Comparison of adaptive neuro-fuzzy inference system and multiple nonlinear regression for the productivity prediction of inclined passive solar still

Ahmed F. Mashaly and A. A. Alazba

ABSTRACT

Solar still productivity (SSP) is of vital importance in solar desalination project planning and management. In this investigation, the applicability of adaptive neuro-fuzzy inference systems (ANFIS) and multiple nonlinear regression (MNLR) in modeling SSP is investigated. Eight different membership functions (MFs) were used with ANFIS approach. Solar radiation, relative humidity, feed flow rate, total dissolved solids of feed, and brine are used as inputs to the models. The outcomes of the ANFIS are compared with those of the MNLR with respect to correlation coefficient (CC), root mean square error (RMSE), overall index of model performance (OI), and mean absolute error (MAE). Comparison results illustrate the generalized bell MF with ANFIS model has better accuracy than the other seven MFs in modeling SSP. Performance evaluation criteria show the predictive abilities of ANFIS and MNLR models were very similar and can be suggested to predict SSP effectively. Using the ANFIS model, the average value of CC, RMSE, OI, and MAE was 0.96, 0.05 L/m^2/h, 0.91, and 0.04 L/m^2/h, respectively. The corresponding values for the MNLR model were CC = 0.97, RMSE = 0.06 L/m^2/h, OI = 0.93, and MAE = 0.05 L/m^2/h. One of the advantages of MNLR model is using explicit equations.

Key words | adaptive neuro-fuzzy inference system, desalination, multiple nonlinear regression, prediction, solar still

Ahmed F. Mashaly (corresponding author)
A. A. Alazba
Alamoudi Water Research Chair,
King Saud University,
Riyadh,
Saudi Arabia
E-mail: *mashaly.ahmed@gmail.com*

A. A. Alazba
Agricultural Engineering Department,
King Saud University,
Riyadh,
Saudi Arabia

INTRODUCTION

Water scarcity is a major danger to humanity in the future. Consequently, converting saline water into pure water is one of the feasible solutions to meet water demand. A solar still is the simplest solar desalination device, but its productivity is very low. Hence, numerous experimental investigations and studies in recent years, such as those by Selvaraj & Natarajan (2018), Panchal & Mohan (2017), Lal *et al.* (2017), and Kaviti *et al.* (2016), have been done to evaluate, enhance, and raise the productivity of solar stills. The main disadvantage of experimental work is its cost and the fact that it is time-consuming. Thus, mathematical modeling for productivity is considered suitable to evaluate the performance of solar stills and facilitate assessment on the basis of efficiency.

In this study, an adaptive neuro-fuzzy inference system (ANFIS) as a mathematical modeling technique was developed and used for the performance analysis of an inclined solar passive solar still. The usage of ANFIS has been applied broadly in most areas of science and technology, especially in the field of solar engineering by several investigators (Amirkhani *et al.* 2015; Jović *et al.* 2016; Mohammadi *et al.* 2016; Yaïci & Entchev 2016; Belhachat & Larbes 2017; Halabi *et al.* 2018; Mashaly & Alazba 2018; Shabaan *et al.* 2018). ANFIS is the combination of artificial neural networks (ANNs) and fuzzy logic (FL) (Jang 1993). Compared with ANNs and FL, ANFIS needs less data learning and performs better in forecasting with its rule-based inference mechanism.

doi: 10.2166/aqua.2019.058

This study describes the application of ANFIS to the problem of predicting solar still productivity (SSP) and discusses the precision of eight different membership functions, namely, triangle, trapezoidal, generalized bell, Gaussian, two-sided Gaussian, Pi curve, product of two sigmoidal functions, and difference between two sigmoidal functions (DSIGMF), in SSP estimation. The second objective of this investigation aimed at comparing ANFIS as an intelligence technique and the multiple nonlinear regression (MNLR) as a classical technique for SSP forecasting. The input parameters to the MNLR model were the same as the ones used in the ANFIS models. The predictive capabilities of the models were assessed using standard statistical performance evaluation criteria.

MATERIALS AND METHODS

Experimental setup

The experiments were conducted at the Agricultural Research and Experiment Station, Department of Agricultural Engineering, King Saud University, Riyadh, Saudi Arabia (24°44′10.90″N, 46°37′13.77″E), during the period of February–April 2013, where the weather data were obtained from a weather station (model: Vantage Pro2, manufacturer: Davis, USA) located close to the experimental site (24°44′12.15″N, 46°37′14.97″E). The solar still system utilized in the experiments consists of one stage of C6000 panel (Carocell Solar Panel, F CUBED Ltd, Australia). The area of the used panel was 6 m^2. The solar still is manufactured as a panel using modern cost-effective materials, such as coated polycarbonate plastic. The panel heats and distills a film of water flowing over the absorber mat of the panel. The panel was fixed at an angle of 29° to horizontal. The basic construction materials were galvanized steel legs, aluminum frame, and polycarbonate covers. The transparent polycarbonate was coated from inside by special coating material to prevent fogging (patent for F CUBED, Australia). A cross-sectional view of the solar still is presented in Figure 1. The working operation of the available system is summarized in the following paragraphs.

The water was fed to the panel using a centrifugal pump (model: PKm 60, 0.5 HP, Pedrollo, Italy) with a constant flow rate of 10.74 L/hr. Eight drippers/nozzles drip the feed causing a film to flow over the absorbent mat. Under the absorbent mat there is an aluminum screen which helps to distribute the dropped water over the absorber mat. Beneath the aluminum screen, there is a plate, also made from aluminum. The aluminum was selected for the manufacturing process because aluminum is a hydrophilic material that assists even distribution of the sprayed water. The water flows through and over the absorbent mat, and as the solar energy is absorbed and partially collected inside the panel, water is heated and hot air is naturally circulated within the panel. The hot air flows in the upper part towards the top, and then reverses the direction towards the bottom. By this circulation, the humid air touches the cooled surfaces of the transparent polycarbonate cover and the bottom polycarbonate layer, thus water condenses and flows down the panel and is collected as a distilled stream. Seawater was used as feed water input to the system. The solar still system was run during the period from 23/02/2013 to 23/04/2013. Raw seawater was obtained from the Gulf, Dammam, East Saudi Arabia (26°26′24.19″N, 50°10′20.38″E). The initial concentration of the total dissolved solids (TDS), pH, density (ρ), and electrical conductivity (EC) of the raw seawater were 41.4 ppt, 8.02, 1.04 $g.cm^{-3}$, and 66.34 $mS\ cm^{-1}$, respectively. The productivity or the amount of distilled water produced (SSP) during a time period by the system was obtained by collecting the cumulative amount of water produced over time. The temperatures of the feed (T_F) and brine (T_B) were measured by using thermocouples (T-type, UK). Temperature data for feed brine water were recorded on a data logger (model: 177-T4, Testo, Inc., UK) at 1 min intervals. The amount of feed water (M_F) was measured by calibrated digital flow meter mounted on the feed water line (micro-flo, Blue-White, USA). The amount of brine water and distilled water were measured by graduated cylinder. TDS and EC were checked using calibrated (TDS) meter (Cole-Parmer Instrument, Vernon Hills, USA). A pH meter (model: 3510 pH meter, Jenway, UK) was utilized to determine the pH. ρ was measured by a digital density meter (model: DMA 35_N, Anton Paar, USA). The seawater was fed separately to the panel using the pump mentioned above. The residence time for the water to pass through the panel was about 20 minutes.

