Application of adaptive neuro-fuzzy inference system (ANFIS) for modeling solar still productivity

Ahmed F. Mashaly and A. A. Alazba

ABSTRACT

Climate change is a major challenge to humankind. Solar desalination is a strategic option for overcoming water scarcity as a result of climate change. Modeling of solar still productivity (SSP) plays a significant role in the success of a solar desalination project to optimize capital expenditures and maximize production. A solar still was used to desalinate seawater in this study. An adaptive neuro-fuzzy inference system (ANFIS) for prediction of SSP was developed with different types of input membership functions (MFs). The investigation used the principal parameters affecting SSP, which are the solar radiation, relative humidity, total dissolved solids (TDS) of feed, TDS of brine, and feed flow rate. The performance of ANFIS models in the training, testing, and validation stages are compared with the observed data. The ANFIS model with Pi-shaped curve MF provides better and higher prediction accuracy than models with other MFs. The ANFIS was an adequate model for the prediction of SSP and yielded root mean square error and correlation coefficient values of 0.0041 L/m^2/h and 99.99%, respectively. Sensitivity analysis revealed that solar radiation is the most effective parameter on SSP. Finally, the ANFIS model can be used effectively as a design tool for solar still systems.

Key words | ANFIS, artificial intelligence, desalination, solar, solar still production

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

INTRODUCTION

Water scarcity is defined as a situation in which insufficient water resources are available to satisfy requirements (EU 2007). A common method of decreasing water scarcity in the world is desalination. Desalination is not, of course, problem free. Desalination – based on fossil fuels – can cause an increase in greenhouse gas emissions, further contributing to the root cause of climate change. Climate change is one of the main threats to humanity in the 21st century. Desalination based on renewable energy (e.g., solar energy), can prove to be a viable alternative to fossil-fueled desalination and a way of reducing greenhouse gas emissions and their contribution to climate change (Khoi & Suetsugi 2012; Mashaly *et al.* 2016).

The simplest example of solar desalination is the solar still. A solar still is an environmentally friendly device that uses the free solar energy to desalinate impure or low-quality water to produce fresh water. A solar still has many merits over other desalination methods. For example, a solar still system exploits free, renewable, and clean energy. Additionally, there are no moving parts in the solar still, making it easier to build, maintain, and use. Moreover, it can be manufactured easily with locally available materials and no skilled labor is needed to make it. On the other hand, the main problem to using a solar still is its low productivity compared with other desalination technologies (Kabeel *et al.* 2012; Ayoub *et al.* 2015).

Consequently, solar stills should be optimally designed and operated in order to obtain a reasonable productivity. Prediction of this productivity is one of the most important parameters to be accurately determined. Modeling or forecasting of solar still productivity (SSP) functionally serves designers, users, and investors in planning and making decisions about the utilization of a solar still. With progress in computational techniques, the application of artificial intelligence (AI) in the field of solar systems, such as solar stills,

doi: 10.2166/aqua.2017.138

could lead to results that are not easily obtained with classical modeling techniques. These techniques are often difficult and require a great deal of calculating time to reach a solution.

In recent years, AI techniques (e.g., artificial neural network (ANN), fuzzy logic (FL), and adaptive neuro-fuzzy inference systems (ANFIS)) have been widely used in many fields of engineering applications (Barua *et al.* 2010; El Shafie *et al.* 2012; Sun *et al.* 2015; Taghavifar & Mardani 2015; Jiang *et al.* 2016; Maachou *et al.* 2016; Mashaly & Alazba 2016a, 2016b). This is due to the ability of these techniques for generating the non-linear relationships between input and output parameters which are characteristic of solar engineering problems. Additionally, in the solar engineering field alone, ANNs have been used to estimate daily global solar radiation (Gairaa *et al.* 2016). ANN and ANFIS models were used to predict the performance of solar power plants (Amirkhani *et al.* 2015). Global solar energy was modeled by FL technique (Rizwan *et al.* 2014). ANN was employed to forecast solar photovoltaic energy production (Dumitru *et al.* 2016). Yaïci & Entchev (2016) used ANFIS procedure for performance prediction of a solar thermal energy system. Santos *et al.* (2012), Mashaly *et al.* (2015), and Mashaly & Alazba (2015, 2016c, 2017) studied some applications of ANNs in solar desalination. All these above studies concluded that these procedures are able to forecast and model with excellent accuracy. To our knowledge, no previous research has assessed ANFIS's ability to model SSP. Consequently, this study presents an ANFIS model to predict SSP based on meteorological and operational parameters. The aims of this research are to: (1) develop ANFIS models to predict SSP; (2) evaluate the performance of the ANFIS models by statistical comparison of the SSP obtained from the model to experimental results; and (3) compare the accuracy of ANFIS models with different membership functions (MFs).

MATERIALS AND METHODS

Experimental work

The experiments were conducted at the Agricultural Research and Experiment Station at the Department of Agricultural Engineering, King Saud University, Riyadh, Saudi Arabia (24°44′10.90″N, 46°37′13.77″E) between February and April 2013. The weather data were obtained from a weather station (model: Vantage Pro2; manufacturer: Davis Instruments, USA) close by the experimental site (24°44′12.15″N, 46°37′14.97″E). The solar still system used in the experiments was constructed from a 6 m^2 single stage of C6000 panel (Carocell Solar Panel, F Cubed, Australia). The solar still panel was manufactured using modern, cost-effective materials such as coated polycarbonate plastic. When heated, the panel distilled a film of water that flowed over the absorber mat of the panel. The panel was fixed at an angle of 29° to the horizontal. The basic construction materials were galvanized steel legs, an aluminum frame, and polycarbonate covers. The transparent polycarbonate was coated on the inside with a special material to prevent fogging (patented by F Cubed, Australia). A cross-sectional view of the solar still is illustrated in Figure 1.

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/h. The feed was supplied by eight drippers/nozzles, creating a film of water that flowed over the absorbent mat. Underneath the absorbent mat was an aluminum screen that helped to distribute the water across the mat. Beneath the aluminum screen was an aluminum plate. Aluminum was chosen for its hydrophilic properties, to assist in the even distribution of the sprayed water. Water flowed through and over the absorbent mat, and solar energy was absorbed and partially collected inside the panel; as a result, the water was heated and hot air circulated naturally within the panel. First, the hot air flowed towards the top of the panel, then reversed its direction to approach the bottom of the panel. During this process of circulation, the humid air touches the cooled surfaces of the transparent polycarbonate cover and the bottom polycarbonate layer, causing condensation. The condensed water flowed down the panel and was collected in the form of a distilled stream. Seawater was used as a feed-water input to the system. The system was run from 02/23/2013 to 04/23/2013. Raw seawater was obtained from the Gulf, Dammam, in eastern Saudi Arabia (26°26′24.19″N, 50°10′20.38″E). The initial concentrations of the total dissolved solids (TDS), pH, density (ρ), and electrical conductivity (EC), are 41.4 ppt, 8.02, 1.04 g·cm^{-3}, and

