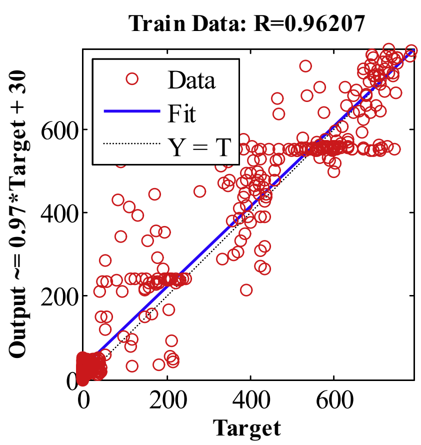

Fig. 13. Train and test data for the FIS in N_2.

Fig. 14 indicates the FIS decision surface for two inputs (x1: local time (hour) and x2: temperature (K)) and target (y: solar irradiance (Wh/m^2)). According to the decision surface, one can see the variation of local time has the maximum effect to predict the hourly solar irradiance.

4.5. ANFIS model

Three types of the ANFIS model were developed to predict the time-series solar irradiance (N_2). The first one is ANFIS-FCM that uses fuzzy c-means clustering to determine the number of rules and MFs for the antecedents and consequents. For finding the optimum performance three different ANFIS-FCM were evaluated that are shown in Table 5. This table represents the results of this investigation in which was used from R and RMSE for evaluating the outputs (for train and test data). ANFIS-FCM (2) has the best performance with 10 clusters and partition matrix exponent = 2, (this parameter controls the amount of fuzzy overlap between clusters that larger values demonstrate a greater degree of overlap). Also, the input and output MFs are *gaussmf* and *linear* respectively. Fig. 15 illustrates the structure of the ANFIS-FCM (2) with four inputs and one output. Ten clusters were generated for each input with 10 rules.

The second one is ANFIS-SC that uses subtractive clustering to generate a Sugeno-type FIS structure. The main parameter for designing this model is the influence radius (radii) which is a vector that specifies a cluster center's range of influence in each of the data dimensions. This model was used to predict the target with different radii that Table 5 shows ANFIS-SC (3) has the maximum performance.

The third one is ANFIS-GP that uses grid partition on the data for generating a single-output Sugeno-type FIS. Input MF type is *gaussmf* and the output MF type is *linear* for this model. Also, the model evaluated with different number of MFs that are shown in Table 5. The values of correlation coefficient and RMSE for this

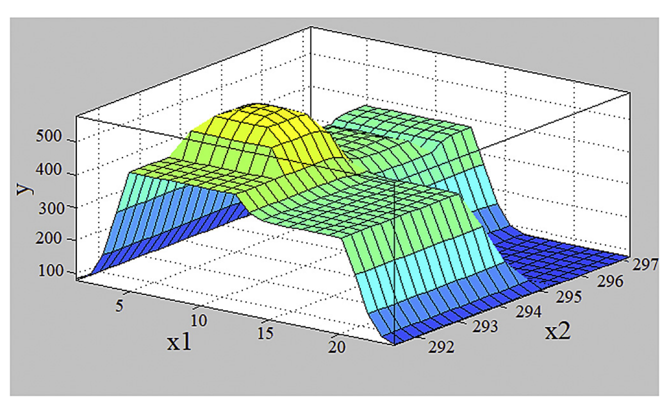

Fig. 14. FIS decision surface for the solar radiation prediction, x1: local time (hour), x2: temperature (K), y: solar irradiance (Wh/m^2).

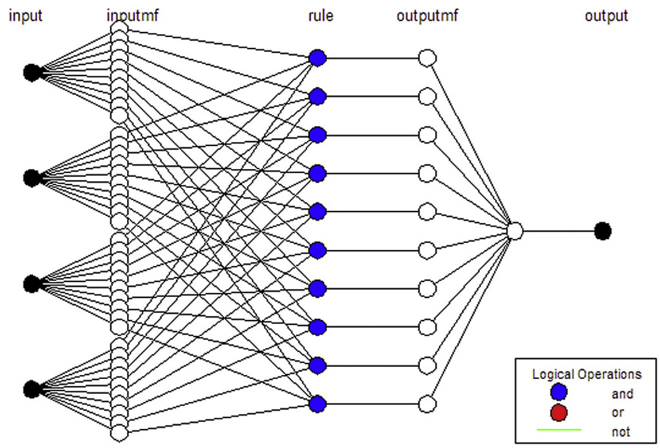

Fig. 15. Structure of the ANFIS-FCM (2).

