ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment

Victor H. Quej, Javier Almorox, Javier A. Arnaldo, Laurel Saito

ARTICLE INFO

ABSTRACT

Daily solar radiation is an important variable in many models. In this paper, the accuracy and performance of three soft computing techniques (i.e., adaptive neuro-fuzzy inference system (ANFIS), artificial neural network (ANN) and support vector machine (SVM) were assessed for predicting daily horizontal global solar radiation from measured meteorological variables in the Yucatán Peninsula, México. Model performance was assessed with statistical indicators such as root mean squared error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2). The performance assessment indicates that the SVM technique with requirements of daily maximum and minimum air temperature, extraterrestrial solar radiation and rainfall has better performance than the other techniques and may be a promising alternative to the usual approaches for predicting solar radiation.

1. Introduction

Solar radiation is the principal renewable energy source for many applications, including agricultural, ecological, hydrological and soil-vegetation-atmosphere transfer models (Liu et al., 2009). Solar researchers have developed many empirical correlations to estimate solar radiation using different meteorological, astronomical and geometrical factors (Besharat et al., 2013). Three major methods include models derived from satellite data, stochastic algorithms and empirical relationships. The satellite-derived method is promising for estimating solar radiation data over large regions, but it is relatively new and may suffer from a shortage of available data. Stochastic weather generators are useful for generating daily estimates from historic averages, but cannot be used for model validation if measured data are not available. Finally, the third approach involves a general practice of using empirical relationships to estimate solar radiation using other readily available meteorological data, such as sunshine duration, maximum and minimum temperatures, precipitation, and humidity. Previous studies have shown that estimates from solar radiation models using insolation performed better than those using other meteorological data. However, because insolation data are inaccessible in some locations, many temperature-based models have been proposed, and numerous evaluations and modifications have subsequently been made (Chen et al., 2011; Olatomiwa et al., 2015b).

Another set of techniques estimate solar radiation with soft-computing techniques. These techniques are within the framework of artificial intelligence that has received much attention for dealing with practical problems (Gopalakrishnan et al., 2011). Soft computing includes artificial neural networks (ANN), genetic algorithms (GAs), fuzzy logic (FL), adaptive neuro fuzzy inference systems (ANFIS), support vector machines (SVM) and data mining (DM). These methods offer advantages over conventional modeling, including the ability to handle large amounts of noisy data from dynamic and nonlinear systems, especially when the underlying physical processes are not fully understood (He et al., 2014).

Several soft computing techniques have been used in recent years to estimate global solar radiation, with ANFIS and ANN being the most popular. Mohammadi et al. (2015) applied ANFIS and SVM techniques to predict global solar radiation based on air temperatures in Bandar Abbas, located in the southern of Iran. The results showed that the SVM models outperform the ANFIS. Chen and Li (2014) investigated the ability of different SVM models in global solar radiation modeling for 15 synoptic stations with different climate conditions located in China. The results conclude that the SVM models could be used successfully in modeling global solar radiation. Piri et al. (2015) performed a comparative investigation between four sunshine durations based on empirical models and SVM models to estimate global solar radiation in two cities (Zahedan and Bojnurd) of Iran. Their

results indicated that all SVM models they used outperformed the empirical models. Olatomiwa et al. (2015a) compared the accuracy of SVM with ANFIS models and one empirical model for global solar radiation modeling in a semi-arid environment in Nigeria. They found that the SVM models performed better than the ANFIS and the empirical models. Ramedani et al. (2014) presented a radial basis SVM (SVM_rbf) model to predict global solar radiation over Teheran, Iran. The SVM_rbf prediction was compared with ANFIS and ANN models. The results demonstrated that the SVM_rbf have higher performance over the other models. Olatomiwa et al. (2015b) developed a novel method at the Chongqing meteorological station, China, by hybridizing SVMs with the Firefly Algorithm (FFA) to predict the monthly average global solar radiation using sunshine hours, maximum temperature and minimum temperature as inputs. The results indicate that the hybrid model can be used as an efficient machine learning technique for accurate estimation of global solar radiation.

Şenkal (2015) modeled global solar radiation using measured data of five stations in Turkey. He used an ANN technique with resilient propagation (RP) in the training phase with geographical and meteorological data as the input variables. The ANN technique had better results compared to other algorithms and input variables. Kumar et al. (2015) compared several models based on regression and ANN models to predict global solar radiation. In general, the ANN models had better results than the regression techniques. Sumithira et al. (2012) conducted a comparative study between ANFIS and other soft computing model techniques for estimating global solar radiation in Tamilnadu, India. The ANFIS model had promising results when compared to other models in the literature.

In the Yucatán Peninsula, several studies have been carried out to estimate solar radiation by using empirical or semi-empirical equations based on available meteorological parameters. However, there have been no studies that have estimated global solar radiation with soft computing techniques and where the precipitation factor in humid climates may play an important role in the modulation of the solar radiation produced by abundant cloud cover. In the present work, ANFIS, SVM and ANN techniques are used to predict daily global solar radiation in six cities of the Yucatán Peninsula. The objective is to evaluate the accuracy and performance of each technique using measured meteorological variables, including daily values of maximum and minimum air temperature, precipitation, and extraterrestrial solar radiation.

2. Methodology

2.1. Study area and climate dataset

This study was carried out using measured data set from six automatic weather stations located in the Yucatán Peninsula, México. The measured daily values used in this study are: minimum and maximum air temperatures, rainfall and global solar radiation. The measured data were provided by the Mexican National Meteorological Service (SMN; Servicio Meterológico Nacional) and Instituto Nacional de Investigaciones Forestales Agrícolas y Pecuarias (INIFAP) for six locations. Table 1 shows the geographical location of these stations, as well as the time periods of data used to evaluate and establish the

Table 1
Geographic information of the study weather stations.

Weather station	Longitude (°W)	Latitude (°N)	Altitude (m)	Time period
Calakmul	−89.8925	18.3650	28	2003–2014
Campeche	−90.5072	19.8361	11	2001–2014
Celestún	−90.3831	20.8580	10	2000–2014
Efraín Hernández	−89.8925	18.1935	90	2006–2014
Mérida	−89.6517	20.9463	18	2000–2006
Tantakín	−89.0472	20.0303	30	2003–2011

models.

The climate in the Yucatán Peninsula is classified as tropical savanna (Aw) according to the Köppen system (Koppen, 1936) with a rainy summer and dry winter. Long term annual average temperature ranges from 25.8 °C to 26.3 °C; rainfall occurs in summer and autumn with a gradient of minimum rainfall in the northwest (600 mm/year) to higher quantities toward the southeast (1400 mm/year).