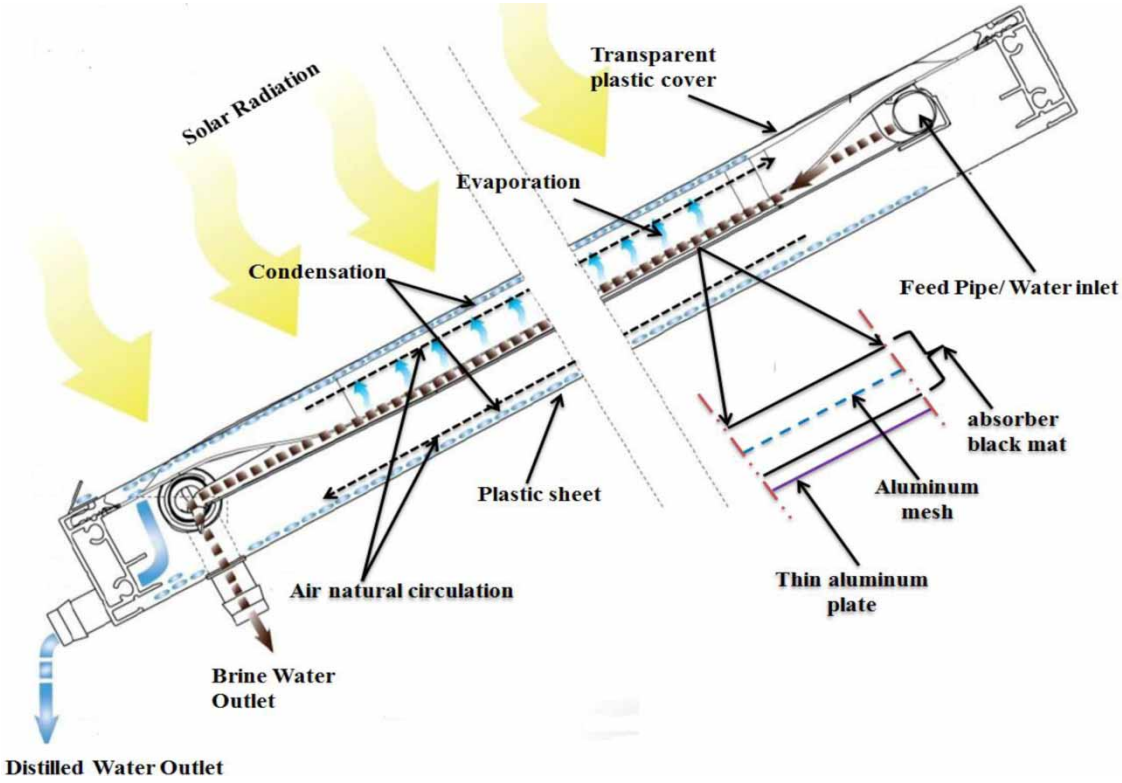

Figure 1 | Cross-sectional view of the solar still.

66.34 mS cm^{-1}, respectively. The SSP or the amount of distilled water produced (SSP) by the system in a given time was obtained by collecting and measuring the amount of water cumulatively produced over time. The temperatures of the feed water (T_F) and brine water (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 minute intervals. The amount of feed water (M_F) was measured by a 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 a graduated cylinder. TDS concentration and EC were measured using a TDS-calibrated meter (Cole-Parmer Instrument Co. Ltd, Vernon Hills, USA). A pH meter (model: 3510 pH meter; Jenway, UK) was used to measure pH. A digital-density meter (model: DMA 35_N, Anton Paar, USA) was used to measure ρ. The seawater was fed separately to the panel using the pump described above. The residence time – the time taken for the water to pass through the still/panel – was approximately 20 minutes. Therefore, the flow rate for the feed water, the distilled water, and the brine water was measured each 20 minutes. Also, the total dissolved solids of feed water (TDS_F) and brine water (TDS_B) were measured every 20 minutes. The weather data, such as ambient temperature (T_o), relative humidity (RH), wind speed (WS), and solar radiation (SR), were obtained from a weather station near the experimental site. In the current study, there is one dependent variable (SSP) and nine independent variables which are T_o, RH, WS, SR, T_F, T_B, TDS_F, TDS_B, and M_F.

Adaptive neuro-fuzzy inference system (ANFIS)

The adaptive neuro-fuzzy inference system (ANFIS) that best incorporates the features of FL and ANN systems is defined by Jang (1993). ANFIS is composed of if-then-else rules and input-output data coupled to a fuzzy set, and uses neural network learning algorithms for training. Moreover, ANFIS simulates complex nonlinear mappings using neural network learning and fuzzy inference methodologies,

and has the capability of working with uncertain, noisy, and inaccurate environments. The ANFIS utilizes the ANN training process to adjust the MF and the associated parameter that approaches the desired data sets. ANFIS's learning algorithm is a hybrid comprising a back-propagation learning algorithm together with a least squares method (Liu & Ling 2003; Inan *et al.* 2007; Wu *et al.* 2009; Svalina *et al.* 2013). To simplify the process, a sample with just two inputs and one output is considered. Five layers were used to build an ANFIS of the first-order Sugeno-type inference system, as presented in Figure 2. Two inputs, x and y, and one output, f, along with two fuzzy if-then-else rules represent the example. In Figure 2, the circle and square show a fixed and adaptive node, respectively. The functions of each of the five layers are explained in the following sections. For a first-order Sugeno fuzzy model, two fuzzy if-then-else rules are as follows:

Rule1: If x is A_1 and y is B_1, then $f_1 = p_1x+q_1y+r_1$ (1)

Rule2: If x is A_2 and y is B_2, then $f_2 = p_2x+q_2y+r_2$ (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 fuzzification layer (layer of input nodes). Every node i is an adaptive node with a node output expressed by:

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

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

where μ_{A_i} and $\mu_{B_{i-2}}$ are the fuzzy MFs and commonly are chosen to be bell-shaped with maximum equal to 1 and minimum equal to 0, and can be computed by:

$$\mu_{A_i}(x), \quad \mu_{B_{i-2}}(y) = \frac{1}{1 + |(x - c_i)/a_i|^{2b_i}} \tag{5}$$

where a_i, b_i, and c_i are the premise parameters of the MF.