model show the ANFIS-GP (3) with 4 MFs has the maximum correlation coefficient as 0.9493 and minimum RMSE as 86.1513 (Wh/m^2) for the test data. The maximum performance for the current investigation in ANFIS-GP model is obtained with the higher MFs (as can be seen in Table 5).

4.6. Comparison of the developed models

Fig. 16 demonstrates a comparison between the developed models to predict the solar radiation (N_1). This graph shows the correlations between the targets and outputs for all data. The SVR model has the best performance to forecast the hourly solar irradiance with the correlation coefficient as 0.99839 for all data. Also, the MLFFNN reports the satisfactory outputs with the $R = 0.98481$ and RMSE $= 44.5356$ (Wh/m^2). The RBFNN illustrates an acceptance correlation between the outputs and targets in the training stage but the test data have lower performance as compared to the training stage. In addition, FIS model can predict the targets with the correlation coefficient of 0.8787 and RMSE of 171.256 (Wh/m^2).

Fig. 17 shows the correlation between targets and outputs for time-series prediction of solar irradiance (N_2) with different proposed models. Although all the developed models can successfully predict the targets with approximately above 0.95% accuracy, SVR has the best performance between the all developed models with $R = 0.9999$, and RMSE $= 3.0266$ (Wh/m^2).

The developed models were applied to estimate the solar irradiance in the case study region for 72 h and the outputs are shown in Fig. 18 (for N_1). As before mentioned the SVR model has the best performance for forecasting the targets and also it can be seen that the other developed models have reported acceptance outputs.

Table 5
Different types of ANFIS methods.

			RMSE		R	
FCM	Number of clusters	Partition Matrix Exponent	Train	Test	Train	Test
FCM(1)	5	2	59.7021	88.719	0.9715	0.94613
FCM(2)	10	2	56.671	83.1365	0.97436	0.95279
FCM(3)	20	4	53.1077	86.8542	0.97752	0.94777
Sub-Clustering	Influence Radius					
Sub-Clustering (1)	0.1		56.0014	10398.3099	0.97497	0.19884
Sub-Clustering (2)	0.2		53.0843	88.8146	0.97754	0.94497
Sub-Clustering (3)	0.4		56.0832	84.8125	0.9749	0.94956
Grid-Part	Number of MFs					
Grid-Part(1)	2		58.1909	86.3136	0.97295	0.9483
Grid-Part(2)	3		54.4955	88.3549	0.97632	0.94535
Grid-Part(3)	4		50.0626	86.1513	0.98005	0.94938