2.2. Missing data analysis

Incorrect and missing records in the database are mainly related to malfunction of the measuring instruments. To cope with this problem, a quality control test was conducted by the following procedure (Jiang, 2009; Khorasanizadeh and Mohammadi, 2013).

a) To identify erroneous or outlier data of air temperature and precipitation, two methods were used, which may be distinguished as follows: (a) analysis of standardized residuals (Norusis, 2008) and (b) adjusted functional boxplots technique (Sun and Genton, 2012).

b) Incorrect records of the global solar radiation are revealed using daily clearness index (K_T) as an indicator. K_T is calculated as the ratio of measured daily global solar radiation intensity to the daily extraterrestrial solar radiation on a horizontal surface (Badescu, 2014). The upper and lower limits for K_T represent a clear sky and completely cloudy sky, respectively. When the daily clearness indices were outside the range of 0.015–1, the data were considered erroneous and deleted.

c) Missing or incorrect records for more than 5 consecutive days in a month were removed completely. In cases when less than five consecutive missing or incorrect records are found within a period of one month, the Piecewise Cubic Hermite Interpolating Polinomials (PCHIP) interpolation method (Fristch and Carlson, 1980; Kahaner et al., 1989) was used to fill in missing or replace incorrect values. Overall, deleted and missing data accounted for approximately 2% of the database.

2.3. Artificial neural network (ANN)

An artificial neural network is an abstract computational model that follows the behavior of the human brain (Haykin and Lippmann, 1994). ANN can be defined as "structures comprised of densely interconnected adaptive simple processing elements (called artificial neurons or nodes) that are capable of performing massively parallel computational data processing and knowledge representation" (Basheer and Hajmeer, 2000). Every neuron in the network computes a weighted by w_{ij} sum of its p input signal y_i, for $i=0,1,2,...n$ hidden layer and then applies a non-linear activation function to produce an output signal u_j. The form of this function is:

$$u_j = \sum_{i=0}^{n} w_{ij} y_i \tag{1}$$

The ANN method most commonly used for estimating solar radiation has been the multi-layer feed forward neural network (MLF) with the back propagation (BP) algorithm (Yadav and Chandel, 2014; Qazi et al., 2015; Rezrazi et al., 2015). This method is popular due to its ability to model problems that are not linearly separable.

The MLF consists of an input layer, an output layer and usually one or more hidden layers. In practice, only a three-layer feed forward neural network (FFNN) is usually necessary as seen in Fig. 1, where the first layer is the input layer representing input variables (i), the second layer is the hidden layer (j), and the third layer is the output layer (k). Each layer is interconnected by weights W_{ij} and W_{jk}, and every unit sums its inputs, adds a bias or threshold term to the sum and
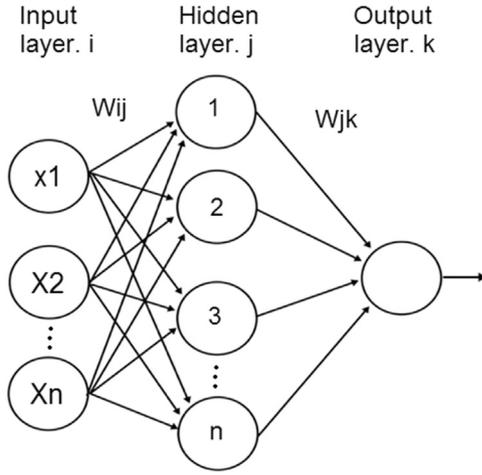
Fig. 1. A three-layer artificial neural network structure.

Fig. 2. Basic structure of ANFIS.

nonlinearity transforms the sum to produce an output. This nonlinear transformation is called the activation function of the node. The output layer nodes often have linear activations. In MLFs, the logistic sigmoid function (Eq. (2)) and linear function (Eq. (3)) are generally used in the hidden and output layer respectively (Rezrazi et al., 2015).

$$f(w) = 1/(1 + e^{-w}) \tag{2}$$

$$f(x) = x \tag{3}$$

where w is the weighted sum of the input and x is the input to the output layer.

The procedure for updating the synaptic weights is called back-propagation (BP). BP refers to the way error computed at the output side is propagated backward from the output to the hidden layer(s), and finally to the input layer (Esmaeelzadeh et al., 2014). The error is minimized across many training cycles called epochs. During each cycle, the network reaches a specified level of accuracy. Generally, the error estimator used here is the sum of the squared error (SSE). In conjunction with the BP procedure, the following algorithms can be used as a second training algorithm: Gradient Descent [Gradient Descent back-propagation algorithm (traingd), Gradient Descent with Momentum (traingdm), Resilence back-propagation (trainrp)], Conjugate Gradient algorithms [Scaled conjugate Gradient (trainscg), Conjugate Gradient back-propagation with Fletcher-Reeves Updates (traincgf), Conjugate Gradient back-propagation with Polak-Riebre Updates (traincgp)], and Quasi-Newton algorithms [Broyden-Fletcher-Goldfarb-Shanno (trainbfg), Levenberg–Marquardt back-propagation (trainlm)]. Selection of an appropriate training algorithm, transfer function and number of neurons in the hidden layer are fundamental characteristics of the ANN model. Each training algorithm has its own characteristics that must be adjusted according to a particular model.

2.4. Adaptive neuro fuzzy inference system (ANFIS)

ANFIS, proposed by Jang (1993), is a hybrid model composed of a fuzzy and artificial neural network, where the nodes in the different layers of a feed-forward network handle fuzzy parameters. This is equivalent to fuzzy inference systems (FIS) with distributed parameters. At its core, the technique splits the representation of prior knowledge into subsets in order to reduce the search space, and uses the back propagation algorithm to adjust the fuzzy parameters. The resulting system is an adaptive neural network functionally equivalent to a first-order Takagi-Sugeno inference system, where the input-output relationship is linear.

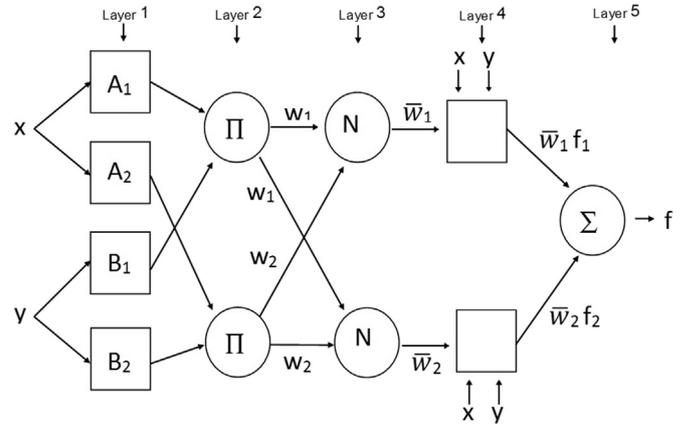In a first-order Sugeno system, a typical rule set with two fuzzy IF/THEN rules can be expressed as:

- Rule 1. If x is A_1 and y is B_1, then $f_1 = p_1 x + q_1 y + r_1$ (4)

- Rule 2. If x is A_2 and y is B_2, then $f_2 = p_2 x + q_2 y + r_2$ (5)

where x and y are the crisp inputs to node i, A_i and B_i are the fuzzy sets in the antecedent, f_i is the output within the fuzzy region specified by the fuzzy rule; and p_i, q_i and r_i are the design parameters that are determined during the training process.