Layer 2 is 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:

$$O_i^2 = w_i = \mu_{A_i}(x)\ \mu_{B_i}(y) \text{ for } i = 1, 2. \tag{6}$$

Layer 3 is normalization layer (layer of average nodes). In this layer, the i^{th} node is a circle labeled N, and computes the normalized firing strength as follows:

$$O_i^3 = \overline{w_i} = \frac{w_i}{w_1 + w_2}, \text{ for } i = 1,\ 2. \tag{7}$$

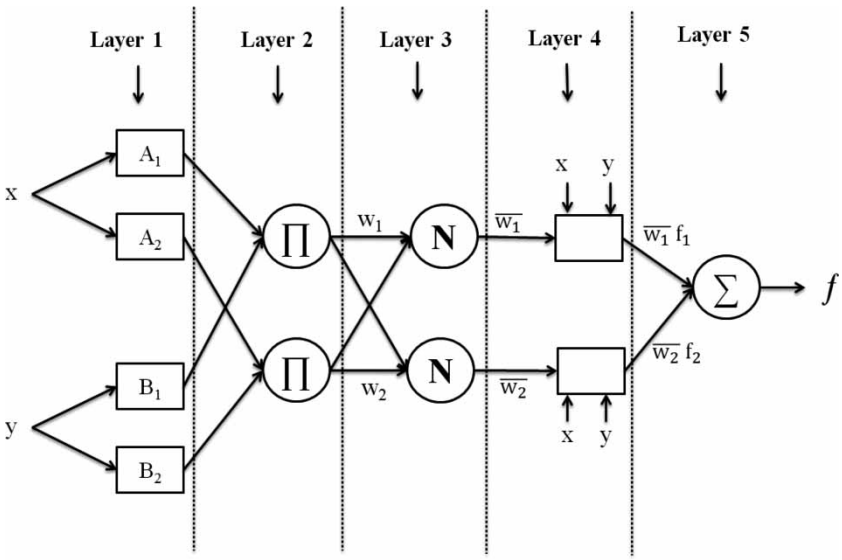

Figure 2 | General ANFIS architecture of two-input first-order Sugeno-type model with two rules.

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

$$O_i^4 = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i) \text{ for } i = 1,\ 2, \tag{8}$$

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

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

$$O_i^5 = f_{out} = \sum_{i=1}^{2} \overline{w_i} f_i = \frac{\sum_{i=1}^{2} w_i f_i}{w_1 + w_2} \tag{9}$$

Finally, the overall output can be formulated as:

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

Substituting Equation (7) into Equation (10):

$$f_{out} = \overline{w_1} f_1 + \overline{w_2} f_2 \tag{11}$$

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

The final output can be written as:

$$f_{out} = (\overline{w_1} x) p_1 + (\overline{w_1} y) q_1 + (\overline{w_1}) r_1 + (\overline{w_2} x) p_2 + (\overline{w_2} y) q_2 + (\overline{w_2}) r_2 \tag{13}$$

ANFIS model development

In the current study, the data obtained from the experiment were randomly divided into three portions: 70% as the training data sets (112 data points) for the learning process, 20% of 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 were 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 applied data were normalized using the following formula:

$$X_n = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{14}$$

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

MATLAB software (MATLAB 8.1.0.604, R2013a, the MathWorks Inc., USA) was used to develop an ANFIS model from the available experimental data to predict SSP. The Sugeno-type fuzzy inference system was used in the modeling of SSP. The grid partition method was used to classify the input data and in making the rules (Jang & Sun 1995). In this investigation, we employed four different types of input MFs: (1) built-in MF composed of the DSIGMF (difference between two sigmoidal MFs); (2) TRIMF (triangular-shaped built-in MF); (3) TRAMF (trapezoidal-shaped built-in MF); (4) PIMF (π-shaped built-in MF). 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 was used 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.

ANFIS models' performance evaluation

In this paper, the performance of the ANFIS models developed was evaluated using different standard statistical performance evaluation criteria. The statistical measures considered were correlation coefficient (CC), root mean square error (RMSE), model efficiency (ME), overall index of model performance (OI), coefficient of residual mass (CRM), and mean absolute percentage error (MAPE):

$$CC = \frac{\sum_{i=1}^{n} (SSP_{o,i} - \overline{SSP_o})(SSP_{p,i} - \overline{SSP_p})}{\sqrt{\sum_{i=1}^{n} (SSP_{o,i} - \overline{SSP_o})^2 \times \sum_{i=1}^{n} (SSP_{p,i} - \overline{SSP_p})^2}} \tag{15}$$

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

$$ME = 1 - \frac{\sum_{i=1}^{n} (SSP_{o,i} - SSP_{p,i})^2}{\sum_{i=1}^{n} (SSP_{o,i} - \overline{SSP_o})^2} \tag{17}$$

$$OI = \frac{1}{2}\left(1 - \frac{RMSE}{SSP_{max} - SSP_{min}} + ME\right) \quad (18)$$

$$CRM = \frac{\left(\sum_{i=1}^{n} SSP_{p,i} - \sum_{i=1}^{n} SSP_{o,i}\right)}{\sum_{i=1}^{n} SSP_{o,i}} \quad (19)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{SSP_{o,i} - SSP_{p,i}}{SSP_{o,i}}\right| \times 100 \quad (20)$$

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

RESULTS AND DISCUSSION

Independent parameters' selection

In this study, field data obtained from the experimental work were used for the training, testing, and validation of the ANFIS models. One of the most important steps in the modeling process for satisfactory prediction of 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 was performed to evaluate relationships between the dependent parameter (SSP) and the independent parameters (T_o, RH, WS, SR, T_F, T_B, M_F, TDS_F, and TDS_B), as presented 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 ANFIS models. Also, 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 CC of +0.734. Furthermore, SSP is found to be well correlated with TDS_F 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). Furthermore, there is a significant correlation between SSP and M_F and TDS_B with CC = 0.25 and −0.172, respectively. On the other hand, a very weak correlation is found between SSP and T_o and T_F 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 T_B due to their high collinearity with other parameters, although 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 CD. The same argument was also invoked to consider RH as an input

Table 1 | CC matrix for studied parameters

	T_o	RH	WS	SR	T_F	T_B	M_F	TDS_F	TDS_B	SSP
T_o	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
T_F	0.91	−0.80	−0.01	−0.09	1.00	0.13	0.48	0.06	−0.11	−0.06
T_B	0.06	0.05	0.33	0.82	0.13	1.00	−0.40	0.49	0.57	0.40
M_F	0.44	−0.72	−0.34	−0.27	0.48	−0.40	1.00	−0.75	−0.84	0.25
TDS_F	−0.01	0.23	0.64	0.22	0.06	0.49	−0.75	1.00	0.94	−0.40
TDS_B	−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

T_o: ambient temperature; RH: relative humidity; WS: wind speed; SR: solar radiation; T_F: temperature of feed water; T_B: temperature of brine water; M_F: feed flow rate; TDS_F: total dissolved solids of feed; TDS_B: total dissolved solids of brine; SSP: solar still productivity.

