Applied Thermal Engineering 96 (2016) 311–319

Contents lists available at ScienceDirect

Applied Thermal Engineering

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

Research Paper

Using ANFIS for selection of more relevant parameters to predict dew point temperature

Kasra Mohammadi [a,*], Shahaboddin Shamshirband [b,*], Dalibor Petković [c], Por Lip Yee [d], Zulkefli Mansor [e]

[a] Department of Mechanical and Industrial Engineering, University of Massachusetts, Amherst, Massachusetts 01003, USA
[b] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] Faculty of Mechanical Engineering, Department for Mechatronics and Control, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia
[d] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[e] Research Center for Software Technology and Management, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

HIGHLIGHTS

- ANFIS is used to select the most relevant variables for dew point temperature prediction.
- Two cities from the central and south central parts of Iran are selected as case studies.
- Influence of 5 parameters on dew point temperature is evaluated.
- Appropriate selection of input variables has a notable effect on prediction.
- Considering the most relevant combination of 2 parameters would be more suitable.

ARTICLE INFO

Article history:
Received 10 April 2015
Accepted 24 November 2015
Available online 2 December 2015

Keywords:
Dew point temperature
ANFIS
Variable selection
Prediction

ABSTRACT

In this research work, for the first time, the adaptive neuro fuzzy inference system (ANFIS) is employed to propose an approach for identifying the most significant parameters for prediction of daily dew point temperature (T_{dew}). The ANFIS process for variable selection is implemented, which includes a number of ways to recognize the parameters offering favorable predictions. According to the physical factors influencing the dew formation, 8 variables of daily minimum, maximum and average air temperatures (T_{min}, T_{max} and T_{avg}), relative humidity (R_h), atmospheric pressure (P), water vapor pressure (V_P), sunshine hour (n) and horizontal global solar radiation (H) are considered to investigate their effects on T_{dew}. The used data include 7 years daily measured data of two Iranian cities located in the central and south central parts of the country. The results indicate that despite climate difference between the considered case studies, for both stations, V_P is the most influential variable while R_h is the least relevant element. Furthermore, the combination of T_{min} and V_P is recognized as the most influential set to predict T_{dew}. The conducted examinations show that there is a remarkable difference between the errors achieved for most and less relevant input parameters, which highlights the importance of appropriate selection of input parameters. The use of more than two inputs may not be advisable and appropriate; thus, considering the most relevant combination of 2 parameters would be more suitable to achieve higher accuracy and lower complexity in predictions. In the final step, comparisons between the predictions of the ANFIS model using the selected inputs and other soft computing techniques demonstrate that ANFIS has a higher accuracy to predict daily dew point temperature.

1. Introduction

Accessibility to the accurate dew point temperature data are of particular importance in different scientific fields such hydrology, climatology and agronomy. Dew point temperature is the temperature in which air must be cooled down at a constant pressure to reach saturation. In fact, it is the temperature at which water vapor

* Corresponding author. Tel.: +60146266763; fax: +60146266763.
E-mail address: kmohammadi@umass.edu (K. Mohammadi); shamshirband@um.edu.my (S. Shamshirband).

http://dx.doi.org/10.1016/j.applthermaleng.2015.11.081

in the air condenses into liquid water. It is also described as the temperature at which the saturation and actual vapor pressures are equal [1]. The dew formation occurs once the surface air temperature decline to the dew point temperature. The radiation exchange between the surface of the Earth and the atmosphere, water vapor pressure as well as turbulent heat are among the major elements influencing dew formation [2,3].

The precise prediction of the dew point temperature would be of indispensable significance for a variety of purposes. Dew point temperature along with relative humidity is typically utilized to identify the level of moisture in the air. It can also be utilized in conjunction with wet bulb-temperature for computing the ambient temperature, which provides the possibility for being prepared against the potential frosts, which may harm crops [1,4]. Dew point temperature can be used to provide a favorable estimate of the near-surface humidity that influences the stomatal closure in plants where a low level humidity may result in decline in the plants' productivity [1,5]. Dew would be really significant for plant survival particularly in the arid areas with rare rainfall [6]. Dew point temperature is an especially significant element in various hydrological and climatological models for the purpose of reference evapotranspiration estimation [1,7]. It may generally be stated that dew point temperature is an element that either explicitly or implicitly contributes to the plants' productivity, crop harm during freezes, loss of human life during heat waves as well as the levels of human comfort [1].