2.4.1. ANFIS structure for two inputs

ANFIS structure consists of five layers, namely: fuzzy layer, product layer, normalized layer, de-fuzzy layer and total output layer. The ANFIS structure for two inputs, one output and two rules is shown in Fig. 2. Each layer performs a particular task in the fuzzy inference system. For identification, the adaptive nodes are represented by squares, and fixed nodes are represented by circles.

Abdulshahed et al. (2015) describes the ANFIS structure as follows:
Layer 1. (Fuzzy layer): Each node i in this layer (denoted with a square) represents a node function:

$$O_{1,i} = \mu_{Ai}(x), \quad \text{for } i = 1, 2, ..., n \qquad O_{1,i} = \mu_{Bi-2}(y), \quad \text{for } i = 1, 2, ..., n \tag{6}$$

where x (or y) is the input to node i, and A_i or (B_{i-2}) is the linguistic label (small, large, etc.) characterized by appropriate membership functions (MF's) $\mu_{Ai}(x)$ and $\mu_{Bi-2}(y)$. The MF's can take any shape or function such as triangular, trapezoidal, Gaussian, or bell-shaped

Table 2
The basic MF's used in the study.

Name of MFs	Equation	
Triangular MF	$\mu_{Ai}(x) = max\left\{min\left(\dfrac{x-a}{b-a}, \dfrac{c-x}{c-b}\right),0\right\}$	(7)
Trapezoidal MF	$\mu_{Ai}(x) = max\left\{min\left(\dfrac{x-a}{b-a},1,\dfrac{d-x}{d-c}\right),0\right\}$	(8)
Gaussian MF	$\mu_{Ai}(x) = e^{\frac{-(x-c)^2}{2\sigma^2}}$	(9)
Bell-Shaped MF	$\mu_{Ai}(x) = \dfrac{1}{1+\left|\frac{x-c}{a}\right|^{2b}}$	(10)

$\{a, b, c, d\}$ is the parameter set that changes the shapes of the MFs with maximum 1 and minimum 0.

(Table 2). The parameters of the MF's in this layer are named as premises.

Layer 2. (Product layer): This layer has the circle nodes represented with Π in Fig. 1. In this layer, the rule operator (AND/OR) is applied to get one output that represents the results of the antecedent for a fuzzy rule that multiplies the incoming signals, such as:

$$O_{2,i} = w_i = \mu_{Ai}(x) * \mu_{Bi}(y), \quad for\ i = 1, 2. \tag{11}$$

where $O_{2,i}$ is the output of the layer, and the output signal w_i represents the firing strength of the rule.

Layer 3. (Normalized layer): The nodes in this layer are denoted with N and they calculate the ratio of the i th rule's firing strength to the sum of firing strengths of all rules by:

$$O_{3,i} = \overline{w} = \frac{w_i}{w_1 + w_2}, \quad for\ i = 1, 2. \tag{12}$$

where the $O_{3,i}$ is the output of Layer 3. The quantity $\overline{w}$ is known as the normalized firing strength.

Layer 4. (De-fuzzy layer): The nodes in this layer are denoted with a square and they calculate the weighted output of each linear function as follows:

$$O_{4,i} = \overline{w_i}. f_i = \overline{w_i}(p_i x + q_i y + r_i), \quad for\ i = 1, 2. \tag{13}$$

where $\overline{w}$ is the output of layer 3, and $\{p_i, \quad q_i, r_i\}$, are the coefficients of a linear combination in the Sugeno inference system. These parameters of this layer are referred to as the consequent parameters.

Layer 5. (Total output layer): The single node denoted with an Σ computes overall output as follows:

$$O_{5,i} = \sum \overline{w_i}. f_i = \frac{\sum_i w_i. f_i}{\sum_i w_i} = f\ out = Estimated\ overall\ output \tag{14}$$

ANFIS uses a hybrid learning algorithm for estimation of the premise and consequent parameters. The hybrid learning algorithm procedure estimates the consequent parameters in a forward pass and the premise parameters in a backward pass. In the forward phase, the information propagates forward until layer 4, where the consequent parameters are optimized by a least square regression algorithm. In the backward phase, the error signals propagate backwards and the premise parameters are updated by a gradient descent (GD) algorithm (Jang et al., 1997). This error measure is usually defined by the sum of the squared difference between measured and modeled values and is minimized to a desired value.

The final overall output in Fig. 1 can be rewritten as:

$$f\ out = (\overline{w}_1. x)p_1 + (\overline{w}_1. y)q_1 y + (\overline{w}_1)r_1 + (\overline{w}_2. x)p_2 + (\overline{w}_2. y)q_2 y + (\overline{w}_2)r_2 \tag{15}$$

where x and y are the input parameters of the model, $\overline{w}_1$, $\overline{w}_2$ are the normalized firing strengths of fuzzy rules and $(p_1; q_1; r_1; p_2; q_2; r_2)$ are the consequent parameters.

2.5. Support vector machines (SVM)

A support vector machine (SVM) was introduced by Vapnik (2013) and is a supervised learning technique from the field of machine learning theory and structural risk minimization, applicable to both classification and regression. SVMs are essentially a sub-discipline of machine learning. SVMs are derived from the concept of structural risk minimization hypothesis to minimize both empirical risk and the confidence interval of the learning machine, which in turn helps to produce good generalization capability.

In addition to their solid mathematical foundation in statistical learning theory, SVMs have demonstrated highly competitive performance in numerous real-world applications. Initially developed for solving classification problems, SVM techniques can also be successfully applied in regression problems, usually called SVR (support vector regression). A regression is estimated by using SVM for a data set

$\{(x_i, y_i)\}N/i = 1$, where x_i is the input vector, y_i is the output value and N is the total number of data sets by mapping x into a feature space via a nonlinear function $\varphi(x)$, and then finding a regression function as follows:

$$f(x) = \omega. \varphi(x) + b \tag{16}$$

which can best approximate the actual output y with an error tolerance ε, where ω and b are regression function parameters known as weight vector and bias value, respectively. φ is known as a nonlinear mapping function.