Table 2 | Summary statistics for input and output parameters

Parameter	Type	Unit	Symbol	Min	Max	Avg	SD	CV
Relative humidity	Input	%	RH	12.90	70.00	23.36	12.90	0.55
Solar radiation	Input	W/m^2	SR	75.10	920.69	587.55	181.93	0.31
Feed flow rate	Input	L/min	M_F	0.13	0.25	0.21	0.04	0.20
Total dissolved solids of feed	Input	PPT	TDS_F	41.40	130.00	80.23	29.42	0.37
Total dissolved solids of brine	Input	PPT	TDS_B	46.20	132.80	95.54	29.59	0.31
SSP	Output	$L/m^2/h$	SSP	0.05	0.97	0.50	0.24	0.48

Min: minimum value; Max: maximum value; Avg: average value; SD: standard deviation; CV: coefficient of variation.

parameter with low CC. The input and output parameters used in these models, and the range of them, are listed in Table 2.

However, model (variable/parameter) selection is the process of selecting an optimum subset of input parameters from the set of potentially useful parameters, which may be available for a given problem. In this study, we also used the stepwise procedure of the regression analysis and several statistics, such as Akaike's information criterion (AIC) and Schwarz's Bayesian information criterion (BIC) for the variables' selection to ensure accuracy. The subset/scenario with the lowest values based on these two criteria is selected as the best subset. The result of stepwise regression according to AIC and BIC values is presented in Table 3. The values −881.913 and −863.462 for AIC and BIC, respectively, are the smallest compared to all other subsets/scenarios. According to findings displayed in Table 3, the best subset includes the parameters *SR*, TDS_F, TDS_B, M_F, and *WS*. The parameters T_o, *RH*, T_F, and T_B were found to be insignificant and then excluded from predicting the SSP.

Table 3 | Results of parameter selection process

Step	Subsets (Models)	Akaike information criterion (AIC)	Schwarz Bayesian criterion (BIC)
1	*SR*	−576.163	−570.012
2	*SR*, TDS_F	−779.492	−770.267
3	*SR*, TDS_F, TDS_B	−810.278	−797.977
4	*SR*, TDS_F, TDS_B, M_F	−878.532	−863.156
5	*SR*, TDS_F, TDS_B, M_F, *WS*	−881.913	−863.462

WS: wind speed; *SR*: solar radiation; M_F: feed flow rate; TDS_F: total dissolved solids of feed; TDS_B: total dissolved solids of brine.

ANFIS model structure

Four ANFIS models were developed with four different types of input MFs. The MFs used were TRIMF, TRAMF, PIMF, and DSIGMF. The ANFIS models developed have five inputs (*RH*, *SR*, M_F, TDS_F, and TDS_B) and one output (SSP). However, the structure of the ANFIS models with the five input parameters is displayed in Figure 3. Therefore, in the input layer, five neurons were incorporated. For each of the neurons, the three same MFs were considered with three linguistic terms {low, medium, high} and accordingly, 243 ($3\times3\times3\times3\times3$) rules were developed for implementation of the ANFIS model. The number of nodes was 524, and of linear parameters was 1,458 for all four models. The number of nonlinear parameters was 60, and of parameters was 1,518 in the models with TRAMF, PIMF, and DSIGMF, while the number of nonlinear parameters was 45, and of parameters was 1,503 in the model with TRIMF. However, the four ANFIS models were trained, tested, and validated to assess the predictability of SSP by ANFIS-based models. The next sections discuss and evaluate the performance of the four developed ANFIS models in the training, testing, and validation stages by the statistical performance indicators.

Training process

Figure 4 shows the relationship between the predicted and observed values of SSP by using different MFs during the training phase. Figure 4 incorporates four graphs representing the four MFs used to compute SSP. It is clear from the figure that the predicted SSP values by using the four MFs,

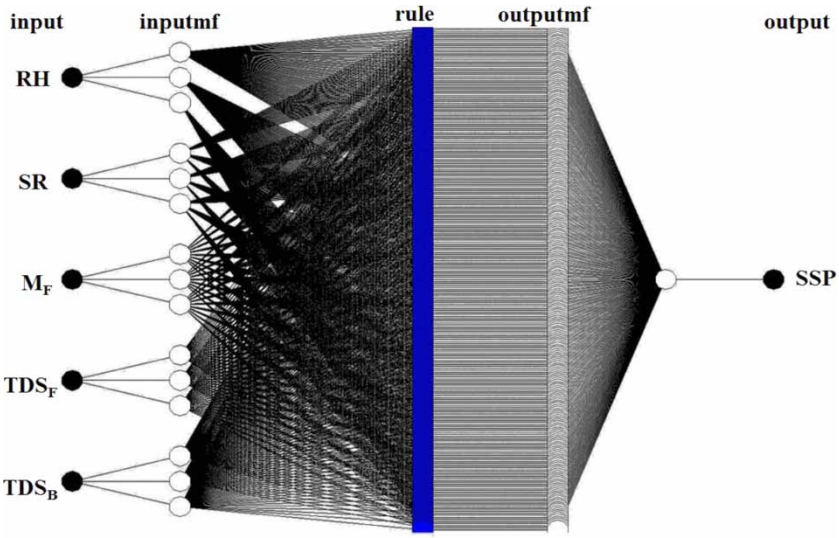

Figure 3 | ANFIS model structure used in this study (*RH*: relative humidity; *SR*: solar radiation; M_F: feed flow rate; TDS_F: total dissolved solids of feed; TDS_B: total dissolved solids of brine; SSP: solar still productivity).

namely, TRIMF, TRAPMF, PIMF, and DSIGMF, are in excellent agreement with the observed SSP values. Also, Table 4 shows the results of the statistical parameters CC, RMSE, ME, OI, CRM, and MAPE, which are numerical indicators used to evaluate the agreement between the observed and predicted SSP. The average values of CC, RMSE, ME, OI, CRM, and MAPE are 99.99%, 0.0034 L/m^2/h, 99.97%, 99.80%, 7.31×10^{-5}, and 0.6125% respectively, for the developed ANFIS models in the training process.

From Table 4, it can be seen that the values of RMSE, CRM, and MAPE are very low, and the values of CC, ME, and OI are very high for SSP predicted from the TRIMF, TRAPMF, PIMF, and DSIGMF MFs during the training process. The CC, ME, and OI values are very close to 1 while RMSE, CRM, and MAPE values are close to 0, showing excellent agreement between the observed results and predicted results from the ANFIS models. The tiny deviations between observed and predicted results, in turn, highlight the effectiveness of the ANFIS technique in the prediction process of SSP. This agrees with the results of Inan *et al.* (2007) and Yaïci & Entchev (2016). The performance, in the training stage, for all MFs is approximately the same, but in relative terms, the highest performance is obtained with DSIGMF in the training process. However, the statistical values between observed and predicted values were excellent with all MFs. These findings emphasize the accuracy and efficiency of ANFIS models for estimating SSP by using the four MFs.