In recent years, soft computing techniques such as artificial neural network (ANN), genetic programming (GP), support vector machine (SVM) and adaptive neuro-fuzzy inference system (ANFIS) have been successfully employed in hydrology, climatology and agrometeorology related studies.

Jeong et al. [8] used the ANFIS technique for forecasting monthly precipitation in Korea. They performed a proper variable selection between climatological and hydrological elements to determine the 3 most important parameters for developing the ANFIS model. Their results specified that ANFIS is a promising technique to forecast precipitation. Kisi and Sanikhani [9] employed the ANFIS, ANN and support vector regression (SVR) techniques to predict monthly precipitation in Iran without the use of climatic data. They utilized longitude, latitude and altitude of 50 Iranian locations to develop the models. Their results demonstrated that ANFIS with grid partition (GP) can be used effectively to predict precipitation in Iran. Hamidi et al. [10] studied the performance of SVM and ANN approaches to model monthly precipitation in Hamedan city of Iran. Their results indicated that the SVM method, by presenting good predictions, shows higher performance compared to the ANN technique; thus, SVM was introduced as a useful technique for precipitation modeling. Citakoglu et al. [11] applied the ANFIS and ANN techniques to estimate monthly mean reference evapotranspiration in Turkey. They used different combinations of climatic data as input data. Their results demonstrate that both techniques provide accurate estimates, but further precision can be achieved by ANFIS. Tabari et al. [12] utilized the ANN technique to forecast 1 day ahead the soil temperature at 6 different depths in humid and arid locations of Iran. Their results illustrated that ANN is an efficient method to offer precise short-term forecasts of soil temperature. Kisi et al. [13] evaluated the precision of three approaches of multi-layer perceptron (MLP), radial basis neural networks (RBNN) and generalized regression neural networks (GRNN) to model soil temperatures at different depths for Turkey. They found that air temperature is the most significant parameters to estimate soil temperature. Also, GRNN is among the most accurate models for soil temperature estimation at different depths. Talaee [14] estimated daily soil temperature at 6 different depths based on the coactive neuro-fuzzy inference system (CANFIS) for two arid and semiarid Iranian stations. For this aim, 6 parameters were selected as input elements for the CANFIS model. The results indicated the adequacy of the CANFIS for soil temperature estimation. Cobaner et al. [15] used ANN, ANFIS and multiple linear regression (MLR) techniques to estimate maximum, minimum and average monthly air temperatures over Turkey. They considered latitude, longitude and altitude of the locations as well as the number of months as input variables. Their results demonstrated that ANFIS provides more accuracy to estimate air temperatures in Turkey.

Shank et al. [16] utilized ANN techniques for prediction of dew point temperature from 1 to 12 h ahead based upon the previous weather data sets. They used measured data of 20 stations in Georgia, USA for developing general models to predict dew point temperature in the whole state of Georgia. Zounemat-Kermani [17] evaluated the capability of multi linear regression (MLR) and Levenberg–Marquardt (LM) feed-forward neural network for estimation of hourly dew point temperature in a location in Ontario, Canada. It was found that LM–NN model provide further accuracy compared to the MLR model. Nadig et al. [18] developed combined air temperature and dew point temperature models using the ANN technique to provide an enhancement in the predictions of both temperatures. Their results demonstrated that the combined method decreased prediction error. Shiri et al. [3] assessed the capability of two ANN models and gene expression programming (GEP) technique to estimate daily dew point temperature in two stations of Korea. Their results indicated that the GEP model outperforms the ANN models. Kim et al. [19] utilized two soft computing techniques for estimation of daily dew point temperature in California, USA. By providing comparisons with a conventional regression model, they found that developed soft computing models are more precise in estimating daily dew point temperature.