The coefficients b and ω are calculated by minimizing the following regularized risk function:

$$R(C) = C \sum_i^N L_\varepsilon(f(x_i), y_i)\frac{1}{2}\|\omega\|^2 \tag{17}$$

where the term $\frac{1}{2}\|\omega\|^2$ improves the generalization of the SVM by regulating the degree of model complexity, which denotes the Euclidean norm. C is a positive trade-off parameter that determines the degree of the empirical error in the optimization problem that is chosen by the user. The most important difference with respect to classic regression is the use of a novel loss function (ε). This is the Vapnik's linear loss function with ε-insensitivity (tube size of SVM) zone defined as:

$$L_\varepsilon(f(x_i), y_i) = \begin{cases} 0 & for |f(x_i) - y_i| \leq \varepsilon \\ |f(x) - y| - \varepsilon & otherwise \end{cases} \tag{18}$$

Thus, the loss is equal to 0 if the difference between the predicted and the measured value is less than ε. If the predicted value is within the tube, the loss error is zero. For all other predicted points outside the tube, the loss equals the magnitude of the difference between the predicted value and the radius ε of the tube. To avoid outliers, slack variables ξ and ξ* are introduced for measurements "above" and "below" of the ε tube respectively. Both slack variables are positive values.

The risk is minimized as follows:
minimize

$$R(\xi, \xi^*, \omega, b) = \frac{1}{2}\|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \tag{19}$$

subjected to $\begin{cases} y_i - \omega\phi(x_i) - b_i \leq \varepsilon + \xi_i \\ \omega\phi(x_i) + b_i - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \quad \xi_i^* \geq 0 \end{cases}$

where the $C \sum_{i=1}^n (\xi_i + \xi_i^*)$ controls the degree of empirical risk.

To solve the optimization problem, Lagrange multipliers α and α* are added to the condition equations, and the equation can be written in its dual form:

$$R(\alpha, \alpha^*) = \sum_{i=1}^n y_i(\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^n (\alpha_i - \alpha_i^*)$$
$$- \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j) \tag{20}$$

with constraints:

$$\sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \qquad 0 \leq \alpha_i \leq C \quad , \quad 0 \leq \alpha_i^* \leq \quad Ci = 1, 2, \ldots\ldots, n$$

where α_i and α_{i*} are Lagrange multipliers to be solved, and $K(x_i, x_j)$ is called a kernel function and is calculated by $K(x_i, x_j) = \varphi(x_i).\varphi(x_j)$ on the feature space. The kernel allows SVMs to form nonlinear boundaries; in other words, it provides the SVM the ability to model complicated separating hyperplanes.

After calculating Lagrange multipliers, an optimal desired weights vector of the regression hyperplane is found as follows:

$$\omega = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \quad \varphi(x) \tag{21}$$

and Eq. (16) can be rewritten as follows:

$$f(x, \alpha, \alpha^*) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \tag{22}$$

where n is the number of support vectors, $(\alpha_i - \alpha_i^*)$ are their Lagrange multipliers, the term $K(x_i, x_j)$ is the kernel function in the input space and the bias b is calculated from training samples.

In general, mathematically, a basic function for the statistical learning process in SVM is

$$y = f(x) = \sum_{i=1}^{M} \alpha_i \varphi(x) = \mathrm{w}\varphi(x) \tag{23}$$

where the output is a linearly weighted sum of M. The nonlinear transformation is carried out by $\varphi(x)$.

The decision function of SVM is represented as

$$y = f(x) = \left\{ \sum_{i=1}^{N} \alpha_i K(x_i, x) \right\} - b \tag{24}$$

where K is the kernel function, α_i and b are parameters, N is the number of training data, x_i are vectors used in the training process, and x is the independent vector. The parameters α_i and b are derived by maximizing their objective functions. A suitable choice of kernel allows the data to become separable in the feature space despite being nonseparable in the original input space. The four basic kernel functions are showed in the Table 3.

2.6. ANN, ANFIS and SVM models' implementation for estimating daily global solar radiation

We conducted computer simulations of ANFIS and ANN techniques using two script files written in MATLAB 2015b software version. The SVM technique was applied using an open source software package named LIBSVM 3.2 (Chang and Lin, 2001) that included the sequential minimal optimization (SMO) algorithm that was implemented using R (RDevelopment, 2012; Meyer and Wien, 2014).

For all three techniques, the daily global solar radiation *(H)* variable was defined as the dependent variable. The daily minimum temperature (T_{min}), daily maximum temperature (T_{max}), rainy days [(W) defined as a binary variable (i.e., rainfall > 0, W=1; rainfall=0, W=0)], and daily extraterrestrial radiation (H_0), were used as input to the ANFIS, ANN and SVM models for *H* modeling. The H_O value was calculated as a function of the day of year, site latitude and solar angle according to the equation proposed by Allen et al. (1998).

Table 3
Different kernel functions used in the study.

Name of kernel function	Equation	
Radial basis function(RBF)	$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \quad \gamma > 0$	(25)
Polynomial	$K(x_i, x_j) = (\gamma x_i \cdot x_j + r)^d, \; \gamma > 0$	(26)
Linear	$K(x_i, x_j) = x_i \cdot \quad x_j$	(27)
Sigmoid	$K(x_i, x_j) = \tanh(\gamma x_i \cdot x_j + r)$	(28)

x_i and x_j are vectors in the input space, d, r, and γ are the kernel parameters.

Two scenarios were evaluated in this study by considering the actual T_{min}, T_{max}, W and H_0: (1) T_{min}, T_{max}, W, and H_0 were used for the SVM1, ANFIS1 and ANN1 models; (2) T_{min}, T_{max} and H_0 were used for the SVM2, ANFIS2 and ANN2 models. The rainfall parameter (W) was included in scenario 1, but not in scenario 2 in order to evaluate its effect on the models. The *H* measured was used as target values for all models.

In order to preclude the possibility of overfitting, all data in the measured dataset were normalized to the range from zero to one. Moreover, to ensure the representativeness of the dataset, the database was split into two subsets randomly, using 70% for training, and the remaining 30% was used to validate the model. The training dataset was used to train all the models, while validation dataset was used to verify the performance of the trained models.

2.6.1. ANN model architecture

This investigation used a three-layer or FFNN for *H* simulation, where the first layer is the input layer representing input variables, the second layer is the hidden layer, and the third layer is the output layer. This topology has proved its ability in modeling many real-world functional problems (Ata, 2015; Piotrowski et al., 2015; Antonopoulos and Antonopoulos, 2017). The selection of hidden neurons is the tricky part in ANN modeling, as it relates to the complexity of the system being modeled. In this study, the optimum numbers of neurons in the hidden layer was determined by a simple trial and errors process. A range of 2–80 neurons were evaluated until a minimum acceptable error was achieved between the predicted and observed output. The next step was to choose the transfer functions for the hidden and output layers. In this study, the logistic sigmoid (Logsig) Eq. (2) was used as the transfer function in the hidden layer and the linear transfer function (Pureline) Eq. (3) was applied in the output layer. Table 4 shows the ANN features for each model and station.