Testing process

Figure 5 displays the 1:1 relationship between the predicted and observed values of SSP by using the developed ANFIS models during the testing stage and comprises four graphs representing the four MFs employed to develop the ANFIS models. The figure clearly reveals that the SSP values forecasted by using the ANFIS models, coupled with the four MFs, are in good agreement with the observed SSP values. The statistical performance of the ANFIS models with different MFs is shown in Table 4 during the testing process. As indicated in Table 4, the ANFIS models' CC values ranged from 89.96% to 92.27%, RMSE values from 0.0871 to 0.1109 L/m^2/h, ME values from 76.19% to 85.32%, OI ranged from 81.40% to 87.41%, CRM ranged from −0.0559 to 0.0156, and MAPE values from 12.65% to 17.32%, in the testing process.

Additionally, from Table 4, the comparisons reveal that the performance of the ANFIS model with PIMF is better than that with other MFs during the testing process, where its statistics were CC = 93.67%, RMSE = 0.0871 L/m^2/h,

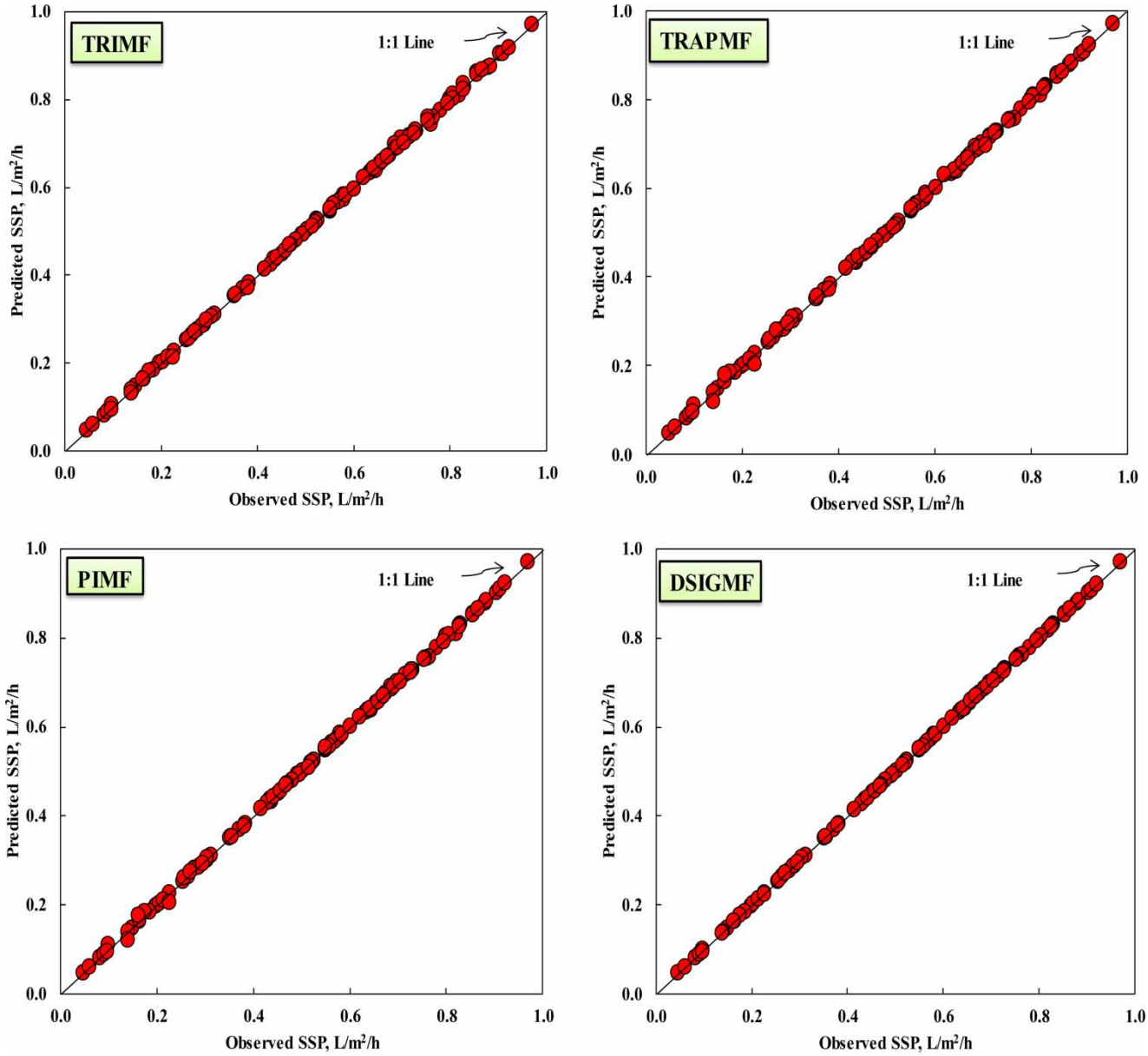

Figure 4 | Predicted versus observed solar still production (SSP) during the training process by using different MFs.

ME = 85.32%, OI = 87.41%, CRM = 0.0107, and MAPE = 14.80%. For PIMF, the CC, ME, and OI values are close to 1 while RMSE, CRM, and MAPE values are close to 0, showing a very good agreement between the observed and predicted values. On the other hand, the relatively lowest performance has the ANFIS model with TRIMF with the relatively highest RMSE, CRM, and MAPE values of 0.1109 L/m^2/h, −0.0559, 17.32%, respectively, and the relatively lowest CC, ME, and OI values of 89.96%, 76.19%, and 81.40%, respectively, in the testing stage. However, the difference between the highest (PIMF) and lowest (TRIMF) performances was relatively small during the testing process.

Validation process

A good agreement is obtained between the observed and predicted SSP with ANFIS models, especially calculated

Table 4 | Statistical performance of the developed ANFIS models with different types of input MFs during training, testing, and validation stages

		MFs			
Stages	**Statistical parameters**	**TRIMF**	**TRAPMF**	**PIMF**	**DSIGMF**
Training	CC	0.9999	0.9998	0.9999	0.9999
	RMSE	0.0040	0.0048	0.0041	0.0007
	ME	0.9997	0.9996	0.9997	0.99999
	OI	0.9977	0.9972	0.9976	0.9996
	CRM	−7.47E−06	0.0001	0.0001	0.0001
	MAPE	0.6376	0.8857	0.7844	0.1424
Testing	CC	0.8996	0.9265	0.9367	0.9227
	RMSE	0.1109	0.0896	0.0871	0.0961
	ME	0.7619	0.8446	0.8532	0.8211
	OI	0.8140	0.8682	0.8741	0.8525
	CRM	−0.0559	0.0156	0.0107	0.0131
	MAPE	17.3199	12.6466	14.8029	16.5259
Validation	CC	0.8861	0.7950	0.8458	0.7629
	RMSE	0.1064	0.1306	0.1195	0.1541
	ME	0.6920	0.5361	0.6110	0.3538
	OI	0.7564	0.6581	0.7049	0.5472
	CRM	−0.0273	0.0596	0.0307	0.0775
	MAPE	16.9053	14.5901	14.5509	18.2047