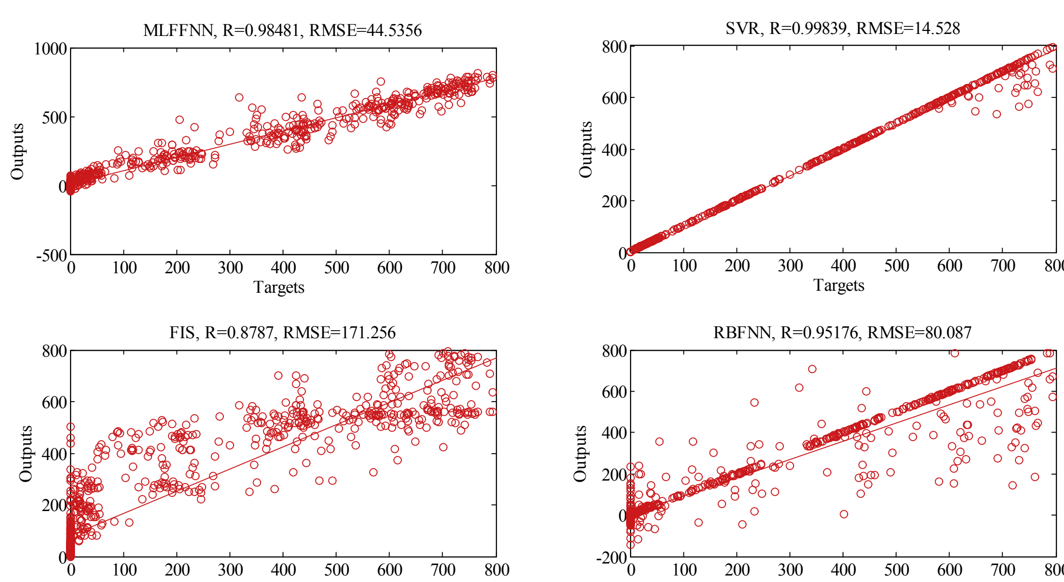

Fig. 16. Comparison of the forecasting model in N_1.

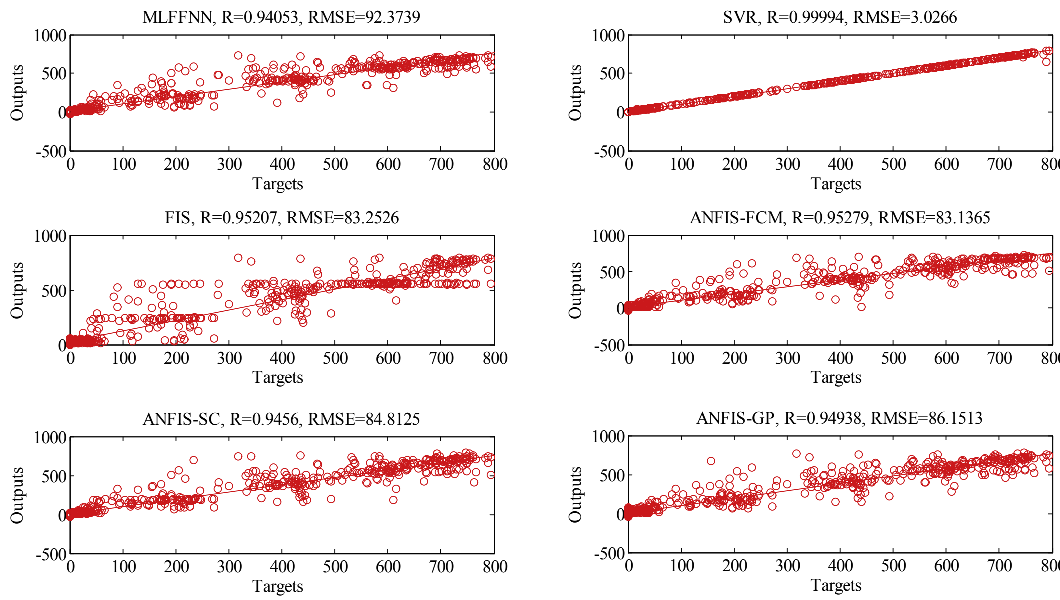

Fig. 17. Time series prediction of different proposed models.

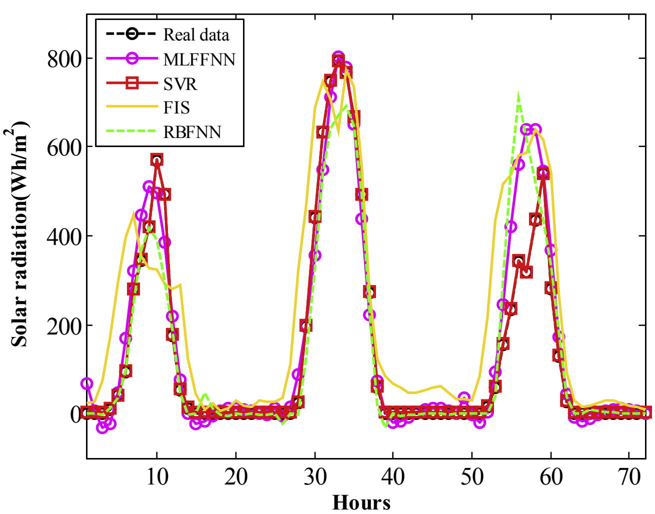

Fig. 18. Prediction of solar irradiance for 72 h (N_1).