Basically, proper selections of more significant input parameters for dew point temperature prediction to provide more precision and less complexity would be of indispensable significance. In fact, there can be drawbacks in the inclusion of many input variables. Some of the drawbacks would include the difficulty in explaining the model, inaccuracies caused by irrelevant parameters, complexity in the developed model due to high number of required inputs and time consuming task for collecting more data. These factors may consequently deteriorate the generalization capacity of the model.

To the best of our knowledge, there is no specific study on determining the most relevant variables that affect dew point temperature and analyzing the influence different combinations of variables on dew point temperature. In fact, the main originality of this research work is to identify the most significant parameters for dew point temperature prediction, which has not been conducted so far. The motivation behind performing this study is the significance of dew point temperature in various scientific fields as well as the importance of proper selections of input parameters for dew point temperature prediction. Therefore, in this research work, for the first time, the adaptive neuro-fuzzy inference system (ANFIS) is applied to select the most influential parameters influencing the daily dew point temperature. ANFIS is a hybrid intelligent system that merges the technique of the learning power of the ANNs with the knowledge representation of fuzzy logic. The major advantages of the ANFIS model are computationally efficiency and adaptability. The main aim of this study is to identify and introduce the most significant parameters for prediction of daily dew point temperature in two Iranian locations of Kerman and Tabass situated in the central and south central parts of Iran. The process, named variable selection, includes a number of ways to determine a subset of the total recorded parameters that show favorable capability of prediction. The ANFIS network is used to perform a variable search and thereafter, it is utilized to examine how eight important parameters of minimum, maximum and average air temperatures, relative humidity, atmospheric pressure, water vapor pressure, sunshine duration and global solar radiation influence dew point temperature prediction in two stations.

2. Materials and method

2.1. Data collection

For this research work, two Iranian cities of Kerman and Tabass, respectively located in the central and south central parts of the country have been nominated as case studies. Figure 1 illustrates the locations of Kerman and Tabass on the map of Iran. As the considered cities enjoy different weather conditions, further reliability in the conducted examinations may be expected. Kerman city as the center of Kerman province is situated in the south central part of Iran at the geographical location of 30°29′N and 57°06′E, and its elevation is 1756 m above sea level. The city of Kerman has a semi-moderate and dry climate [20]. Based upon the Köppen classification, the climate condition of Kerman City is categorized as BWk, which relates to arid desert cold [21]. Tabass City is situated in the central desert of Iran in the South-Khorasan province at the geographical location of 33°36′N and 56°55′E, and its elevation is 711 m above sea level. Tabass climate is generally characterized with hot summers and rare snowfall in the winters [22]. According to the Köppen classification, its climate is categorized as BWh, which relates to arid desert hot [21].

In this study, 7 years of measured data provided by the Iranian Meteorological Organization (IMO) for the period of 1998–2004 have been utilized. Based upon the physical factors influencing the formation of dew, 8 input elements have been considered to assess their effects on the accurate prediction of daily dew point temperature. The used data sets consist of measured daily dew point temperature (T_{dew}), minimum air temperature (T_{min}), maximum air temperature (T_{max}), average air temperature (T_{avg}), relative humidity (R_h), atmospheric pressure (P), water vapor pressure (V_P), sunshine duration (n), and global solar radiation on a horizontal surface (H). The available data for this study were divided into two parts of training and checking data sets.

Fig. 1. Locations of Kerman and Tabass on the map of Iran (Kerman and South-Khorasan provinces have been highlighted).

2.2. Input and output variables

As stated earlier, identifying the most significant parameters, which are potentially influential in predicting the daily dew point temperature (T_{dew}) is the main objective of this research work. Eight important variables of T_{min} (input 1), T_{max} (input 2), T_{avg} (input 3), R_h (input 4), P (input 5), V_P (input 6), n (input 7) and H (input 8) were considered to assess their influence on prediction of T_{dew} (output). Tables 1 and 2 present some descriptive statistics including minimum, average and maximum values, standard deviation as well as the range of the 8 input parameters and the output parameters, respectively for Kerman and Tabass.