CC: correlation coefficient; RMSE: root mean square error; ME: model efficiency; OI: overall index of model performance; CRM: coefficient of residual mass; MAPE: mean absolute percentage error.

by the TRIMF, as shown by the 1:1 curve in Figure 6. The performances of the ANFIS models with different MFs are listed in Table 4. This good agreement is clearly reflected by the values of the statistical parameters presented in Table 4. The ANFIS model using TRIMF had a CC value that was about 10.28%, 4.55%, and 13.90% more accurate than that from the ANFIS models with TRAPMF, PIMF, and DSIGMF, respectively. These results are in conformity with the findings of Taghavifar & Mardani (2015) and Jiang *et al.* (2016). According to Table 4, the RMSE values indicate that the TRIMF is the most accurate. The RMSE value for TRAPMF in the ANFIS model (0.1306 L/m^2/h) is 1.23 times that given by the TRIMF (0.1064 L/m^2/h), while the value of 0.1195 L/m^2/h for PIMF is 1.12 times the corresponding TRIMF value. Also, the RMSE value for the TRIMF was closer to zero than for the DSIGMF, where the DSIGMF RMSE value was almost 1.5 times that of the value for the TRIMF model. However, the best RMSE value was achieved by the ANFIS model with TRIMF, followed by, respectively, the ANFIS model with PIMF, TRAPMF, and DSIGMF.

The ME and OI values of the ANFIS models with different MFs are 69.20% and 75.64% for TRIMF, 53.61% and 65.81% for TRAPMF, 61.10% and 70.49% for PIMF, and 35.38% and 54.72% for DSIGMF, respectively, during the validation process. The ME values for the ANFIS model with TRIMF were, respectively, 22.53%, 11.71%, and 48.87% more accurate than those from TRAPMF, PIMF, and DSIGMF. The OI value for the ANFIS model with TRIMF was closer to 1 than its values for the other MFs. The best MAPE value (14.55%) was achieved by the ANFIS model with PIMF. The CRM values for TRAPMF, PIMF, and DSIGMF were, respectively, 2.18, 1.12, and 2.84 times that of the value for the ANFIS model with TRIMF. This means that the TRIMF has a lower residual mass coefficient and gives more accurate estimation values of SSP. Moreover, the CRM (−0.0273) value for the ANFIS model with TRIMF is the closest to 0 compared to other MFs. This indicates good agreement between the observed results and predicted results from the ANFIS, especially with TRIMF, and confirms that the ANFIS model with TRIMF performed better than that with other MFs when using the validation data set.

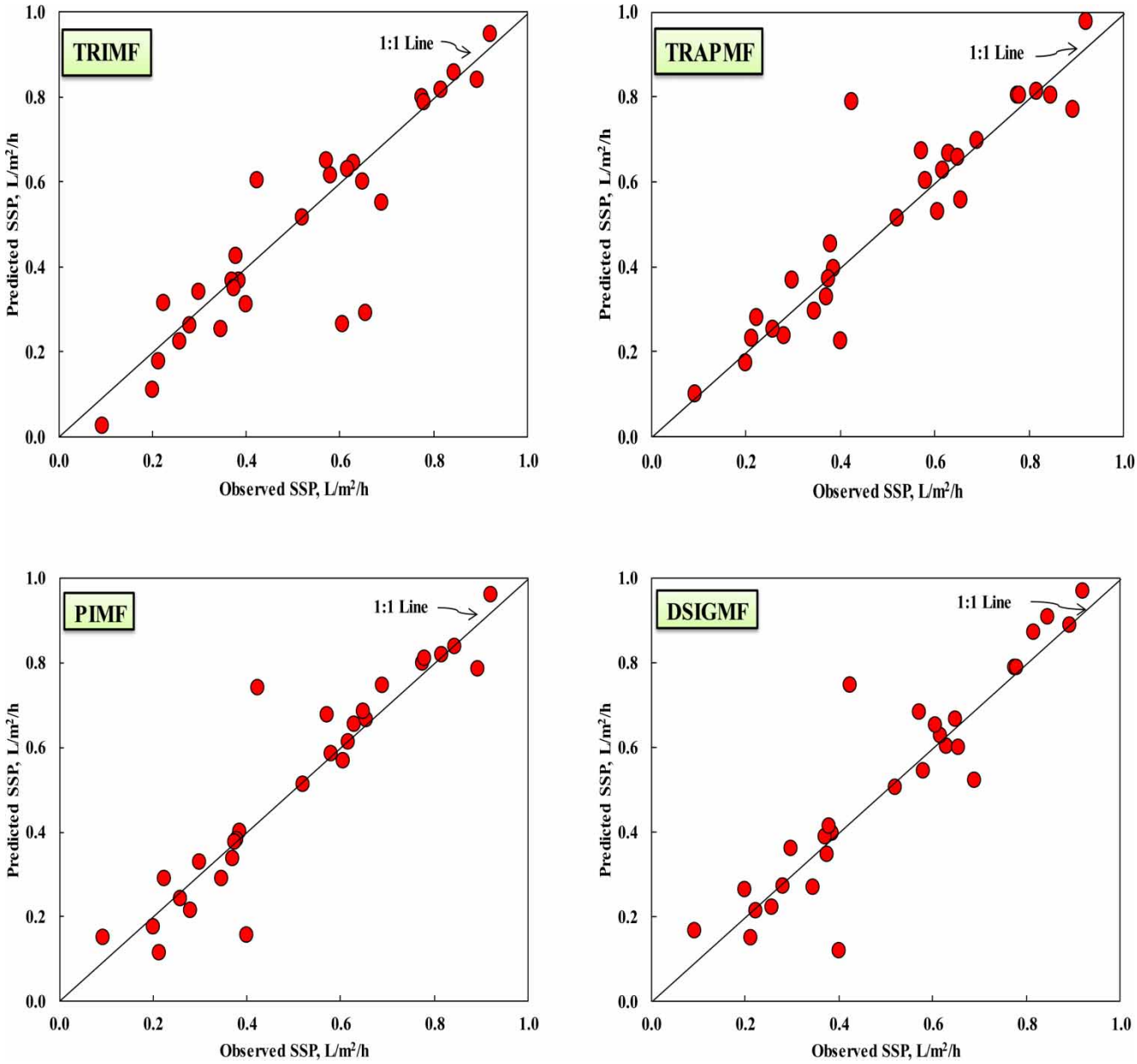

Figure 5 | Predicted versus observed solar still production (SSP) during the testing process by using different MFs.