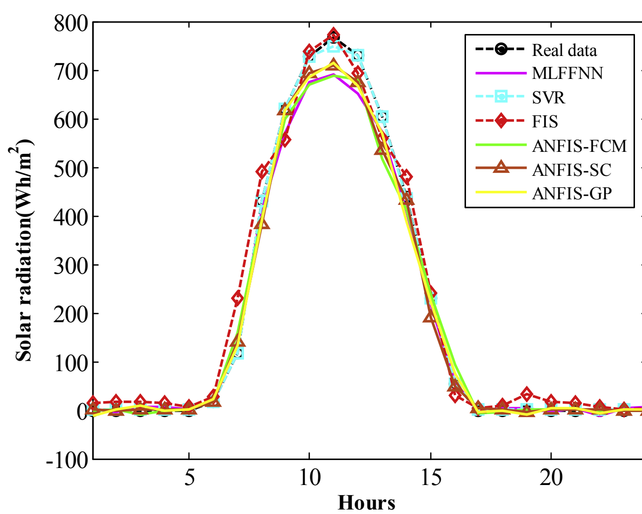

Fig. 19. Prediction of solar irradiance for 24 h (N_2).

Also, Fig. 19 indicates the solar irradiance prediction based time-series data (N_2) for the future 24 h in the study zone. The outputs are very close to each other but for this prediction FIS and SVR have the better performance.

Table 6 demonstrates the correlation coefficient for some different proposed models in order to estimate solar irradiance from literature studies. Hourly solar radiation prediction is a dynamic system and statistical models cannot estimate the targets with the high accuracy. As can be seen in Table 6, Ref. (Chen et al., 2013) reported approximately $R = 0.7$ for statistical methods in order to forecast solar radiation. But other studies have shown that machine learning algorithms can predict this system with high efficiency. This article was shown that the developed models can predict the targets with the high performance. Ref. (Gupta and Singhal, 2016) developed an ANN with maximum accuracy ($R = 0.9551$) to predict the solar radiation, but this study obtained an ANN with $R = 0.9887$ in the test data. Also developed SVR model reported a performance higher than Ref. (Ramedani et al., 2014) model.

5. Conclusion

Solar power is a vast, free and renewable resource that can be used to produce electricity. Solar-generated electricity produces no greenhouse gases or emissions of any kind. Solar energy is a commercially-proven, rapidly growing form of electricity generation. Machine learning algorithms were used in order to predict the solar radiation in the study region. In first methodology (N_1), pressure, temperature, relative humidity, wind speed and local time were used to predict solar irradiance. In this network, the amount of the solar radiation that has been recorded during the past few months or years was used to train the network in order to find the solar irradiance with the new inputs. Indeed, it is trained a network to predict the solar irradiance with the before mentioned inputs and without solar irradiance data collection. The second methodology (N_2) is time-series prediction that developed models used the past values measured solar radiation in order to estimate the future values. Thus the developed models always require solar

Table 6
Performance of the different proposed models for predicting solar irradiance from literature review.

	Model	R		Model	R
Chen et al. (Chen et al., 2013)	Statistical method	0.7	Quej et al. (Quej et al., 2017a)	SVM	0.8209
	Fuzzy algorithm	0.8613		ANFIS	0.8024
	Neural network	0.8915		ANN	0.8012
Olatomiwa et al. (Olatomiwa et al., 2015a)	SVM-Firefly Algorithm	0.8918	Ibrahim et al. (Ibrahim and Khatib, 2017)	ANN	0.828413
	ANN	0.8635		ANN-FFA	0.8699
	Genetic Programming	0.7895		Random forests	0.7947
Ramedani et al. (Ramedani et al., 2014)	ANN	0.8938	Bou-Rabee et al. (Bou-Rabee et al., 2017)	ANN	85.6
	ANFIS	0.8949	Gupta et al. (Gupta and Singhal, 2016)	ANN	0.9551
	SVR-RBF	0.8888	Bakirci (Bakirci, 2017)	Empirical	0.8831

irradiance database to forecast the future values. This type of forecasting model is valuable for grid operators in order to make decisions of grid operation and also for electric market operators. For the N_1, the MLFFNN, RBFNN, SVR, and FIS models were proposed and compared to each other. For the N_2, the MLFFNN, SVR, FIS, ANFIS-FCM, ANFIS-SC, and ANFIS-GP were developed to forecast solar radiation. This study has led to conclusions in the following, for the N_1:

- The MLFFNN was trained with six data training algorithms. The results illustrated that LM and BR have highest correlation coefficient for the test data respectively with 0.9887 and 0.9819. The RBFNN was implemented with six different spread parameter that for this analysis the best performance was obtained in spread to be 6 in which R and RMSE for the test data were 0.88011 and 150.1416, respectively. Also, the SVR model with different user-determined parameters was applied to predict the targets. This model has shown the highest R close to 1 and RMSE = 1.053 (Wh/m^2). In addition, the FIS model was applied in order to estimate the solar irradiance. The best performance was achieved by considering the four MFs for each input and target. For this state R and RMSE for the test data are 0.8787 and 171.256 (Wh/m^2), respectively. The comparison of the developed models illustrated that the SVR model and MLFFNN have the maximum efficiency for forecasting the solar radiation.

For the N_2:

- Time-series data of hourly solar irradiance was predicted with the MLFFNN. For this purpose, different data training algorithms were used in order to obtain the better performance. The results illustrated that the LM and BR training algorithms have the maximum correlation coefficient for the test data respectively with 0.9526 and 0.9570. The maximum performance was obtained with 70 neurons in the hidden layer. It can be observed, this model has shown a better performance for N_1 compared to N_2. The SVR models reported the maximum performance for time-series prediction with R = 0.9999 and RMSE = 9.1363 (Wh/m^2). Also, the FIS model demonstrated the better outputs for the N_2 compared to the N_1. This model estimated the targets with approximately 95% accuracy in the test data. In addition, three types of ANFIS model were evaluated in order to forecast the target. ANFIS-FCM with 10 clusters has shown the best performance with R = 0.95279 and RMSE = 83.1365 (Wh/m^2) for the test data. As a result, for the N_2, although the SVR and ANFIS-FCM have shown the better performance for time-series forecasting, the other developed models also reported the outputs with around 94% accuracy in the test data.

Nomenclature

ANN	artificial neural network
ANFIS	adaptive neuro-fuzzy inference system
FIS	fuzzy inference system
FCM	fuzzy c-means
GP	grid partition
MF	membership function
MLFFNN	multilayer feed-forward neural network
MSE	mean square error
purelin	Linear transfer function
P	pressure (hPa)
R	correlation coefficient
RH	relative humidity (%)
RMSE	root mean square error
RBF	radial basis function
SC	subtractive clustering
SVR	support vector regression
T	temperature (K)
t	local time
tansig	hyperbolic-tangent sigmoid
v	wind speed (m/s)
R	correlation coefficient

Appendix A

A.1. Lagrange multipliers

In Eq. (9) α_i, α_i^* indicate the Lagrange multipliers. They can be obtained by maximizing the dual function of Eq. (7) and then can be introduced by:

$$\max_{\alpha_i,\alpha_i^*} \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i - \alpha_i^*) \times \left(\alpha_j - \alpha_j^*\right) x_i, x_j \tag{A1}$$

$$s.t.\begin{cases} \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i \le C \\ 0 \le \alpha_i^* \le C \\ i = 1, 2, \dots, n \end{cases} \tag{A2}$$

In Eq. (7), only some $(\alpha_i - \alpha_i^*)$ are not equal to zero, which comes from the Karush-Kuhn-Tucker's conditions of solving a quadratic programming problem. The support vector itself refers to the approximation error of data point on non-zero coefficient equal or larger than ε. Because errors lower than ε are acceptable, the data from the training set inside the "ε-tube" do not contribute to the cost nor the solution of the problem.