To build a system with the best characteristics, it is necessary to identify the most relevant and influential subset of parameters. This process of selection is usually called variable selection. The purpose of this process is to find a subset of the total set of parameters which offers favorable capability of prediction [23–26]. Essentially, with neural network (NN) as the foundation the complex system's architecture in the function of approximation and regression can be modeled. NN is an architecture that is made up of extremely parallel adaptive processing elements. These are interconnected through structured networks. Therefore, the accuracy of the NN models, which are created as a result of these data, relies heavily on the accuracy of the chosen input data in the representation of the system. To achieve a successful generation and creation of a model that is capable of estimating a special process output, the selection process of the subset of parameters that are really pertinent is crucial. This is achieved in the process of variable selection. The problems faced in the process of the selection of parameters could possibly be resolved by integrating and applying prior knowledge to segregate and remove parameters that are irrelevant.

Among many NN systems, the adaptive neuro-fuzzy inference system (ANFIS) is one of the most utilized and powerful. Thus, in this study, the ANFIS is employed to select the most influential variables [27,28].

To determine how the eight above-mentioned parameters influence the dew point temperature, a parameter search by employing the ANFIS was conducted. ANFIS [29], a hybrid intelligent system that increases the capability of learning and adapting automatically,

Table 1
Descriptive statistics for the input and output parameters for Kerman station.

Variable	Min	Max	Mean	St. dev	Range
T_{min} (°C)	−16.20	26.60	7.14	7.99	42.80
T_{max} (°C)	−4.00	42.00	25.16	8.86	46.00
T_{avg} (°C)	−10.10	32.10	16.15	8.07	42.20
R_h (%)	9.80	98.40	32.51	17.52	88.60
P (mb)	987.30	1033.30	1009.75	9.24	46.00
V_P (mb)	1.03	16.14	5.63	2.25	15.11
n (hr)	0.20	13.40	9.01	3.27	13.20
H (MJ/m^2)	0.90	33.46	20.75	7.01	32.56
T_{dew} (°C)	−21.80	14.20	−2.61	5.83	36.00

Table 2
Descriptive statistics for the input and output parameters for Tabass station.

Variable	Min	Max	Mean	St. dev	Range
T_{min} (°C)	−5.20	35.40	15.68	9.58	40.60
T_{max} (°C)	2.20	48.60	28.41	10.91	46.40
T_{avg} (°C)	0.70	40.50	22.04	10.15	39.80
R_h (%)	9.40	93.90	32.24	17.74	84.50
P (mb)	989.70	1037.10	1011.24	9.38	47.40
V_P (mb)	1.82	20.57	7.27	2.62	18.75
n (hr)	0.20	13.50	9.07	3.17	13.30
H (MJ/m^2)	1.28	37.65	20.15	7.16	36.37
T_{dew} (°C)	−12.60	18.00	1.83	4.91	30.60

has been used by researchers for many different purposes in a variety of engineering systems such as in modeling [30–33], for prediction [34–36] and for control [37–40]. This neuro-adaptive learning methodology allows the fuzzy modeling process to obtain information regarding the data gathered [41,42]. This is the foundational idea underlying all neuro-adaptive learning methodologies. The ANFIS methodology aims to organize the FIS (fuzzy inference system) by analyzing the input/output data pairs [43,44]. It gives fuzzy logic the ability to adjust the MF parameters so that it is optimal in allowing the associated FIS to detect and trace the given input/output data [45,46].

2.3. Variable selection using ANFIS

Generating predetermined input–output subsets requires the construction of a set of fuzzy 'IF THEN' rules with the suitable MFs (membership function). The ANFIS can serve as the foundation for such a construction. The input–output data are converted membership functions. In accordance to the collection of input–output data, the ANFIS takes the initial FIS and adjusts it through a back propagation algorithm. The FIS is comprised of three components: (1) a rule base, (2) a database and (3) a reasoning mechanism. The rule base consists of a choice of fuzzy rules. The database assigns the MFs, which are employed in the fuzzy rules. Finally, the last component is the reasoning mechanism and it infers from the rules and input data to come to a feasible outcome. These intelligent systems are a combination of knowledge, methods and techniques from a variety of different sources. They adjust to perform better in environments, which are changing. These systems have similar-human intelligence within a specific domain. The ANFIS recognizes patterns and assists in the revision of environments. FIS integrates human comprehension, does interfacing, and makes decisions.