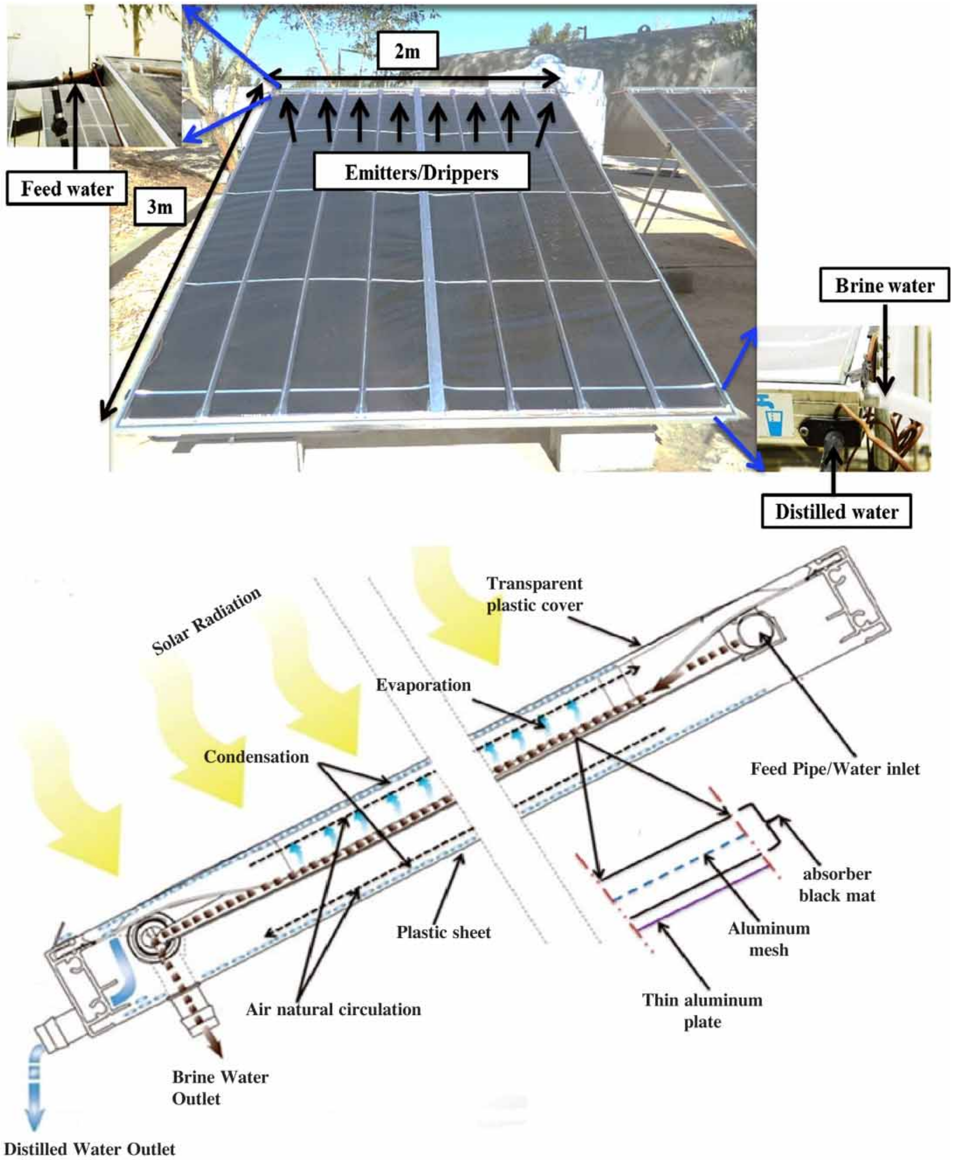


Figure 1 | Photograph and cross-sectional view of the inclined passive solar still used in the study.

Consequently, the flow rate for feed water, distilled water, and brine water were measured at 20 min. Also, the total dissolved solids of feed water (TDSF) and of brine water (TDSB) were measured every 20 minutes. The weather data, such as ambient temperature (To), relative humidity (RH), wind speed (WS), and solar radiation (SR) were obtained from the weather station mentioned above. Here, there is one dependent variable which is the SSP and nine independent variables which are To, RH, WS, SR, TF, TB, TDSB, TDSF, and MF.

Input parameters

In this study, field data obtained from the experimental work were used for the training, testing, and validation of the MNLR and ANFIS models. One of the most important steps in the modeling process for satisfactorily predicting results is the selection of the input parameters, since these parameters determine model structures and affect the weighted coefficient and the overall performance of the model. For this purpose, a correlation matrix is performed to assess relationships between the dependent parameter (SSP) and the independent parameters (To, RH, WS, SR, TF, TB, MF, TDSF, and TDSB) as given in Table 1. This matrix allows us to recognize how each parameter affects SSP and eventually which parameter(s) should be used as input in MNLR and ANFIS models. Furthermore, this matrix displays the findings of correlation analysis conducted between each pair of parameters. The strongest correlation is observed between SSP and SR with Pearson's correlation coefficient (CC) of +0.734. Furthermore, SSP is found to be well correlated with TDSF with CC = −0.402. The + and − signs refer to positive correlation and negative correlation, respectively. This agrees with the findings of Mashaly *et al.* (2016). Also, there is a significant correlation between SSP and MF and TDSB with CC = 0.25, and −0.172, respectively. On the other hand, a very weak correlation is found between SSP and To and TF with CC = −0.072, and −0.061, respectively; consequently, we do not consider them as input parameters. The correlation analysis also led to the exclusion of the WS and TB due to their high collinearity with other parameters, though there are significant correlations with the SSP. Although some of the parameters also appear correlated to others, these were included in the modeling process since their inclusion was found to improve its prediction performance, primarily by enhancing the CC. The same argument was also invoked to consider RH as an input parameter with low CC. The descriptive statistics of the input parameters used in the training, testing, and validation are shown in Table 2.

Adaptive neuro-fuzzy inference system (ANFIS)

The adaptive neuro-fuzzy inference system, which incorporates the best features of fuzzy logic and artificial neural network systems, is defined by Jang (1993). Architecturally, ANFIS is composed of if–else rules and input–output data couples of fuzzy and it uses neural network learning algorithms for training. Furthermore, ANFIS is an approach to simulate complex nonlinear mappings using neural network learning and fuzzy inference methodologies and has the capability of working with uncertain noisy and inaccurate environments. ANFIS utilizes the ANN training process to adjust the membership function and the associated parameter that approaches the desired data sets. The learning algorithm of ANFIS is a hybrid learning algorithm comprising the use of back-propagation learning algorithm and least squares method together (Jang 1993). In order to understand and simplify the process, a sample having two inputs and an output is considered. Five layers are used to build the ANFIS architecture of the first-order Sugeno-type inference system and presented in Figure 2. Two inputs, x and y, and one output, f, along with two fuzzy IF–THEN rules are taken into account as an example. In Figure 2 the circle and rectangle show a fixed node and an adaptive node, respectively. The functions of each layer of the five layers are explained in the following sections. For a first-order Sugeno fuzzy model, two fuzzy IF–THEN