The neural networks were trained using backpropagation (BP) with the Levenberg-Marquardt (LM) algorithm, which is considered to be one of the faster and more accurate algorithms. It combines the speed of the Newton algorithm with the stability of the steepest decent method (Wilamowski et al., 1999). The LM algorithm uses Newton's method to calculate Jacobian matrices without computing the hessian matrices. This makes the LM algorithm have a faster convergence with minimal error.

2.6.2. ANFIS model architecture

The ANFIS technique was used for *H* modeling, in which a set of parameters in ANFIS were identified through a hybrid learning rule

Table 4
Features of the ANN models.

Station/Model ID	Activation functions		Number of hidden layer units
	Hidden layer	Output layer	
Campeche			
ANN1	Logsig	Pureline	13
ANN2	Logsig	Pureline	12
Calakmul			
ANN1	Logsig	Pureline	37
ANN2	Logsig	Pureline	26
Celestún			
ANN1	Logsig	Pureline	27
ANN2	Logsig	Pureline	52
Efraín Hernández			
ANN1	Logsig	Pureline	19
ANN2	Logsig	Pureline	24
Tantakín			
ANN1	Logsig	Pureline	18
ANN2	Logsig	Pureline	32

that combined the back-propagation gradient descent and a least square method. The choice of appropriate membership function (MFs) and a given number of fuzzy IF-THEN rules for each variable is essential to achieve good estimates during training. In the present study, the bell-shaped function (Eq. (10)) better represented the variability of data points and was used for the MFs. The bell-shaped function was chosen by an arbitrary process of testing the four types of MFs cited in Table 2. The "IF" part of fuzzy rules is related to input variables to MF's and is called premise or antecedent; the "THEN" part, which is related to the output variable for the MFs, is called the conclusion or consequent part of the rule (Cobaner, 2011). The number of fuzzy IF-THEN rules in a system depends on the number of MFs and the number of input variables. So, two IF-THEN rules were employed for each input variable, since the best results were obtained with this value as determined by iterative processes. The grid partition method on the data was used to generate the Takagi-Sugeno fuzzy inference system (FIS) structure (Cobaner, 2011; Shiri et al., 2012). In the grid partition method, the input is divided into different spaces using the axis–paralleled method in which each input represents a fuzzy MF.

2.6.3. SVM model architecture

For H modeling, the SVMs were trained using the LIBSVM software with $\varepsilon - SV$ regression. There are four major kernels predominantly used in SVMs such as linear, polynomial, radial basis and sigmoid which are shown in Table 3. In this study, the radial basis function (RBF) kernel was chosen as the best after a process of trial and error experiments. The SVM hypothesis suggested that the performance of SVM depends on the slack parameter (ε), the cost factor (C), and kernel parameter (γ) while using RBF as the kernel function. The C assigns a penalty for the number of vectors falling between the two hyperplanes in the hypothesis. The C suggests that if the data are of good quality, the distance between the two hyperplanes is narrower. If the data are noisy it is preferable to have a smaller value of C that will not penalize the vectors. So it is important to find the optimum C value for SVM modeling. In this study, these parameters (ε, C, γ) were optimized by using the genetic algorithm (GA) with five-fold cross-validation (CV), varying the ε values between ε=0.002 to ε=2, and the varying the C parameters between C=0.0001 to C=10. In addition, the parameter γ of the kernel function is adjusted via GA, varying the γ values between γ=0.0001 to γ=2. GA is a search-based optimization technique based on the principles of genetics and natural selection. GA combines selection operators, generation, crossover and mutation for identifying the best solution for the optimization problem (Antonanzas-Torres et al., 2015; Zhang et al., 2015).

The optimized parameters to train the SVM models are presented in Table 5.

2.7. Model performance evaluation

The performance of the models was evaluated using the following statistical tests: root mean squared error (RMSE; Eq. (29)), mean absolute error (MAE; Eq. (30)) and coefficient of determination (R^2; Eq. (31)).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(H_c - H_m)^2}\ \text{MJ m}^{-2}\text{day}^{-1} \tag{29}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(|H_c - H_m|)\ \text{MJ m}^{-2}\text{day}^{-1} \tag{30}$$

$$R^2 = \frac{[\sum_{i=1}^{n}(H_c - H_{c,avg})(H_m - H_{m,avg})]^2}{\sum_{i=1}^{n}(H_c - H_{c,avg})^2\sum_{i=1}^{n}(H_m - H_{m,avg})^2}\ \text{Unitless} \tag{31}$$

where n represents the total number of evaluating data, H_m, H_c and $H_{m,avg}$ are the measured, calculated and mean of measured H values, respectively.

Table 5
Optimum SVM parameters obtained by the GA-based.

Station/Model ID	Optimum values		
	C	γ	ε
Campeche			
SVM1	1.423	0.532	0.407
SVM2	2.779	0.147	0.495
Calakmul			
SVM1	1.162	0.358	0.342
SVM2	2.293	0.465	0.474
Celestún			
SVM1	1.084	0.409	0.533
SVM2	3.662	0.518	0.519
Efraín Hernández			
SVM1	3.675	0.235	0.287
SVM2	3.295	0.466	0.512
Mérida			
SVM1	6.514	0.046	0.437
SVM2	1.308	0.702	0.423
Tantakín			
SVM1	9.747	0.132	0.453
SVM2	5.826	0.128	0.275

Smaller values of RMSE and MAE imply closer approximation of measured values by the models. Larger R^2 values indicate a stronger matching of trends in the measured data by the model results. The RMSE and MAE are widely used basic metrics for assessing the performance of prediction models (Teke et al., 2015).

3. Results

The SVM, ANFIS and ANN computing techniques were used for H prediction in six meteorological stations located in the Yucatán Peninsula, México. The statistical performance measures for each model are presented in Table 6 for both training and testing data sets. In the first scenario, when rainfall data are included, the SVM1 model achieved the best performance for all locations in the testing phase according to mean performance statistics (R^2=0.689, RMSE=2.678 MJ m^{-2} day^{-1} and MAE=1.973 MJ m^{-2} day^{-1}), followed by ANN1 (R^2=0.652, RMSE=2.775 MJ m^{-2} day^{-1} and MAE=2.150 MJ m^{-2} day^{-1}). The ANFIS1 model had a similar performance to the ANN1 (R^2=0.645, RMSE=2.801 MJ m^{-2} day^{-1}, and MAE=2.153 MJ m^{-2} day^{-1}).