FIS in MATLAB is employed in the whole process of the FIS training and evaluation. Figure 2 shows all inputs parameters in the ANFIS selection procedure. The ANFIS model should select the parameter or a set of parameters that are the most influential to the output (i.e. daily dew point temperature). In fact, the ANFIS network is trained for each input separately, and RMSE for the ANFIS predictions is observed for each input separately. It should be mentioned that training RMSE is a relevant indicator for determining the influence of inputs on the dew point temperature. While checking RMSE is used to track overfiting between the training and checking data. Furthermore, the ANFIS network is trained with combinations of two inputs and then achieved RMSEs are observed and the most influential combination of two inputs on the dew point temperature prediction is determined. As a sample, an ANFIS network for 2 input variables is depicted in Fig. 3 and then its description is presented in the following.

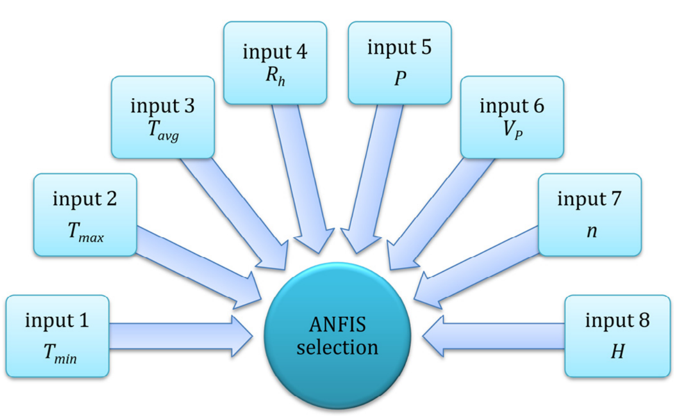

Fig. 2. Input parameters for ANFIS selection procedure.

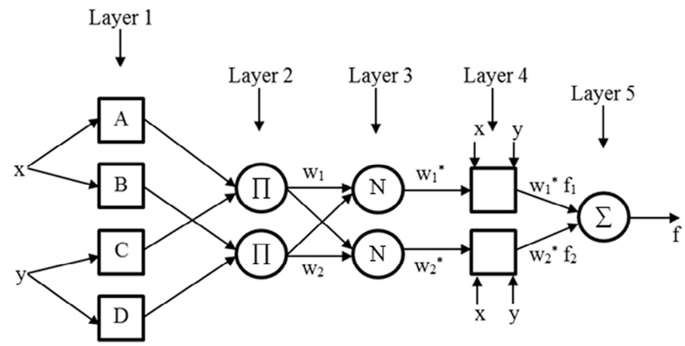

Fig. 3. ANFIS structure with two inputs, one output and two rules.

The fuzzy IF-THEN rules of Takagi and Sugeno's class and two inputs for the first-order Sugeno is employed for the purposes of this study:

$$\text{if } x \text{ is } A \text{ and } y \text{ is } C \text{ then } f_1 = p_1 x + q_1 y + r_1 \tag{1}$$

The 1st layer is made up of input parameters of MFs, and it provides the input values to the following layer. In another word, the MFs convert the crisped inputs in fuzzy values between 0 and 1. Each node here is considered an adaptive node having a node function $O = \mu_{AB}(x)$ and $O = \mu_{CD}(x)$ where $\mu_{AB}(x)$ and $\mu_{CD}(x)$ are membership functions. Bell-shaped membership functions having the maximum value (1.0) and the minimum value (0.0) are selected, such as:

$$\mu(x) = bell(x; a_i, b_i, c_i) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i}\right)^2\right]^{b_i}} \tag{2}$$

where $\{a_i, b_i, c_i\}$ is the set of parameters set. The bell-shaped MF has the best capabilities for the generalization of nonlinear data. The parameters of this layer are designated as premise parameters. Here, x and y are the inputs to nodes. They represent a combined version of the two most impactful variables on the dew point temperature.