Table 1 | Correlation coefficient matrix for studied parameters

	To	RH	WS	SR	TF	TB	MF	TDSF	TDSB	SSP
To	**1.00**	−0.66	−0.14	−0.15	0.91	0.06	0.44	−0.01	−0.15	−0.07
RH	−0.66	**1.00**	−0.08	0.15	−0.80	0.05	−0.72	0.23	0.45	0.01
WS	−0.14	−0.08	**1.00**	0.22	−0.01	0.33	−0.34	0.64	0.49	−0.31
SR	−0.15	0.15	0.22	**1.00**	−0.09	0.82	−0.27	0.22	0.39	0.73
TF	0.91	−0.80	−0.01	−0.09	**1.00**	0.13	0.48	0.06	−0.11	−0.06
TB	0.06	0.05	0.33	0.82	0.13	**1.00**	−0.40	0.49	0.57	0.40
MF	0.44	−0.72	−0.34	−0.27	0.48	−0.40	**1.00**	−0.75	−0.84	0.25
TDSF	−0.01	0.23	0.64	0.22	0.06	0.49	−0.75	**1.00**	0.94	−0.40
TDSB	−0.15	0.45	0.49	0.39	−0.11	0.57	−0.84	0.94	**1.00**	−0.17
SSP	−0.07	0.01	−0.31	0.73	−0.06	0.40	0.25	−0.40	−0.17	**1.00**

To, ambient temperature; RH, relative humidity; WS, wind speed; SR, solar radiation; TF, temperature of feed water; TB, temperature of brine water; MF, feed flow rate; TDSF, total dissolved solids of feed; TDSB, total dissolved solids of brine; SSP, solar still productivity.

Table 2 | Descriptive statistics of the input parameters used in the modeling process

	Input parameters														
	Training					Testing					Validation				
Statistical parameters	RH (%)	SR (W/m²)	M_F (L/min)	TDS_F (PPT)	TDS_B (PPT)	RH (%)	SR (W/m²)	M_F (L/min)	TDS_F (PPT)	TDS_B (PPT)	RH (%)	SR (W/m²)	M_F (L/min)	TDS_F (PPT)	TDS_B (PPT)
AVG	23.36	590.03	0.21	81.36	96.69	23.29	566.11	0.21	77.08	91.72	23.51	613.12	0.21	78.59	95.09
SE	1.24	17.10	0.00	2.79	2.77	2.43	33.56	0.01	5.17	5.54	2.57	44.87	0.01	7.66	7.27
MED	18.88	635.60	0.24	75.60	95.80	18.52	585.88	0.24	68.95	86.05	20.40	654.88	0.24	74.95	89.45
SD	13.08	180.99	0.04	29.49	29.34	13.77	189.86	0.04	29.24	31.32	10.29	179.46	0.04	30.66	29.09
MIN	12.90	75.10	0.13	41.40	46.20	13.00	188.29	0.13	41.60	49.00	13.00	157.62	0.13	41.70	51.80
MAX	66.95	920.69	0.25	130.00	132.80	70.00	890.50	0.25	129.30	130.50	50.19	880.60	0.25	128.90	130.40

AVG, average value; SE, standard error; MED, median; SD, standard deviation; MIN, minimum value; MAX, maximum value; RH, relative humidity; SR, solar radiation; MF, feed flow rate; TDSF, total dissolved solids of feed; TDSB, total dissolved solids of brine.

rules are as follows:

$$\textbf{Rule 1}\text{:If x is } A_1 \text{ and y is } B_1, \text{ then } f_1 = p_1x + q_1y + r_1 \quad (1)$$

$$\textbf{Rule 2}\text{:If x is } A_2 \text{ and y is } B_2, \text{ then } f_2 = p_2x + q_2y + r_2 \quad (2)$$

where x and y are the inputs and A_1, B_1, A_2, B_2 are fuzzy sets p_1, p_2, q_1, q_2, r_1, and r_2 are the coefficients of the output function that are determined during the training.

Layer 1 is the fuzzification layer (layer of input nodes). Every node i is an adaptive node with a node output expressed by:

$$O_i^1 = \mu_{A_i}(x), \text{ for } i = 1, 2 \quad (3)$$

$$O_i^1 = \mu_{B_{i-2}}(y), \text{ for } i = 3, 4 \quad (4)$$

where μ_{A_i} and $\mu_{B_{i-2}}$ are the fuzzy membership functions.

Layer 2 is the rule layer (layer of rule nodes). Every node i in this layer is a fixed node, marked by a circle and labeled Π, representing the simple multiplication. The output of this layer is the product of all the incoming signals and can be formulated as:

$$O_i^2 = w_i = \mu_{A_i}(x)\ \mu_{B_i}(y) \text{ for } i = 1, 2 \quad (5)$$

Layer 3 is the normalization layer (layer of average nodes). In this layer, the ith node is a circle labeled N, which computes the normalized firing strength as follows:

$$O_i^3 = \overline{w_i} = \frac{w_i}{w_1 + w_2}, \text{ for } i = 1, 2 \quad (6)$$

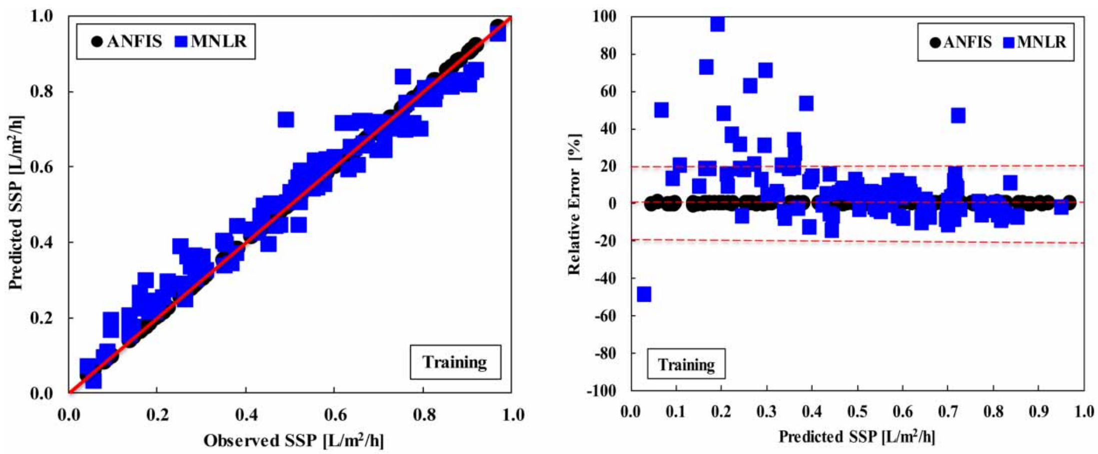
Layer 4 is the defuzzification layer (layer of consequent nodes). In this layer, every node i marked by a rectangle is an adaptive node with a node function. The output of this layer is calculated by:

$$O_i^4 = \overline{w_i}f_i = \overline{w_i}(p_ix + q_iy + r_i) \text{ for } i = 1, 2 \quad (7)$$

where $\{p_i, q_i, r_i\}$ is the parameter set of this node.