A.2. Kernel trick

The key to non-linear extension of the SVR is Eq. (A.1) and the existence of the so-called kernel trick. The dot product of x_i, x_j from Eq. (A.1) becomes a kernel function $\varnothing(x_i), \varnothing(x_j) = k(x_i, x_j)$ in the case of non-linearity. The function $\varnothing : R^d \rightarrow \mathscr{H}$ presents the idea of mapping the input space into a feature space with a higher dimension. The common kernel function that satisfy this issue are:

Linear kernel: $k_{lin}(x, x') = x, x'$ (A3)

Polynomial kernel: x, x' , $p \in N$ (A4)

Radial basis function (RBF): $K(x, x') = exp(-\sigma x - x'), \ \sigma \in R, \ \sigma > 0$ (A.5)

Appendix B

B.1. Performance Evaluation Measures

For evaluating the performance of the developed models, mean square error (MSE), root mean square error (RMSE), and correlation coefficient (R) were used that are given by the following equations:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \tag{B1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 \tag{B2}$$

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{B3}$$

In which x_i, y_i, $\bar{x}$, $\bar{y}$, and n are observed value, predicted value, mean of observed data, mean of predicted data, and number of data respectively.

References

Alzahrani, A., Kimball, J.W., Dagli, C., 2014. Predicting solar irradiance using time series neural networks. Proc. Comput. Sci. 36 (C), 623–628.

Ayodele, T.R., Ogunjuyigbe, A.S.O., Monyei, C.G., 2016. On the global solar radiation prediction methods. J. Renew. Sustain. Energy 8 (2), 23702.

Bakirci, K., 2017. Prediction of global solar radiation and comparison with satellite data. J. Atmos. Solar-Terrestrial Phys. 152–153, 41–49. December 2016.

Bou-Rabee, M., Sulaiman, S.A., Saleh, M.S., Marafi, S., 2017. Using artificial neural networks to estimate solar radiation in Kuwait. Renew. Sustain. *Energy Rev.* 72, 434–438. November 2016.

Ceylan, I., Gedik, E., Erkaymaz, O., Gürel, A.E., 2014. The artificial neural network model to estimate the photovoltaic modul efficiency for all regions of the Turkey. Energy Build. 84, 258–267.

Chen, S.X., Gooi, H.B., Wang, M.Q., 2013. Solar radiation forecast based on fuzzy logic and neural networks. Renew. Energy 60, 195–201.

Chiteka, K., Enweremadu, C.C., 2016. Prediction of global horizontal solar irradiance in Zimbabwe using artificial neural networks. J. Clean. Prod. 135, 701–711.

Danish, B., Zhang, B., Wang, Wang, Z., Jul. 2017. Role of renewable energy and non-renewable energy consumption on EKC: evidence from Pakistan. J. Clean. Prod. 156, 855–864.

Gupta, S., Singhal, R., 2016. Prediction of global solar radiation in UAE using artificial neural networks. J. Sustain. Dev. Energy Water Environ. Syst. 4 (2), 94–106.

Hirata, Y., Aihara, K., 2017. Improving time series prediction of solar irradiance after sunrise: comparison among three methods for time series prediction. Sol. Energy 149, 294–301.

Hossain, M., Mekhilef, S., Danesh, M., Olatomiwa, L., Shamshirband, S., Nov. 2017. Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. J. Clean. Prod. 167, 395–405.

Ibrahim, I.A., Khatib, T., 2017. A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm. Energy Convers. Manag. 138, 413–425.

Izadyar, N., Ong, H.C., Chong, W.T., Mojumder, J.C., Leong, K.Y., Dec. 2016. Investigation of potential hybrid renewable energy at various rural areas in Malaysia. J. Clean. Prod. 139, 61–73.

Jović, S., Aničić, O., Marsenić, M., Nedić, B., 2016. Solar radiation analyzing by neuro-fuzzy approach. Energy Build. 129, 261–263.

Katambara, Z., Ndiritu, J., 2009. A fuzzy inference system for modelling streamflow: case of Letaba River, South Africa. Phys. Chem. Earth 34 (10–12), 688–700.