In the second scenario, the SVM2 model with inputs of T_{min}, T_{max}, and H_0 performed the best for all locations according to overall mean errors (R^2=0.624, RMSE=2.877 MJ m^{-2} day^{-1}, MAE=2.203 MJ m^{-2} day^{-1}), and the ANN2 and ANFIS2 had similar performance measures (R^2=0.596, RMSE=3.023 MJ m^{-2} day^{-1}, MAE=2.352 MJ m^{-2} day^{-1} for ANN2 and R^2=0.587, RMSE=3.052 MJ m^{-2} day^{-1}, MAE=2.365 MJ m^{-2} day^{-1} for ANFIS2). In all cases, the first scenario models performed better than the second scenario models.

Estimated daily solar global radiation values by the SMV1 models for the testing phase are plotted against the measured values for the six meteorological stations in Fig. 3. Table 6 also shows the statistical results for the data set during the training phase. The superiority of the SVM technique over the ANFIS and ANN techniques is evident from the performance measures.

The MAE and RMSE values found in this study are similar to those found by other authors (Tabari et al., 2012; Mohammadi et al., 2015). The results demonstrate the ability of SVM1 models to adapt to existing conditions in Yucatán.

Table 6
The R^2, RMSE and MAE statistics for each model during training and testing phases for six locations in Yucatán, Peninsula, México.

| Station/Model ID	Training data			Testing data		
	R^2	RMSE (MJ m^{-2}d^{-1})	MAE (MJ m^{-2}d^{-1})	R^2	RMSE (MJ m^{-2}d^{-1})	MAE (MJ m^{-2}d^{-1})
Campeche						
SVM1	0.674	2.701	1.999	0.663	2.786	2.049
SVM2	0.573	3.089	2.295	0.562	3.171	2.326
ANFIS1	0.644	2.851	2.046	0.652	2.769	2.057
ANFIS2	0.571	3.132	2.234	0.539	3.188	2.364
ANN1	0.642	2.704	2.112	0.655	2.758	2.064
ANN2	0.562	2.988	2.366	0.535	3.206	2.419
Calakmul						
SVM1	0.697	2.158	1.650	0.698	2.142	1.668
SVM2	0.679	2.128	1.706	0.643	2.330	1.809
ANFIS1	0.640	2.271	1.001	0.636	2.423	1.888
ANFIS2	0.638	2.361	1.799	0.597	2.549	1.999
ANN1	0.678	2.222	1.003	0.644	2.406	1.871
ANN2	0.656	2.300	1.761	0.623	2.477	1.952
Celestún						
SVM1	0.666	2.604	1.001	0.635	2.827	2.164
SVM2	0.552	3.223	2.275	0.517	3.320	2.493
ANFIS1	0.648	2.656	1.979	0.513	3.381	2.715
ANFIS2	0.560	3.133	2.336	0.483	3.668	3.017
ANN1	0.650	2.804	2.010	0.526	3.344	2.706
ANN2	0.546	3.036	2.371	0.497	3.614	2.982
Efraín Hernández						
SVM1	0.709	2.745	2.126	0.732	2.697	2.103
SVM2	0.654	2.992	2.345	0.701	2.972	2.350
ANFIS1	0.688	2.936	2.368	0.718	2.884	2.200
ANFIS2	0.643	3.100	2.510	0.672	3.081	2.399
ANN1	0.694	2.688	2.185	0.727	2.824	2.189
ANN2	0.649	2.904	2.364	0.671	3.105	2.422
Mérida						
SVM1	0.737	2.509	1.918	0.704	2.535	1.933
SVM2	0.661	2.837	2.169	0.653	2.914	2.232
ANFIS1	0.697	2.704	2.019	0.689	2.561	1.935
ANFIS2	0.608	3.097	2.295	0.617	2.849	2.150
ANN1	0.715	2.499	2.000	0.697	2.526	1.952
ANN2	0.670	2.807	2.139	0.625	2.818	2.113
Tantakín						
SVM1	0.719	2.418	1.883	0.700	2.482	1.918
SVM2	0.691	2.538	1.974	0.667	2.556	2.010
ANFIS1	0.651	2.810	2.234	0.660	2.788	2.121
ANFIS2	0.639	2.894	2.307	0.616	2.976	2.259
ANN1	0.705	2.390	1.928	0.661	2.795	2.116
ANN2	0.681	2.487	2.019	0.626	2.921	2.222
Overall average						
SVM1	0.700	2.523	1.763	0.689	2.578	1.973
SVM2	0.635	2.801	2.127	0.624	2.877	2.203
ANFIS1	0.661	2.705	1.941	0.645	2.801	2.153
ANFIS2	0.610	2.953	2.247	0.587	3.052	2.365
ANN1	0.680	2.551	1.873	0.652	2.775	2.150
ANN2	0.627	2.753	2.170	0.596	3.023	2.352

Thus, according to the statistical indicators, the SVM approach performed better than the ANN and ANFIS models. In addition, the rainfall input parameter had a positive effect on model performance in estimating daily solar radiation for this tropical area.

4. Discussion and conclusions

In this study, we compared the performance of SVM, ANFIS and ANN techniques to predict daily global solar radiation in Yucatán, Mexico. T_{min}, T_{max}, W and H_0 were used as explanatory variables. Three statistical indicators (R^2, RMSE and MAE) were used to evaluate model performance.

The results indicated that the SVM technique with RBF kernel is superior to other approaches for estimating global solar radiation in Yucatán, Mexico. Therefore, it can be used successfully to estimate daily solar radiation in Mexican sub-humid tropics environments.