Layer 5 is the output layer. The single node in this layer is a fixed circle node labeled ∑, which calculates the final overall output as the summation of all incoming signals. The overall output is computed by this formula:

$$O_i^5 = f_{out} = \sum_{i=1}^{2} \overline{w_i}f_i = \frac{\sum_{i=1}^{2} w_if_i}{w_1 + w_2} \quad (8)$$



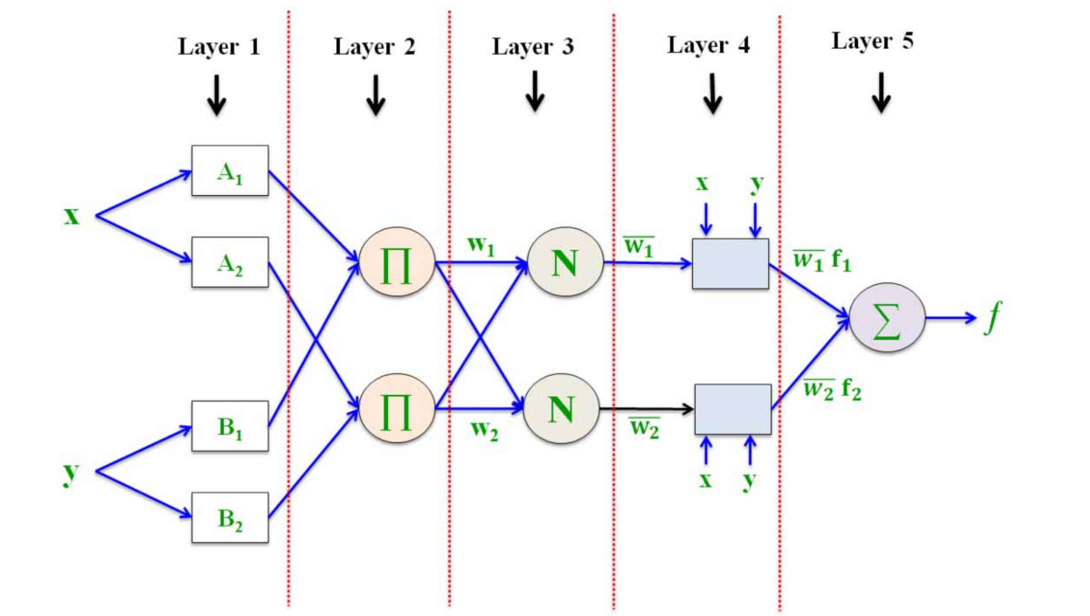


Figure 2 | Schematic diagrams of adaptive neuro-fuzzy inference system (ANFIS).

Finally, the overall output can be formulated as:

$$f_{out} = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \tag{9}$$

Substituting Equation (7) into Equation (10):

$$f_{out} = \overline{w_1} f_1 + \overline{w_2} f_2 \tag{10}$$

$$f_{out} = \overline{w_1}(p_1 x + q_1 y + r_1) + \overline{w_2}(p_2 x + q_2 y + r_2) \tag{11}$$

The final output can be written as:

$$f_{out} = (\overline{w_1} x) p_1 + (\overline{w_1} y) q_1 + (\overline{w_1}) r_1 + (\overline{w_2} x) p_2 + (\overline{w_2} y) q_2 + (\overline{w_2}) r_2 \tag{12}$$

In this investigation, the available data obtained from the experimental work were randomly divided into three portions: 70% as the training data sets (112 data points) for learning process, 20% as the testing data sets (32 data points) to test the precision of the model and 10% for the validation procedure (16 data points). Before the training, the data are normalized to be in a range between 0 and +1 in order to decrease their range and increase the precision of the findings. After the normalization process, the data are ready for the training process. The normalization of data was done according to the following equation:

$$X_n = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{13}$$

where X_n is normalized value, X_i is measured value of the variable, X_{max} is maximum measured value, and X_{min} is minimum measured value.

MATLAB software (MATLAB 8.1.0.604, R2013a, MathWorks Inc., USA) was used to develop the ANFIS model from the experimental data to forecast SSP. The Sugeno-type fuzzy inference system was used in the modeling of SSP. The grid partition method is employed to classify the input data and in making the rules (Jang & Sun 1995). In the modeling process, we employ eight different types of input MFs, including triangle (TRIMF), trapezoidal (TRAPMF), generalized bell (GBELLMF), Gaussian (GAUSSMF), two-sided Gaussian (GASUSS2MF), Pi curve (PIMF), the product of two sigmoidal functions (PSIGMF), and DSIGMF. The output MF was selected as a linear function. Moreover, a hybrid learning algorithm that combines the least-squares estimator and the gradient descent method is utilized to estimate the optimum values of the FIS parameters of the Sugeno-type (Jang & Sun 1995). The number of epochs was chosen as 50 owing to their small error.

Multiple nonlinear regression (MNLR)

Finding an appropriate relationship between a response parameter and a set of regression parameters is one of the classical problems in statistical analysis. Regression analysis is usually employed to find out the quantitative relationships of a response parameter with one or more explanatory parameters. Nonlinear regression is a form of regression analysis in which response data are modeled by a function, which is a nonlinear combination of the independent parameters. It depends on one or more independent parameter. Divergent from traditional multiple linear regression (MLR), which is restricted to determining linear models, MNLR can estimate models with arbitrary relationships between independent and dependent parameters. The generalized MNLR model is assumed to be:

$$\begin{aligned} Y = {} & A_0 + A_1X_1 + A_2X_2 + A_3X_3 + A_4X_4 + A_5X_5 \\ & + A_6X_1^2 + A_7X_2^2 + A_8X_3^2 + A_9X_4^2 + A_{10}X_5^2 + A_{11}X_1X_2 \\ & + A_{12}X_1X_3 + A_{13}X_1X_4 + A_{14}X_1X_5 + A_{15}X_2X_3 \\ & + A_{16}X_2X_4 + A_{17}X_2X_5 + A_{18}X_3X_4 + A_{19}X_3X_5 \\ & + A_{20}X_4X_5 + A_{21}X_1^2X_2 + A_{22}X_1^2X_3 + A_{23}X_1^2X_4 \\ & + A_{24}X_1^2X_5 + A_{25}X_2^2X_1 + A_{26}X_2^2X_3 + A_{27}X_2^2X_4 \\ & + A_{28}X_2^2X_5 + A_{29}X_3^2X_1 + A_{30}X_3^2X_2 + A_{31}X_3^2X_4 \\ & + A_{32}X_3^2X_5 + A_{33}X_4^2X_1 + A_{34}X_4^2X_2 + A_{35}X_4^2X_3 \\ & + A_{36}X_4^2X_5 + A_{37}X_5^2X_1 + A_{38}X_5^2X_2 + A_{39}X_5^2X_3 \\ & + A_{40}X_5^2X_4 + A_{41}X_1^2X_2^2 + A_{42}X_1^2X_3^2 + A_{43}X_1^2X_4^2 \\ & + A_{44}X_1^2X_5^2 + A_{45}X_2^2X_3^2 + A_{46}X_2^2X_4^2 + A_{47}X_2^2X_5^2 \\ & + A_{48}X_3^2X_4^2 + A_{49}X_3^2X_5^2 + A_{50}X_4^2X_5^2 \end{aligned} \tag{14}$$

where X_1 is RH (%), X_2 is SR (W/m^2), X_3 is MF (L/min), X_4 is TDSF (PPT), X_5 is TDSB (PPT), $X_1^2 - X_5^2$ are the second-orders of these parameters, X_1X_2, $X_1^2X_2^2$, $X_1^2X_2$, etc. are the interaction between each two parameters, Y is the dependent parameter (i.e., SSP, L/m^2/h) and $A_0 - A_{50}$ are regression coefficients.