Overall ANFIS models' performance

Overall, in all modeling stages, the model with the highest prediction capability was ANFIS with PIMF, followed by TRIMF, TRAPMF, and DSIGMF. Generally, the best value of CC was achieved with TRIMF, while the best values of RMSE, ME, OI, and CRM were achieved with PIMF. The best MAPE values were achieved by TRAPMF and PIMF. The CRM values for the ANFIS model with TRIMF in the training, testing, and validation were negative, showing under-prediction by the model, while they were positive for the ANFIS model with the other MFs, indicating over-prediction. Mostly, the use of ANFIS models with different MFs

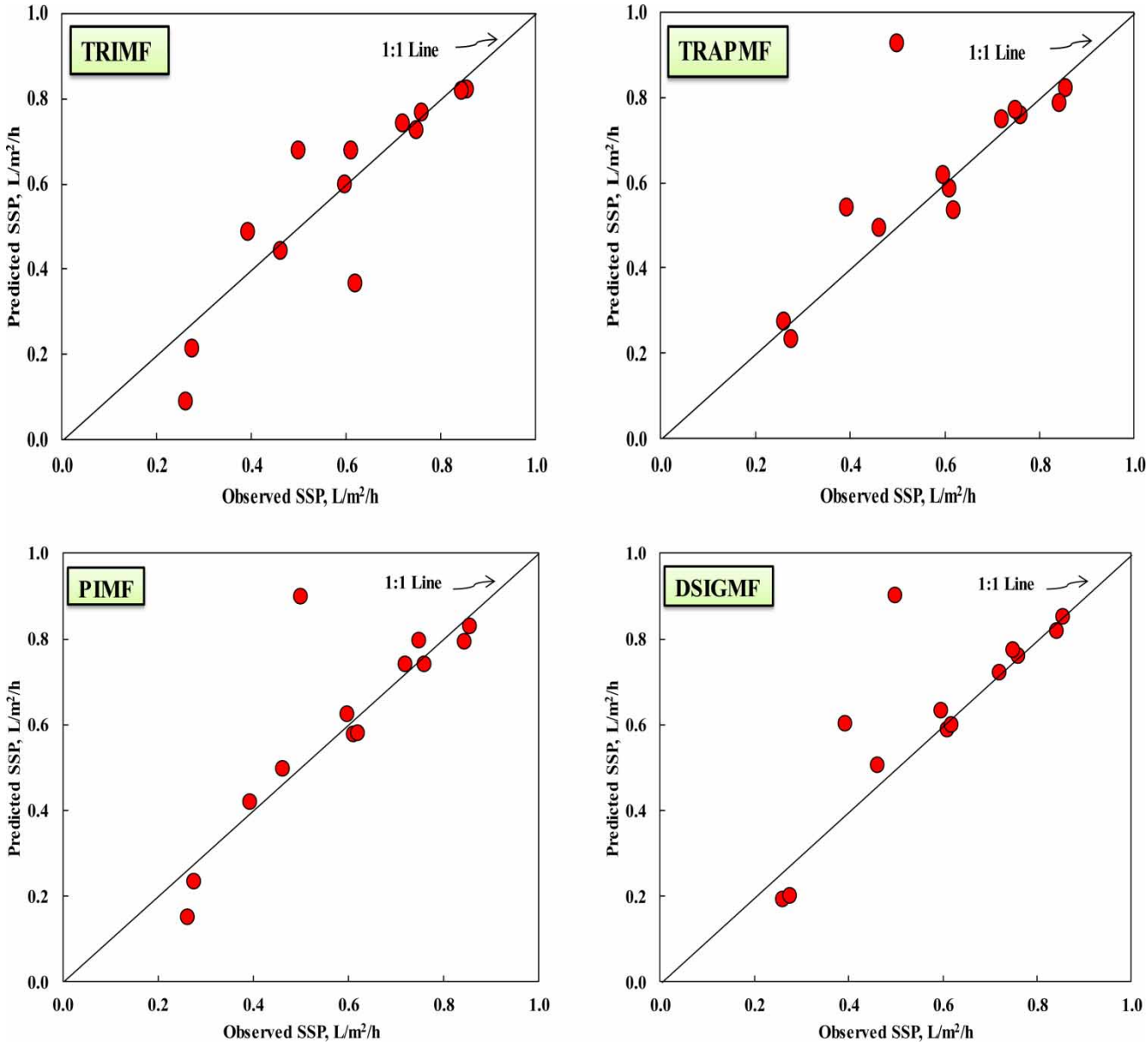

Figure 6 | Predicted versus observed solar still production (SSP) during the validation process by using different MFs.

gives high accuracy in the modeling process for SSP, and this is reflected by the values of the statistical parameters displayed in Table 4. However, as seen from the complete findings in Figures 4–6 and Table 4, the ANFIS models are reliable and powerful tools for predicting SSP, and they provide us with results higher in performance and accuracy. These satisfactory results confirm the promising ability of the ANFIS technique for SSP forecasting.

Sensitivity analysis

To explain the effect of each input variable (*WS*, *SR*, M_F, TDS_F, and TDS_B) on SSP, it was necessary to perform a sensitivity analysis. Sensitivity analysis was performed using the best ANFIS model (PIMF) to deduce the effects of input variables on output (SSP), as depicted in Figure 7. As can be seen, the most effective variable on the SSP is *SR*, whereas *RH* is the least effective variable. The results

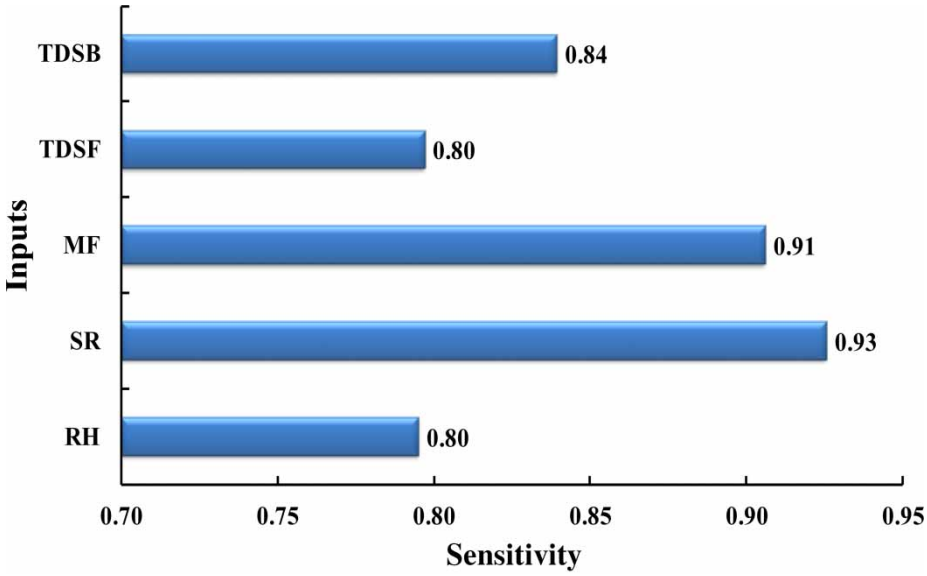

Figure 7 | Sensitivity analysis for input variables using the best ANFIS model (*RH*: relative humidity; *SR*: solar radiation; M_F: feed flow rate; TDS_F: total dissolved solids of feed; TDS_B: total dissolved solids of brine).

show that the *SR* is the most influential variable on SSP, followed by M_F, TDS_B, TDS_F and *RH*.