Moreover, the results are generally in accordance with the recent published studies, which mostly concluded that SVM technique for estimating solar radiation are superior to the ANN and ANFIS methods. The SVM technique is able to select an optimal hyperplane in the training process as its support vector and remove the nonsupport vector from the model, resulting in fast training and low computational cost. This makes the model cope well with noisy conditions (Tabari et al., 2012), such as the
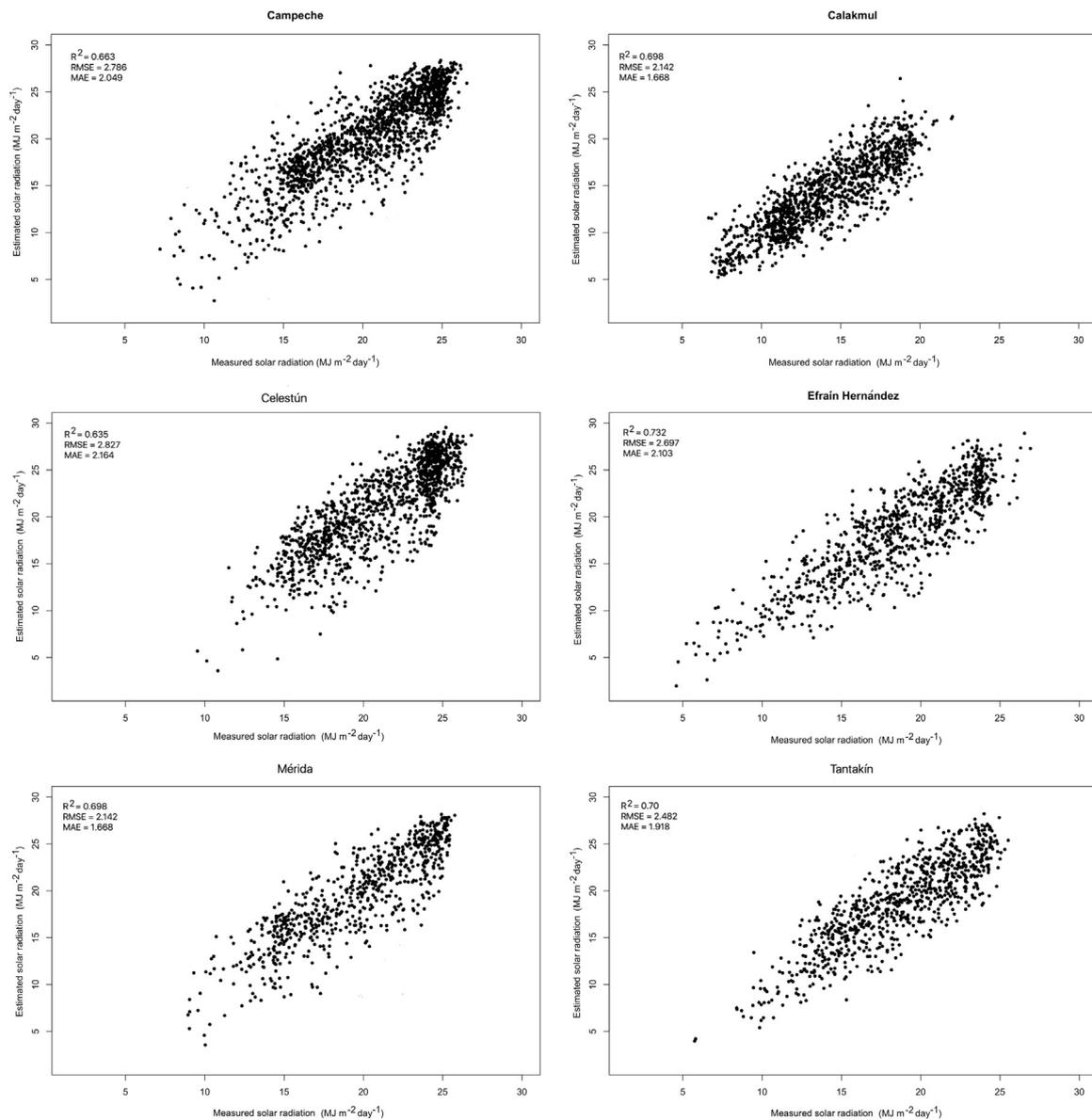
Fig. 3. Regression analysis plot for the SVM1 model between measured and estimated values of daily global solar radiation during test phase.

local events of rain or cloudiness. In addition, the SVM solves a convex optimization problem which will always converge to a global minimum.

The low values of R^2 are likely due to the type of sub-humid environment with abundant rainfall or cloudiness events that occur mainly during the summer and autumn. Other studies with higher R^2 values have occurred mainly in arid and semi-arid climates where annual precipitation is less than 300 mm (Tabari et al., 2012; Antonanzas-Torres et al., 2015; Mohammadi et al., 2015; Belaid and Mellit, 2016), but few, if any, studies have been conducted in wetter climates. The effect of climate is supported by the improved performance of the models when information of rainfall was added as compared to the models based solely on temperatures and extraterrestrial radiation. ANFIS and ANN techniques had similar results, but did not perform as well as the SVM technique. Thus, the results of this study indicate that SVM increases the prediction accuracies of global solar radiation estimations in tropical warm sub-humid climates such as the Yucatán in México, especially when precipitation is included in the model.

References

Liu, X., Mei, X., Li, Y., Zhang, Y., Wang, Q., Jensen, J.R., Porter, J.R., 2009. Calibration of the Ångström–Prescott coefficients (a, b) under different time scales and their impacts in estimating global solar radiation in the Yellow River basin. Agric. For. Meteorol. 149, 697–710.

Besharat, F., Dehghan, A.A., Faghih, A.R., 2013. Empirical models for estimating global solar radiation: a review and case study. Renew. Sustain. Energy Rev. 21, 798–821.

Chen, J.-L., Liu, H.-B., Wu, W., Xie, D.-T., 2011. Estimation of monthly solar radiation from measured temperatures using support vector machines–a case study. Renew. Energy 36, 413–420.

Olatomiwa, L., Mekhilef, S., Shamshirband, S., Mohammadi, K., Petković, D., Sudheer, C., 2015a. A support vector machine–firefly algorithm-based model for global solar radiation prediction. Sol. Energy 115, 632–644. http://dx.doi.org/10.1016/j.solener.2015.03.015.

Gopalakrishnan, K., Khaitan, S.K., Kalogirou, S., 2011. Soft computing in green and renewable energy systems. Springer.

He, Z., Wen, X., Liu, H., Du, J., 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. J. Hydrol. 509, 379–386.

Mohammadi, K., Shamshirband, S., Danesh, A.S., Abdullah, M.S., Zamani, M., 2015. Temperature-based estimation of global solar radiation using soft computing methodologies. Theor. Appl. Climatol.. http://dx.doi.org/10.1007/s00704-015-1487-x.

Chen, J.-L., Li, G.-S., 2014. Evaluation of support vector machine for estimation of solar radiation from measured meteorological variables. Theor. Appl. Climatol. 115, 627–638.

Piri, J., Shamshirband, S., Petković, D., Tong, C.W., ur Rehman, M.H., 2015. Prediction of the solar radiation on the Earth using support vector regression technique. Infrared Phys. Technol. 68, 179–185. http://dx.doi.org/10.1016/j.infrared.2014.12.006.

Olatomiwa, L., Mekhilef, S., Shamshirband, S., Petkovic, D., 2015b. Potential of support vector regression for solar radiation prediction in Nigeria. Nat. Hazards 77,

1055–1068.

Ramedani, Z., Omid, M., Keyhani, A., Shamshirband, S., Khoshnevisan, B., 2014. Potential of radial basis function based support vector regression for global solar radiation prediction. Renew. Sustain. Energy Rev. 39, 1005–1011. http://dx.doi.org/10.1016/j.rser.2014.07.108.

Şenkal, O., 2015. Solar radiation and precipitable water modeling for Turkey using artificial neural networks. Meteorol. Atmos. Phys., 1–8.