As in the ANFIS modeling, the MNLR modeling process includes three stages, namely, training, testing, and validation. The data division is the same used in the ANFIS modeling. Therefore, the training, testing, and validation sets have 112, 32, and 16 data points, respectively. The Statistical Package for Social Science (IBM SPSS Statistics 23) program (SPSS Inc., Chicago, IL, USA) was used to develop the nonlinear regression analysis by stepwise method. Stepwise method is a technique to build a model by adding or removing predictor parameters, typically through a series of F-tests or T-tests. The parameters to be added or removed are selected based on the test statistics of the estimated coefficients.

Statistical evaluation of ANFIS and MNLR models' performance

The performance of ANFIS and MNLR models was evaluated through coefficient of correlation (CC), the root mean square error (RMSE), the overall index of model performance (OI), and the mean absolute error (MAE). Detailed descriptions about these statistical numerical indicators to assess the performance of ANFIS and MNLR models are available in Mashaly & Alazba (2016). The CC, RMSE, OI, and MAE were estimated as follows:

$$CC = \frac{\sum_{i=1}^{n}(SSP_{o,i} - \overline{SSP}_o)(SSP_{p,i} - \overline{SSP}_p)}{\sqrt{\sum_{i=1}^{n}(SSP_{o,i} - \overline{SSP}_o)^2 \times \sum_{i=1}^{n}(SSP_{p,i} - \overline{SSP}_p)^2}} \quad (15)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(SSP_{o,i} - SSP_{p,i})^2}{n}} \quad (16)$$

$$OI = \frac{1}{2}\left(2 - \frac{RMSE}{SSP_{max} - SSP_{min}} - \frac{\sum_{i=1}^{n}(SSP_{o,i} - SSP_{p,i})^2}{\sum_{i=1}^{n}(SSP_{o,i} - \overline{SSP}_o)^2}\right) \quad (17)$$

$$MAE = \frac{\sum_{i=1}^{n}|SSP_{o,i} - SSP_{p,i}|}{n} \quad (18)$$

where $SSP_{o,i}$ denotes the measured value, $SSP_{p,i}$ is the predicted value, $\overline{SSP}_o$ is the mean of measured values, $\overline{SSP}_p$ is the mean of predicted values, SSP_{max} is maximum measured value, SSP_{min} is minimum measured value, and n is the whole number of observations.

RESULTS AND DISCUSSION

Performances of ANFIS models

We developed eight ANFIS models with eight different types of input membership functions (MFs). The used MFs were TRIMF, TRAPMF, GBELLMF, GAUSSMF, GASUSS2MF, PIMF, DSIGMF, and PSIGMF. As we discussed, the eight developed ANFIS models during the training process with additional details oriented to the best ANFIS model. The ANFIS models developed have five inputs (RH, SR, MF, TDSF, and TDSB) and one output (SSP). Table 3 shows the outcomes of the statistical parameters, CC, RMSE, OI, and MAE, which are numerical indicators used to assess the agreement between observed and predicted SSP values during the training stage. For all the ANFIS models, in the input layer, five neurons were incorporated. For each neuron, the same three MFs were considered with three linguistic terms {low, medium, high} and accordingly, 243 ($3 \times 3 \times 3 \times 3 \times 3$) rules were developed for implementation of ANFIS models. The model properties of the ANFIS model structures are listed in Table 3. However, the ANFIS models' CC values ranged from 0.9998 to 0.9999, RMSE values from 0.0007 to 0.0048 L/m^2/h, OI values from 0.9972 to 0.9996, and MAE values from 0.0005 to 0.0027 L/m^2/h. The CC and OI values are very close to 1 while RMSE and MAE values are close to 0, indicating excellent agreement between the measured results and predicted results from the ANFIS models during the training stage. These findings emphasize the accuracy and efficiency of ANFIS models for estimating SSP by using the eight MFs.

The performance, in the training stage, for all MFs is approximately the same, but in relative terms, the highest performance is obtained with GBELLMF in the training process. The CC, RMSE, OI, and MAE for GBELLMF

Table 3 | Performance of the ANFIS model with various membership functions

	Model properties					Statistical parameters			
MFs	**NN**	**NLP**	**NNP**	**TNP**	**NFR**	**CC**	**RMSE**	**OI**	**MAE**
TRIMF	524	1,458	45	1,503	243	0.9999	0.0040	0.9977	0.0027
TRAPMF	524	1,458	60	1,518	243	0.9998	0.0048	0.9972	0.0026
GBELLMF	**524**	**1,458**	**45**	**1,503**	**243**	**0.9999**	**0.0013**	**0.9993**	**0.0007**
GAUSSMF	524	1,458	30	1,488	243	0.9999	0.0014	0.9992	0.0007
GAUSS2MF	524	1,458	60	1,518	243	0.9999	0.0012	0.9994	0.0007
PIMF	524	1,458	60	1,518	243	0.9999	0.0041	0.9976	0.0022
PSIGMF	524	1,458	60	1,518	243	0.9999	0.0007	0.9996	0.0005
DSIGMF	524	1,458	60	1,518	243	0.9999	0.0007	0.9996	0.0005

NN, number of nodes; NLP, number of linear parameters; NNP, number of nonlinear parameters; TNP, total number of parameters; NFR, number of fuzzy rules; CC, coefficient of correlation; RMSE, root mean square error; OI, overall index of model performance; MAE, mean absolute error.

Table 4 | Equations representing stepwise MNLR models generated for estimating SSP

Model	Mathematical expression	CC	SEE
MNLR-1	$SSP = -0.118 + 0.005\ SR \cdot MF$	0.859	0.1256
MNLR-2	$SSP = -0.276 + 0.006\ SR \cdot MF + 0.001\ MF \cdot RH^2$	0.926	0.0931
MNLR-3	$SSP = -0.105 + 0.005\ SR \cdot MF + 0.001\ MF \cdot RH^2 - 4.12 \times 10^{-4}\ MF^2 \cdot TDSF^2$	0.955	0.0734
MNLR-4	$SSP = -0.330 + 0.004 + SR \cdot MF - 0.001 MF \cdot RH^2 + 2.18 \times 10^{-4}\ MF^2 \cdot TDSF^2 - 1.80 \times 10^{-5}\ MF \cdot TDSB$	0.972	0.0589
MNLR-5	$SSP = -0.281 + 0.004\ SR \cdot MF - 2.21 \times 10^{-4}\ MF \cdot RH^2 - 0.001\ MF^2 \cdot TDSF^2 + 0.032\ MF \cdot TDSB + 1.31 \times 10^{-10} RH^2 \cdot SR^2$	0.977	0.0533
MNLR-6	$SSP = -0.266 + 0.004\ SR \cdot MF - 3.61 \times 10^{-5}\ MF \cdot RH^2 - 0.001\ MF^2 \cdot TDSF^2 + 0.031\ MF \cdot TDSB + 1.68 \times 10^{-10} RH^2 \cdot SR^2 - 2.91 \times 10^{-9} RH^2 \cdot TDSF^2$	0.978	0.0518
MNLR-7	$SSP = -0.267 + 0.004\ SR \cdot MF - 0.001\ MF^2 \cdot TDSF^2 + 0.030\ MF \cdot TDSB + 1.65 \times 10^{-10} RH^2 \cdot SR^2 - 5.16 \times 10^{-9} RH^2 \cdot TDSF^2$	0.978	0.0516
MNLR-8	$SSP = -0.297 + 0.004\ SR \cdot MF - 0.002\ MF^2 \cdot TDSF^2 + 0.039\ MF \cdot TDSB + 2.02 \times 10^{-10} RH^2 \cdot SR^2 - 1.12 \times 10^{-8} RH^2 \cdot TDSF^2 + 1.48 \times 10^{-4}\ MF \cdot TDSF^2$	0.979	0.0498
MNLR-9	$SSP = -0.185 + 0.002\ SR \cdot MF - 0.002\ MF^2 \cdot TDSF^2 + 0.036\ MF \cdot TDSB + 2.28 \times 10^{-10} RH^2 \cdot SR^2 - 1.19 \times 10^{-8} RH^2 \cdot TDSF^2 + 1.40 \times 10^{-4}\ MF \cdot TDSF^2 + 8.39 \times 10^{-6}\ SR^2 \cdot MF^2$	0.979	0.0480
MNLR-10	$SSP = -0.160 + 0.001\ SR \cdot MF - 0.002\ MF^2 \cdot TDSF^2 + 0.035\ MF \cdot TDSB + 2.52 \times 10^{-12} RH^2 \cdot SR^2 - 1.89 \times 10^{-8} RH^2 \cdot TDSF^2 + 1.83 \times 10^{-4}\ MF \cdot TDSF^2 + 1.09 \times 10^{-5}\ SR^2 \cdot MF^2 + 2.47 \times 10^{-7} SR \cdot RH^2$	0.980	0.0464
MNLR-11	$SSP = -0.160 + 0.001\ SR \cdot MF - 0.002\ MF^2 \cdot TDSF^2 + 0.035\ MF.TDSB - 1.90 \times 10^{-8}\ RH^2 \cdot TDSF^2 + 1.83 \times 10^{-4}\ MF \cdot TDSF^2 + 1.10 \times 10^{-5}\ SR^2 \cdot MF^2 + 2.50 \times 10^{-7}\ SR \cdot RH^2$	0.981	0.0442