Khosravi, A., Koury, R.N.N., Machado, L., 2017. Thermo-economic analysis and sizing of the components of an ejector expansion refrigeration system. Int. J. Refrig.

Martín, L., Zarzalejo, L.F., Polo, J., Navarro, A., Marchante, R., Cony, M., 2010. Prediction of global solar irradiance based on time series analysis: application to solar thermal power plants energy production planning. Sol. Energy 84 (10), 1772–1781.

Mohammadi, K., Shamshirband, S., Tong, C.W., Alam, K.A., Petković, D., 2015. Potential of adaptive neuro-fuzzy system for prediction of daily global solar radiation by day of the year. Energy Convers. Manag. 93, 406–413.

Najafi, G., Ghobadian, B., Mamat, R., Yusaf, T., Azmi, W.H., 2015. Solar energy in Iran: current state and outlook. Renew. Sustain. Energy Rev. 49, 931–942.

NASA-SSE [Online]. Available: www.soda-pro.com http://www.soda-pro.com/web-services/radiation/nasa-sse [Accessed 1 September 2017].

Olatomiwa, L., Mekhilef, S., Shamshirband, S., Mohammadi, K., Petković, D., Sudheer, C., 2015. A support vector machine-firefly algorithm-based model for global solar radiation prediction. Sol. Energy 115, 632–644.

Olatomiwa, L., Mekhilef, S., Shamshirband, S., Petkovic, D., 2015. Adaptive neuro-fuzzy approach for solar radiation prediction in Nigeria. Renew. Sustain. Energy Rev. 51, 1784–1791.

Petkovic, D., Pavlovic, N.T., Cojbasic, Z., 2016. Wind farm efficiency by adaptive neuro-fuzzy strategy. Int. J. Electr. Power Energy Syst. 81, 215–221.

Quej, V.H., Almorox, J., Arnaldo, J.A., Saito, L., 2017. ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. J. Atmos. Solar-Terrestrial Phys. 155, 62–70. February.

Quej, V.H., Almorox, J., Ibrakhimov, M., Saito, L., 2017. Estimating daily global solar radiation by day of the year in six cities located in the Yucatán Peninsula, Mexico. J. Clean. Prod. 141, 75–82.

Ramedani, Z., Omid, M., Keyhani, A., Shamshirband, S., Khoshnevisan, B., Nov. 2014. Potential of radial basis function based support vector regression for global solar radiation prediction. Renew. Sustain. Energy Rev. 39, 1005–1011.

Rasheed, D.I.B., AL-Naimi, H., AL-Salihi, Ali M., 2014. Neural network based global solar radiation estimation using limited meteorological data for Baghdad, Iraq. Int. Energy Environ. Found. 5 (1), 79–84.

Renno, C., Petito, F., Gatto, A., 2016. ANN model for predicting the direct normal irradiance and the global radiation for a solar application to a residential building. J. Clean. Prod. 135, 1298–1316.

Sharma, A., Kakkar, A., 2017. Forecasting daily global solar irradiance generation using machine learning. Renew. Sustain. Energy Rev. 0–1.

Vapnik, V., 2013. The Nature of Statistical Learning Theory. Springer science & business media.

Vi, R., Melo Reis, De, Oliviera, R.N., Machado, L., Koury, R.N.N., 2012. Using a heat pump as an alternative to support solar collector for water heating in Brazil. Int. J. Air-Condition. Refrig. 20 (3), 1250013.

Yadav, A.K., Sharma, V., Malik, H., Chandel, S.S., Jun. 2017. Daily array yield prediction of grid-interactive photovoltaic plant using relief attribute evaluator based Radial Basis Function Neural Network. Renew. Sustain. Energy Rev.

Yaiici, W., Entchev, E., 2016. Adaptive Neuro-Fuzzy Inference System modelling for performance prediction of solar thermal energy system. Renew. Energy 86, 302–315.

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. Renew. Energy 106, 343–353.

Machine Learning in MATLAB - MATLAB & Simulink." [Online]. Available: https://www.mathworks.com/help/stats/machine-learning-in-matlab.html. [Accessed: 28-Sep-2017].