CONCLUSIONS

In the current investigation, the ANFIS modeling method was assessed for predicting SSP. The data were obtained through an experiment using a solar still to desalinate seawater. Various ANFIS models with different MFs were trained, tested, and validated to determine the best ANFIS model for SSP modeling. The ANFIS models were trained using 70% of the available data, tested using 20% of the data, and validated using the remaining 10%. In the modeling process, five input parameters were used, namely, *RH*, *SR*, M_F, TDS_F, and TDS_B. We used four different types of MF in this investigation, namely, a triangle (TRIMF), a trapezoid (TRAPMF), a Pi curve (PIMF), and the difference between two sigmoidal functions (DSIGMF). The performance of the ANFIS models was evaluated by the CC, RMSE, ME, OI, CRM, and MAPE performance indicators. Generally, the results reveal the high accuracy, effectiveness, and reliability of the ANFIS method for modeling SSP. On the basis of statistical performance criteria, it was found that PIMF and TRIMF were the best MFs in all modeling stages for ANFIS models. Overall, the highest performance is obtained with PIMF in the modeling process. The results also shed light on some principles in view of ANFIS-based model application for estimating SSP using meteorological and operational variables. It is hoped that the results will serve as a reference for future attempts to assess the used parameters in this investigation using other soft computing methodologies to forecast SSP.

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.

Ayoub, G. M., Al-Hindi, M. & Malaeb, L. 2015 A solar still desalination system with enhanced productivity. *Desalin. Water Treat.* **53** (12), 3179–3186.

Barua, S., Perera, B. J. C., Ng, A. W. M. & Tran, D. 2010 Drought forecasting using an aggregated drought index and artificial neural network. *J. Water Clim. Change* **1** (3), 193–206.

Dumitru, C. D., Gligor, A. & Enachescu, C. 2016 Solar photovoltaic energy production forecast using neural networks. *Procedia Technol.* **22**, 808–815.

El Shafie, A. H., El-Shafie, A., Almukhtar, A., Taha, M. R., El Mazoghi, H. G. & Shehata, A. 2012 Radial basis function neural networks for reliably forecasting rainfall. *J. Water Clim. Change* **3** (2), 125–138.

EU 2007 Addressing the Challenge of Water Scarcity and Droughts in the European Union, Communication from the Commission to the European Parliament and the Council, Eur. Comm., DG Environ., Brussels.

Gairaa, K., Khellaf, A., Messlem, Y. & Chellali, F. 2016 Estimation of the daily global solar radiation based on Box–Jenkins and ANN models: a combined approach. *Renew. Sust. Energ. Rev.* **57**, 238–249.

Inan, G., Göktepe, A. B., Ramyar, K. & Sezer, A. 2007 Prediction of sulfate expansion of PC motor using adaptive neuro-fuzzy methodology. *Build. Environ.* **42**, 1264–1269.

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.

Jiang, Z., Zheng, H., Mantri, N., Qi, Z., Zhang, X., Hou, Z., Chang, J., Lu, H. & Liang, Z. 2016 Prediction of relationship between surface area, temperature, storage time and ascorbic acid retention of fresh-cut pineapple using adaptive neuro-fuzzy inference system (ANFIS). *Postharvest Biol. Tec.* **113**, 1–7.

Kabeel, A. E., Hamed, M. H. & Omara, Z. M. 2012 Augmentation of the basin type solar still using photovoltaic powered turbulence system. *Desalin. Water Treat.* **48** (1–3), 182–190.

Khoi, D. N. & Suetsugi, T. 2012 Hydrologic response to climate change: a case study for the Be River Catchment, Vietnam. *J. Water Clim. Change* **3** (3), 207–224.

Liu, M. & Ling, Y. Y. 2003 Using fuzzy neural network approach to estimate contractors' markup. *Build. Environ.* **38**, 1303–1308.

Maachou, R., Lefkir, A., Khouider, A. & Bermad, A. 2016 Modeling of activated sludge process using artificial neuro-fuzzy-inference system (ANFIS). *Desalin. Water Treat.* **57** (45), 21182–21188.

Mashaly, A. F. & Alazba, A. A. 2015 Comparative investigation of artificial neural network learning algorithms for modeling solar still production. *J. Water Reuse Desal.* **5** (4), 480–493.

Mashaly, A. F. & Alazba, A. A. 2016a MLP and MLR models for instantaneous thermal efficiency prediction of solar still under hyper-arid environment. *Comput. Electron. Agric.* **122**, 146–155.

Mashaly, A. F. & Alazba, A. A. 2016b 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. 2016c Neural network approach for predicting solar still production using agricultural drainage as a feedwater source. *Desalin. Water Treat.* **57** (59), 28646–28660.

Mashaly, A. F. & Alazba, A. A. 2017 Artificial intelligence for predicting solar still production and comparison with stepwise regression under arid climate. *J. Water Supply Res. T.* **66** (3), 166–177.

Mashaly, A. F., Alazba, A. A., Al-Awaadh, A. M. & Mattar, M. A. 2015 Predictive model for assessing and optimizing solar still performance using artificial neural network under hyper arid environment. *Sol. Energy* **118**, 41–58.

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.

Rizwan, M., Jamil, M., Kirmani, S. & Kothari, D. P. 2014 Fuzzy logic based modeling and estimation of global solar energy using meteorological parameters. *Energy* **70**, 685–691.

Santos, N. I., Said, A. M., James, D. E. & Venkatesh, N. H. 2012 Modeling solar still production using local weather data and artificial neural networks. *Renew. Energ.* **40**, 71–79.

Sun, Y., Tang, D., Dai, H., Liu, P., Sun, Y., Cui, Q., Zhang, F., Ding, Y., Wei, Y., Zhang, J., Wang, M., Li, A. & Meng, Z. 2015 A real-time operation of the three gorges reservoir with flood risk analysis. *Water Sci. Tech. Water Supply* **16** (2), 551–562. doi:10.2166/ws.2015.172.

Svalina, I., Galzina, V., Lujić, R. & Šimunović, G. 2013 An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: the case of close price indices. *Expert Syst. Appl.* **40** (15), 6055–6063.

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.

Wu, J.-D., Hsu, C.-C. & Chen, H.-C. 2009 An expert system of price forecasting for used cars using adaptive neuro-fuzzy inference. *Expert Syst. Appl.* **36**, 7809–7817.

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 21 December 2016; accepted in revised form 12 May 2017. Available online 30 June 2017