CC: coefficient of correlation; SEE: standard error of the estimate.

were 0.9999, 0.0013 L/m^2/h, 0.9993, and 0.0007 L/m^2/h, respectively. This agrees with the results of Taghavifar & Mardani (2015). However, the best/selected ANFIS structure for SSP prediction was obtained by using GBELLMF and consisted of five layers. The 1st layer of the developed model included the input parameter (RH, SR, M_F, TDS_F, and TDS_B) membership functions (MFs). This layer provides the input parameter values to the following layer. The 2nd layer was a MF layer, which determined and checked the weights for each MF. For the best ANFIS model, the fuzzification layer (2nd layer) contained 15 nodes and 45 nonlinear parameters. The rule layer (3rd layer) with 524 nodes achieved a pre-condition matching process for fuzzy rules. The 4th layer (the defuzzification layer) with 524 nodes and 1,458 linear parameters took the inference of the rules and generated output values. The 5th layer summed up and combined the inputs and transformed the fuzzy classification into a binary outcome. Overall, the total number of parameters and fuzzy rules were 1,503 and 243, respectively, for the ANFIS model using GBELLMF.

Performances of MNLR models

The final equations obtained from MNLR are presented in Table 4. The SE, t-stat, and *p*-value for independent variables in the equations are illustrated in Table 5. MNLR models are developed by iteratively adding and removing the terms from a multi-nonlinear model based on their significance within a regression by using SPSS software. All five independent variables (*RH*, *SR*, *MF*, *TDSF*, and *TDSB*), their second-orders and interaction terms were used to create predicting models for SSP. MNLR produced 11 models with 1–8 predictor variables, where $SR \cdot MF$ (interaction between *SR* and *MF*) was involved in each set of predictor variables, as demonstrated in Table 4. Overall, the interaction between *SR* and *MF* has the most effect on SSP modeling.

The CC values associated with each of the 11 models ranged from 0.859 to 0.981. Corresponding standard errors of the estimate (SEE) ranged from 0.0442 to 0.1256 L/m^2/h. It can be noted from Table 4 that the absence or presence of some of the input variables in the MNLR models significantly affects the performances of these models. MNLR-1 with just the $SR \cdot MF$ performed relatively worst, with CC = 0.859 and SSE = 0.1256 L/m^2/h. MNLR-2 performed better than MNLR-1, owing to the presence of $MF.RH^2$. The CC value of MNLR-2 was increased by 7.80% than that for MNLR-1. Furthermore, the SEE value of MNLR-2 was decreased by 25.88% than that for MNLR-1. In MNLR-3, adding the $MF^2.TDSF^2$ caused an increase of 3.13% in CC's value, decrease of 21.16% in SEE's value compared to MNLR-2. The same trend was followed with MNLR-4 by the inclusion of both $MF^2.TDSF^2$ and $MF.TDSB$ (CC = 0.972, SEE = 0.0589 L/m^2/h). Following MNLR-4, the accuracy was dramatically unchanged. The value of CC was between 0.977 and 0.981, and SEE ranged from 0.0533 L/m^2/h to 0.0442 L/m^2/h. However, MNLR-11 showed the best model for estimating SSP (CC = 0.981, SEE = 0.0442 L/m^2/h) which involved $SR.MF$, $MF.TDSF$, $MF.TDSB$, $RH^2.TDSF^2$, $MF.TDSF^2$, $SR^2.MF^2$, and $SR.RH^2$ as predictor variables. Furthermore, the MNLR-11 model provides better performance in estimating SSP, as is reflected in the values of the statistical parameters (Table 5). The selected model has shown the significance of all independent variables at the significance level of 5% (*p*-value <0.05). The standard error (SE) determines the accuracy of the estimate of the coefficient. The smaller the SE, the more accurate the estimate. The significance of each coefficient of the obtained MNLR model was determined by t-stat and *p*-value, which are displayed in Table 5. The larger the t-stat and the smaller the *p*-value, the more significant is the corresponding coefficient. The absolute value of the t-stat should be greater than the critical t-value.

Comparison between ANFIS and MNLR models

In this section, performance comparison is made between the ANFIS and MNLR models. The performance of these models was assessed according to statistical criteria such as CC, RMSE, OI, and MAE. The findings of applying these models are compared in Table 6. During the training process, it is clear from Table 6 that the values predicted by the ANFIS fit perfectly with the observed values during the training process as reflected in the values of the statistical indicators. For the ANFIS model, the CC (0.999) value is very close to 1, and the RMSE value (0.001 L/m^2/h) is close to 0. The corresponding OI value was 0.999, which is very close to 1. The MAE value, as well, was 0.001 L/m^2/h, which was very close to 0. These results indicate that the ANFIS model is better than the performance of the MNLR model (CC = 0.981, RMSE = 0.053 L/m^2/h, OI = 0.948, and MAE = 0.041 L/m^2/h) in the training process, as indicated in Table 6.

In the testing stage, ANFIS statistics were CC = 0.959, RMSE = 0.070 L/m^2/h, OI = 0.910, and MAE = 0.045 L/m^2/h. However, the CC values in the MNLR model were about 0.31% more accurate than for the ANFIS model, as shown in Table 6. The RMSE value of the ANFIS model was almost 1.06 times that of the value of the MNLR model (0.066 L/m^2/h). The OI value for the ANFIS model was less than its value for the MNLR model (0.921). The value of MAE for the MNLR model (0.051 L/m^2/h) was almost 1.13 times that of the ANFIS model. It is quite clear that the results for both models are very close at this stage.