Kumar, R., Aggarwal, R.K., Sharma, J.D., 2015. Comparison of regression and artificial neural network models for estimation of global solar radiations. Renew. Sustain. Energy Rev. 52, 1294–1299. http://dx.doi.org/10.1016/j.rser.2015.08.021.

Sumithira, T.R., Kumar, A.N., Rameshkumar, R., 2012. An adaptive neuro-fuzzy inference system (ANFIS) based prediction of solar radiation. J. Appl. Sci. Res 8, 346–351.

Koppen, W., 1936. Das geographische System der Klimate. vol. 1, Part C. Koppen, W., R. geiger.handb. der Klimatologie.Gebr.Bornträger, Berlin.

Jiang, Y., 2009. Estimation of monthly mean daily diffuse radiation in China. Appl. Energy 86, 1458–1464.

Khorasanizadeh, H., Mohammadi, K., 2013. Prediction of daily global solar radiation by day of the year in four cities located in the sunny regions of Iran. Energy Convers. Manag. 76, 385–392. http://dx.doi.org/10.1016/j.enconman.2013.07.073.

Norusis, M.J., 2008. SPSS 16.0 Guide to Data Analysis Value Package (includes SPSS 16.0 Student Version for Windows).

Sun, Y., Genton, M.G., 2012. Adjusted functional boxplots for spatio-temporal data visualization and outlier detection. Environmetrics 23, 54–64. http://dx.doi.org/10.1002/env.1136.

Badescu, V., 2014. Modeling solar radiation at the earth's surface.

Fritsch, F.N., Carlson, R.E., 1980. Monotone piecewise cubic interpolation. SIAM J. Numer. Anal. 17, 238–246.

Kahaner, D., Moler, C., Nash, S., 1989. Numerical methods and software 1989. Prentice Hall, Englewood Cliffs, 1.

Jang, J.-S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. Syst. Man Cybern. IEEE Trans. 23, 665–685. http://dx.doi.org/10.1109/21.256541.

Abdulshahed, A.M., Longstaff, A.P., Fletcher, S., 2015. The application of ANFIS prediction models for thermal error compensation on CNC machine tools. Appl. Soft Comput. 27, 158–168.

Jang, J.-S.R., Sun, C.-T., Mizutani, E., 1997. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice Hall, New Jersey.

Vapnik, V., 2013. The Nature of Statistical Learning Theory. Springer Science & Business Media.

Haykin, S., Lippmann, R., 1994. Neural networks, A Comprehensive Foundation. Int. J. Neural Syst. 5, 363–364.

Basheer, I.A., Hajmeer, M., 2000. Artificial neural networks: fundamentals, computing, design, and application. J. Microbiol. Methods 43, 3–31. http://dx.doi.org/10.1016/S0167-7012(00)00201-3.

Yadav, A.K., Chandel, S.S., 2014. Solar radiation prediction using artificial neural network techniques: a review. Renew. Sustain. Energy Rev. 33, 772–781. http://dx.doi.org/10.1016/j.rser.2013.08.055.

Qazi, A., Fayaz, H., Wadi, A., Raj, R.G., Rahim, N.A., Khan, W.A., 2015. The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review. J. Clean. Prod. 104, 1–12. http://dx.doi.org/10.1016/j.jclepro.2015.04.041.

Rezrazi, A., Hanini, S., Laidi, M., 2015. An optimisation methodology of artificial neural network models for predicting solar radiation: a case study. Theor. Appl. Climatol. 123, 1–15. http://dx.doi.org/10.1007/s00704-015-1398-x.

Esmaeelzadeh, S.R., Adib, A., Alahdin, S., 2014. Long-term streamflow forecasts by adaptive neuro-fuzzy inference system using satellite images and K-fold cross-validation (case study: Dez, Iran). KSCE J. Civ. Eng., 1–9.

Chang, C., Lin, C., 2001. {LIBSVM}: A Library for Support Vector Machines (Version 2.3).

RDevelopment, C., 2012. TEAM 2009: R: A Language and Environment for Statistical Computing. Vienna, Austria.Internet: ⟨http://www.R-project.org⟩.

Meyer, D., Wien, F.H.T., 2014. Support vector machines. Interface libsvm Packag., e1071.

Ata, R., 2015. Artificial neural networks applications in wind energy systems: a review. Renew. Sustain. Energy Rev. 49, 534–562. http://dx.doi.org/10.1016/j.rser.2015.04.166.

Piotrowski, A.P., Napiorkowski, M.J., Napiorkowski, J.J., Osuch, M., 2015. Comparing various artificial neural network types for water temperature prediction in rivers. J. Hydrol. 529, 302–315. http://dx.doi.org/10.1016/j.jhydrol.2015.07.044.

Antonopoulos, V.Z., Antonopoulos, A.V., 2017. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. Comput. Electron. Agric. 132, 86–96. http://dx.doi.org/10.1016/j.compag.2016.11.011.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration-guidelines for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56. FAO, Rome 300, 6541.

Wilamowski, B.M., Chen, Y., Malinowski, A., 1999. Efficient algorithm for training neural networks with one hidden layer, In: IJCNN. Proceedings of the 1999 International Joint Conference. Citeseer, pp. 1725–1728.

Cobaner, M., 2011. Evapotranspiration estimation by two different neuro-fuzzy inference systems. J. Hydrol. 398, 292–302. http://dx.doi.org/10.1016/j.jhydrol.2010.12.030.

Shiri, J., Kisi, O., Landeras, G., Javier Lopez, J., Nazemi, A.H., Stuyt, L.C.P.M., 2012. Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). J. Hydrol. 414, 302–316. http://dx.doi.org/10.1016/j.jhydrol.2011.11.004.

Zhang, L., Juyang, L.E.I., Qilin, Z., Yudong, W., 2015. Using genetic algorithm to optimize parameters of support vector machine and its application in material fatigue life prediction. Adv. Nat. Sci., 8.

Antonanzas-Torres, F., Urraca, R., Antonanzas, J., Fernandez-Ceniceros, J., Martinez-De-Pison, F.J., 2015. Generation of daily global solar irradiation with support vector machines for regression. Energy Convers. Manag. 96, 277–286. http://dx.doi.org/10.1016/j.enconman.2015.02.086.

Teke, A., Yıldırım, H.B., Çelik, Ö., 2015. Evaluation and performance comparison of different models for the estimation of solar radiation. Renew. Sustain. Energy Rev. 50, 1097–1107.

Tabari, H., Kisi, O., Ezani, A., Hosseinzadeh Talaee, P., 2012. SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. J. Hydrol. 444–445, 78–89. http://dx.doi.org/10.1016/j.jhydrol.2012.04.007.

Belaid, S., Mellit, A., 2016. Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate. Energy Convers. Manag. 118, 105–118. http://dx.doi.org/10.1016/j.enconman.2016.03.082.