The membership layer is the second layer. It looks for the weights of every membership function. This layer gets the receiving signals from the preceding layer and then it acts as membership function to the representation of the fuzzy sets of each input variable, respectively. Second layer nodes are non-adaptive. The layer acts as a multiplier for receiving signals and sends out the outcome in $w_i = \mu_{AB}(x) * \mu_{CD}(y)$ form. Every output node exhibits the firing strength of a rule.

The next layer, the third, is known as the rule layer. All neurons here act as the pre-condition matching the fuzzy rules i.e. each rule's activation level is calculated whereby the number of fuzzy rules is equal to the quantity of layers. Every node computes the normalized weights. The nodes in the 3rd layer are also considered non-adaptive. each of the node computes the value of the rule's firing strength over the sum of all rules' firing strengths in the form of $w_i^* = \frac{w_i}{w_1 + w_2}$, $i = 1, 2$. The outcomes are refered to as the normalized firing strengths.

The 4th layer is responsible for providing the output values as a result of the inference of rules. This layer is also known as the defuzzification layer. Every 4th layer node is an adaptive node having the node function $O_i^4 = w_i^* \cdot f = w_i^* \cdot (p_i x + q_i y + r_i)$. In this layer, the $\{p_i, q_i, r\}$ is the variable set. The variable set is designated as the consequent parameters.

The 5th and final layer is known as the output layer. It adds up all the receiving inputs from the preceding layer. Thereafter, it converts the fuzzy classification outcomes into a binary (crisp). The single node of the 5th layer is considered non-adaptive. This node calculates the total output as the wholesum of all receiving signals:

$$O_i^5 = \sum_i w_i^* xf = \frac{\sum_i w_i f}{\sum_i w_i} \tag{3}$$

In the process of identification of variables in the ANFIS architectures, the hybrid learning algorithms were applied. In the forward pass of the hybrid learning algorithm, functional signals go forward until Layer 4 and the consequent parameters are determined by the least squares estimate. In the backward pass, the error rates circulate backwards and the premise variables are updated by the gradient descent.

3. Results and discussion

A thorough search was performed based upon the considered input parameters to select the set of the optimal combination of inputs that has the most influence on prediction of daily dew point temperature. For this aim, an ANFIS model was built by the functions for each possible combination of the used inputs. Then they were trained respectively for single epoch and then the achieved performance on the basis of each combination was reported. Root mean square error (RMSE) was utilized as a reliable benchmark to show the accuracy level of daily dew point temperature prediction using each input and subsequently determine the rank of each input parameter from the most relevant to the least relevant set. The definition of RMSE is presented in Appendix. The input or set of inputs that provides the lowest errors is considered as the most relevant in regards to the outcome (i.e. daily dew point temperature). The examinations were performed separately for each considered case study to offer further reliability in the evaluations since each case study enjoys different climate conditions. Figure 4 (a) and (b) present the results of ANFIS regression error on the basis of RMSE for both training and testing phases for Kerman and Tabass, respectively, which illustrate the influence of each considered parameter on dew point temperature prediction. It should be mentioned that the training data are 50% of the whole dataset with even sample numbers while the testing data are 50% of the remaining datasets with odd sample numbers. As it is clear from the results offered in Fig. 4 (a) and (b), for both Kerman and Tabass stations V_P with a great difference has the lowest RMSE; thus, it is the most relevant parameter to predict daily dew point temperature. Additionally, it is found that the R_h is the less relevant input parameter since it has the highest RMSE. It is also noticed that for both Kerman and Tabass stations, T_{min} is the second influential parameter. Clearly, there is a significant difference between the RMSE values for the first and even the second most significant parameters. This emphasizes the significance of proper selection of input element for prediction of dew point temperature.

The fact that both the training and checking RMSE are comparable is an indirect indication that suggests that there is no over fitting. This means that the selection of more than one input parameter in the construction of the ANFIS model can be explored. To achieve this, an evaluation on the best combination of 2 input parameters can be conducted. It is worth mentioning that R_h, as the least relevant input, was not eliminated from the possible sets of parameters to analyze the importance of more input combinations. To recognize the best combination of 2 parameters that are the most relevant for prediction of dew point temperature, 28 possible sets of inputs were considered and analyzed.