In the validation stage, the MNLR model had a CC value of about 6.23% more accurate than that of the ANFIS model, as shown in Table 6. The value of RMSE for the ANFIS model (0.085 L/m^2/h) was almost 1.55 times that of the MNLR model (0.055 L/m^2/h). The OI value for the MNLR model was 12.65% more accurate than that of the ANFIS model. The MAE value of

Table 5 | Standard error (SE) of regression coefficients, t statistic (t-stat) and probability (*p*-value) of independent variables for MNLR models

Model		**Intercept**	**Independent parameters** SR.MF	MF.RH^2	MF^2.TDSF^2	MF.TDSB	RH^2.SR^2	RH^2.TDSF^2	MF.TDSF^2	SR^2.MF^2	SR.RH^2
MNLR-1	SE	0.037	2.88×10^{-04}								
	t-stat	−3.154	17.589								
	p-value	0.002	9.52×10^{-34}								
MNLR-2	SE	0.032	2.22×10^{-04}	7.50×10^{-05}							
	t-stat	−8.559	25.435	9.555							
	p-value	8.08×10^{-14}	1.19×10^{-47}	4.43×10^{-16}							
MNLR-3	SE	0.033	1.78×10^{-04}	6.30×10^{-05}	5.00×10^{-05}						
	t-stat	−3.181	30.360	8.224	−8.217						
	p-value	0.002	1.06×10^{-54}	4.80×10^{-13}	4.98×10^{-13}						
MNLR-4	SE	0.039	1.85×10^{-04}	5.70×10^{-05}	1.04×10^{-04}	0.004					
	t-stat	−8.409	24.311	5.618	−11.119	7.766					
	p-value	1.97×10^{-13}	2.22×10^{-45}	1.54×10^{-07}	1.50×10^{-19}	5.18×10^{-12}					
MNLR-5	SE	0.037	1.79×10^{-04}	1.20×10^{-04}	9.60×10^{-05}	0.003	2.62×10^{-11}				
	t-stat	−7.627	23.277	−1.834	−13.063	9.185	4.986				
	p-value	1.09×10^{-11}	1.82×10^{-43}	0.069	8.14×10^{-24}	3.84×10^{-15}	2.00×10^{-06}				
MNLR-6	SE	0.036	1.82×10^{-04}	1.36×10^{-04}	9.70×10^{-05}	0.003	2.91×10^{-11}	1.84×10^{-09}			
	t-stat	−7.328	22.117	−0.266	−12.204	9.026	5.776	−2.663			
	p-value	5.02×10^{-11}	2.59×10^{-41}	0.791	7.37×10^{-22}	9.33×10^{-15}	7.88×10^{-08}	0.009			
MNLR-7	SE	0.036	1.74×10^{-04}		8.50×10^{-05}	0.003	2.60×10^{-11}	1.58×10^{-09}			
	t-stat	−7.464	23.253		−13.852	9.523	6.316	−3.266			
	p-value	2.45×10^{-11}	1.99×10^{-43}		1.58×10^{-25}	6.64×10^{-16}	6.41×10^{-09}	0.001			
MNLR-8	SE	0.034	1.72×10^{-04}		2.43×10^{-04}	0.004	2.56×10^{-11}	2.02×10^{-09}	3.40×10^{-05}		
	t-stat	−8.784	21.961		−8.947	11.013	7.881	−5.545	4.335		
	p-value	3.24×10^{-14}	4.80×10^{-41}		1.41×10^{-14}	3.28×10^{-19}	3.19×10^{-12}	2.21×10^{-07}	3.40×10^{-05}		
MNLR-9	SE	0.050	0.001		2.38×10^{-04}	0.004	2.62×10^{-11}	1.96×10^{-09}	3.30×10^{-05}	3.00×10^{-06}	
	t-stat	−3.729	2.896		−8.538	10.258	8.705	−6.080	4.249	2.993	
	p-value	3.13×10^{-04}	0.005		1.21×10^{-13}	1.78×10^{-17}	5.19×10^{-14}	2.01×10^{-08}	4.70×10^{-05}	0.003	
MNLR-10	SE	0.049	0.001		2.33×10^{-04}	0.003	7.92×10^{-11}	3.00×10^{-09}	3.50×10^{-05}	3.00×10^{-06}	8.23×10^{-08}
	t-stat	−3.284	2.016		−9.258	10.190	0.032	−6.309	5.249	3.863	3.004
	p-value	0.001	0.046		3.31×10^{-15}	2.79×10^{-17}	0.975	7.13×10^{-09}	8.25×10^{-07}	1.96×10^{-04}	0.003
MNLR-11	SE	0.048	0.001		2.31×10^{-04}	0.003		2.45×10^{-09}	3.30×10^{-05}	3.00×10^{-06}	2.61×10^{-08}
	t-stat	−3.303	2.042		−9.373	10.307		−7.725	5.540	3.953	9.566
	p-value	0.001	0.044		1.70×10^{-15}	1.39×10^{-17}		7.31×10^{-12}	2.30×10^{-07}	1.41×10^{-04}	6.27×10^{-16}

Table 6 | Statistical performance of the selected ANFIS and MNLR models during training, testing and validation stages

	Training		Testing		Validation	
Statistical criteria	**ANFIS**	**MNLR**	**ANFIS**	**MNLR**	**ANFIS**	**MNLR**
CC	0.999	0.981	0.959	0.962	0.915	0.972
RMSE (L/m^2/h)	0.001	0.053	0.070	0.066	0.085	0.055
OI	0.999	0.948	0.910	0.921	0.830	0.935
MAE (L/m^2/h)	0.001	0.041	0.045	0.051	0.063	0.044

CC, coefficient of correlation; RMSE, root mean square error; OI, overall index of model performance; MAE, mean absolute error.

0.063 L/m^2/h for the ANFIS model was almost 1.43 times that of the MNLR (0.044 L/m^2/h). The CC, RMSE, OI, and MAE values confirm that ANFIS model performed relatively poorly during the validation phase. However, generally, Table 6 clearly shows that the predictive performance of ANFIS and MNLR models is somewhat similar. These outcomes indicate that both models can be used successfully for SSP modeling.

Another representation of the results generated using the ANFIS and MNLR models, is demonstrated in Figure 3, where the observed and predicted SSP values from the models during the training, testing, and validation stages are compared. These are in the form of a scatter (1:1) plot (left panel) and a relative error plot (right panel). The data were mostly evenly and tightly distributed around the 1:1 line, as illustrated in Figure 3, and there was a very close agreement between the observed and predicted values by the ANFIS model during the training process, while most of the values were inaccurate with the MNLR model during the same process. In the testing stage, Figure 3 illustrates that the most of the values predicted by the ANFIS model were acceptable, while the majority of predicted values were somewhat imprecise when using the MNLR model. In the validation stage,