The results listed in Tables 3 and 4 show the attained RMSE values for all single input parameters and two input combinations utilized

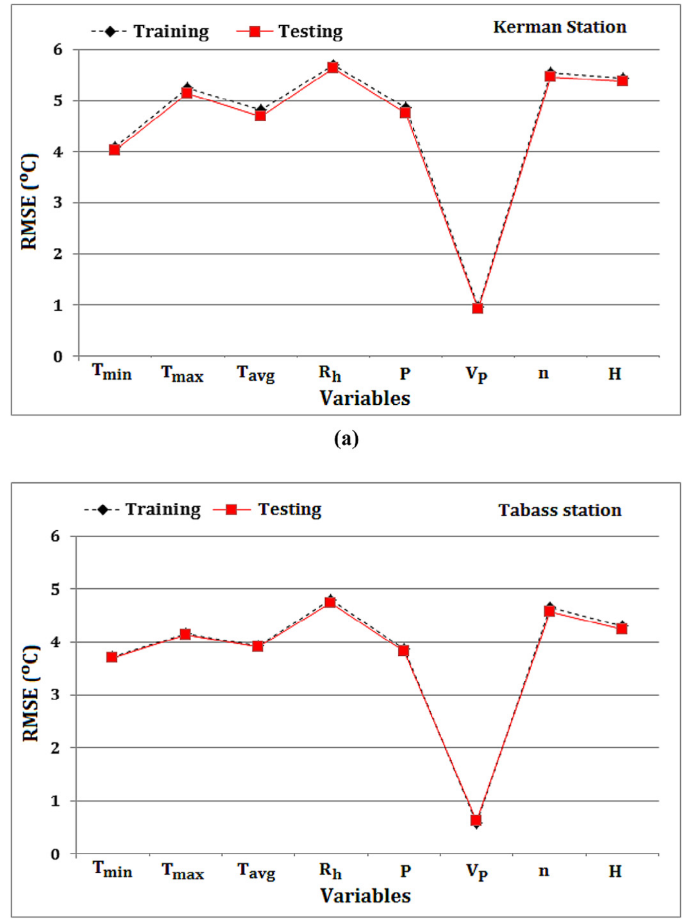

Fig. 4. Achieved RMSE showing the influence of each input parameters on dew point temperature for (a) Kerman and (b) Tabass stations.

for prediction of daily dew point temperature, respectively for Kerman and Tabass. The results indicate that of all the parameters examined, the combination of T_{min} and V_P is the most influential for dew point temperature prediction, and the best predictors of accuracy for both case studies. The RMSE values for the cases of the best single input and the best combination of two inputs are shown in bold in Tables 3 and 4.

For Kerman station, the RMSE values for V_P as the best single input are 0.9662 °C and 0.9321 °C, respectively for the training and checking phases. While for Tabass station, the RMSE values for V_P as the most relevant single input are achieved as 0.5912 °C and 0.6182 °C, respectively for the training and checking phases. Regarding the most influential combinations of 2 inputs, for Kerman station the RMSE for the combination of T_{min} and V_P as the most relevant set of 2 inputs is 0.8068 °C in the training phase and is 0.8335 °C in the checking phase. Also, for Tabass station the RMSE values for the combination of T_{min} and V_P as the most relevant set of 2 inputs are 0.4903 °C and 0.5445 °C, respectively for the training and checking phases.

It is observed that by increasing the number of the most relevant input from 1 to 2, the amount of errors decreases slightly, which provide a small enhancement in the accuracy of prediction. From this viewpoint, it can be concluded that increasing the number of inputs to higher than 2 is not advisable due to minor improvements achieved and further complexity in the required inputs for prediction of daily dew point temperature.

It should be noted that a model with more simplicity in terms of required inputs is always preferable; thus, the use of more than

Table 3
ANFIS regression errors (RMSE in °C) obtained for one and two-input combinations for the training (tr) and checking (ch) phases for Kerman station.