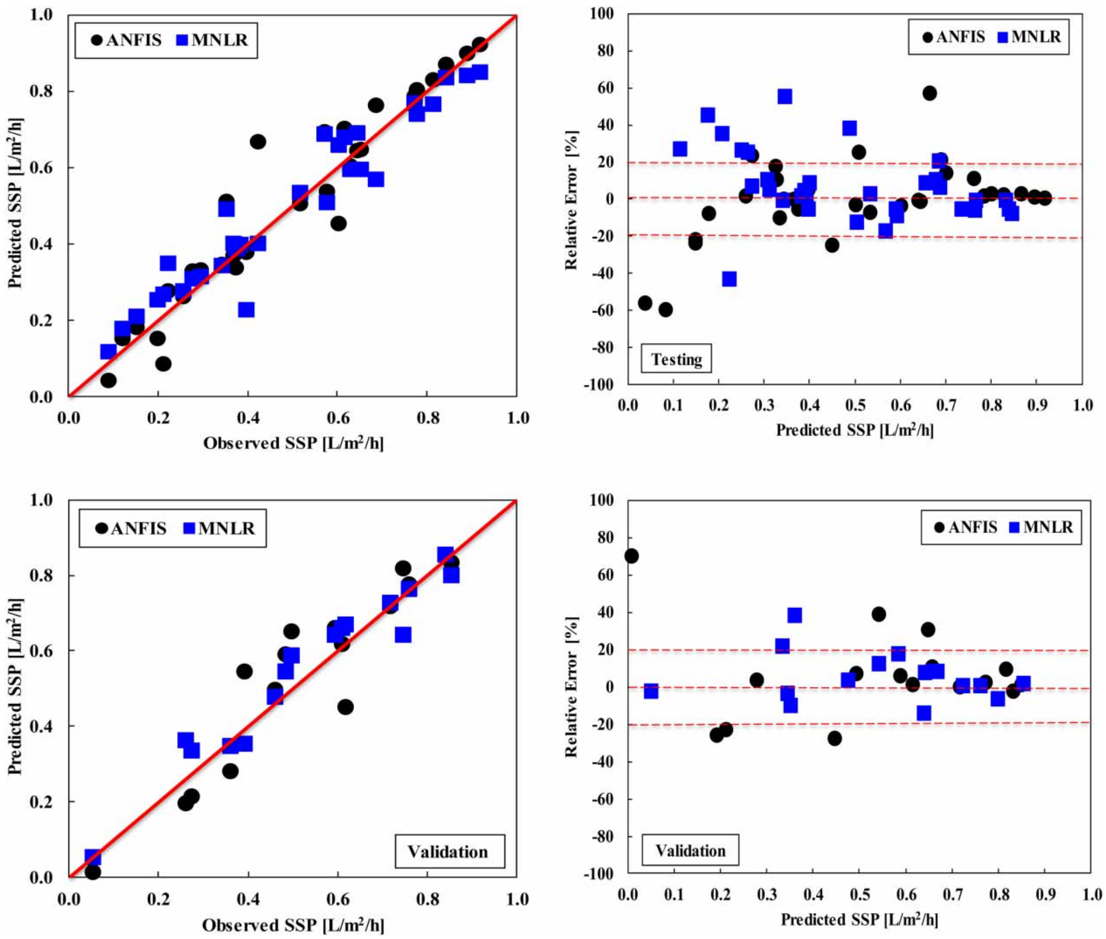


Figure 3 | Performance of the ANFIS and MNLR models during all processes.

there is a similarity in performance between the two models. However, the average relative errors were 8.11%, 6.71%, and 5.05% for the MNLR model during the training, testing, and validation phases, respectively. The corresponding values for the ANFIS model were generally lower at 0.01%, −1.12%, and 6.23% during the training, testing, and validation phases, respectively. In general, both models showed efficiency and accuracy in the SSP prediction process as in Table 6 and Figure 3.

CONCLUSIONS

The main objective of this study was to assess the performance of ANFIS and MNLR models for estimating SSP of an inclined passive solar still. Five parameters, relative humidity, SR, feed flow rate, total dissolved solids of feed, and total dissolved solids of brine, were used as inputs for the ANFIS and MNLR models. The SSP was used as a target variable. The performance of the models was assessed by comparing the prediction results for both approaches with the experimental results by using a variety of standard statistical parameters, namely, CC, RMSE, OI, and MAE. These comparisons revealed that the ANFIS and MNLR models can be suggested to forecast SSP effectively because of fast, accurate, and reliable results. There was agreement between the predicted and observed data for the ANFIS and MNLR models. Generally, the trend of the MNLR results is the same as that of the ANFIS model. The ANFIS model gave accurate estimates, but the MNLR models are easier to use, as they compute the SSP using explicit algebraic equations.

ACKNOWLEDGEMENTS

The project was financially supported by King Saud University, Vice Deanship of Research Chairs.

REFERENCES

Amirkhani, S., Nasirivatan, S., Kasaeian, A. B. & Hajinezhad, A. 2015 ANN and ANFIS models to predict the performance of solar chimney power plants. *Renew. Energ.* **83**, 597–607.

Belhachat, F. & Larbes, C. 2017 Global maximum power point tracking based on ANFIS approach for PV array configurations under partial shading conditions. *Renew. Sust. Energ. Rev.* **77**, 875–889.

Halabi, L. M., Mekhilef, S. & Hossain, M. 2018 Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation. *Appl. Energy* **213**, 247–261.

Jang, R. J. S. 1993 ANFIS: adaptive-network-based fuzzy inference system. *IEEE T. Syst. Man. Cy-S.* **23**, 665–685.

Jang, J. S. R. & Sun, C. T. 1995 Neuro-fuzzy modeling and control. *Proc. IEEE* **83** (3), 378–405.

Jović, S., Aničić, O., Marsenić, M. & Nedić, B. 2016 Solar radiation analyzing by neuro fuzzy approach. *Energ. Buildings* **129**, 261–263.

Kaviti, A. K., Yadav, A. & Shukla, A. 2016 Inclined solar still designs: a review. *Renew. Sust. Energ. Rev.* **54**, 429–451.

Lal, R. K., Mishra, S., Dwivedi, J. P. & Dwivedi, H. 2017 A comprehensive study of the different parameters of solar still. *Mater. Today* **4** (2), 3572–3580.

Mashaly, A. F. & Alazba, A. A. 2016 Comparison of ANN, MVR, and SWR models for computing thermal efficiency of a solar still. *Int. J. Green Energy* **13** (10), 1016–1025.

Mashaly, A. F. & Alazba, A. A. 2018 Membership function comparative investigation on productivity forecasting of solar still using adaptive neuro-fuzzy inference system approach. *Environ. Prog. Sustain.* **37** (1), 249–259.

Mashaly, A. F., Alazba, A. A. & Al-Awaadh, A. M. 2016 Assessing the performance of solar desalination system to approach near-ZLD under hyper arid environment. *Desalin. Water Treat.* **57**, 12019–12036.

Mohammadi, K., Shamshirband, S., Kamsin, A., Lai, P. C. & Mansor, Z. 2016 Identifying the most significant input parameters for predicting global solar radiation using an ANFIS selection procedure. *Renew. Sust. Energ. Rev.* **63**, 423–434.

Panchal, H. & Mohan, I. 2017 Various methods applied to solar still for enhancement of distillate output. *Desalination* **415**, 76–89.

Selvaraj, K. & Natarajan, A. 2018 Factors influencing the performance and productivity of solar stills – a review. *Desalination* **435** (1), 181–187.

Shabaan, S., El-Sebah, M. I. A. & Bekhit, P. 2018 Maximum power point tracking for photovoltaic solar pump based on ANFIS tuning system. *J. Elect. Syst. Info. Tech.* **5** (1), 11–22.

Taghavifar, H. & Mardani, A. 2015 Evaluating the effect of tire parameters on required drawbar pull energy model using adaptive neuro-fuzzy inference system. *Energy* **85**, 586–593.

Yaïci, W. & Entchev, E. 2016 Adaptive neuro-fuzzy inference system modelling for performance prediction of solar thermal energy system. *Renew. Energ.* **86**, 302–315.

First received 7 May 2018; accepted in revised form 21 December 2018. Available online 1 February 2019