	T_{min}	T_{max}	T_{avg}	R_h	P	V_P	n	H
T_{min}	tr = 4.0907, ch = 4.0197	tr = 3.6486, ch = 3.6580	tr = 3.6474, ch = 3.6570	tr = 1.8885, ch = 1.8910	tr = 3.8845, ch = 3.8633	**tr = 0.8068, ch=0.8335**	tr = 3.6024, ch = 3.5898	tr = 3.8042, ch = 3.7173
T_{max}		tr= 5.2433, ch=5.1498	tr = 3.6566, ch = 3.6813	tr = 1.9645, ch = 1.9108	tr = 4.7173, ch = 4.6635	tr = 0.9114, ch=0.8924	tr = 4.4775, ch = 4.4053	tr=5.0494, ch = 4.9583
T_{avg}			tr = 4.8149, ch = 4.6946	tr = 1.2256, ch = 1.1424	tr = 4.6986, ch = 4.5844	tr = 0.8385, ch=0.8435	tr = 4.0101, ch = 3.9410	tr = 4.4964, ch = 4.3790
R_h				tr=5.7045, ch=5.6269	tr=2.8855, ch=2.8660	tr = 0.9278, ch=0.9017	tr=5.2860, ch=5.2453	tr = 4.5100, ch = 4.5237
P					tr = 4.8567, ch = 4.7547	tr = 0.8406, ch=0.8658	tr = 4.3059, ch = 4.2312	tr = 4.6780, ch = 4.5673
V_P						**tr = 0.9662, ch=0.9321**	tr = 0.9581, ch=0.9260	tr = 0.9098, ch=0.9163
n							tr=5.5419, ch=5.4606	tr = 4.8235, ch = 4.7485
H								tr=5.4479, ch=5.3816

Table 4
ANFIS regression errors (RMSE in °C) obtained for one and two-input combinations for the training (tr) and checking (ch) phases for Tabass station.

	T_{min}	T_{max}	T_{avg}	R_h	P	V_P	n	H
T_{min}	tr = 3.7306, ch = 3.6987	tr = 3.3732, ch = 3.3335	tr = 3.3734, ch = 3.3343	tr = 1.2936, ch = 1.2781	tr = 3.6321, ch = 3.6185	**tr = 0.4903, ch=0.5445**	tr = 3.4483, ch = 3.4153	tr = 3.7114, ch = 3.6727
T_{max}		tr = 4.1615, ch = 4.1474	tr = 3.3739, ch = 3.3398	tr = 1.2244, ch = 1.2604	tr = 3.8349, ch = 3.8032	tr = 0.5133, ch=0.5614	tr = 3.8549, ch = 3.8058	tr = 4.1036, ch = 4.0694
T_{avg}			tr = 3.9380, ch = 3.9095	tr = 0.8116, ch=0.8000	tr = 3.7581, ch = 3.7330	tr = 0.5012, ch=0.5534	tr = 3.6312, ch = 3.5932	tr = 3.9114, ch = 3.8725
R_h				tr = 4.8020, ch = 4.7409	tr=2.7243, ch=2.7171	tr = 0.5749, ch=0.6069	tr = 4.5004, ch = 4.4345	tr = 3.7587, ch = 3.6789
P					tr = 3.8672, ch = 3.8301	tr = 0.5016, ch=0.5535	tr = 3.7137, ch = 3.6991	tr = 3.8455, ch = 3.8068
V_P						**tr = 0.5912, ch=0.6182**	tr = 0.5856, ch=0.6147	tr = 0.5377, ch=0.5772
n							tr = 4.6502, ch = 4.5775	tr = 4.1833, ch = 4.0796
H								tr = 4.3125, ch = 4.2500

two inputs in the construction of the ANFIS model may not be advisable and appropriate. It can be concluded that considering the most relevant 2 inputs would be the more proper possibility in terms of optimum number of inputs to provide a balance between the simplicity and high precision. Thus, the two-input ANFIS model will be the basis for further examination.

Once the inputs are fixed, the 100 epochs that are the quantity of epoch on ANFIS training can then be increased. The error curves for these 100 epochs of training and checking for the most influential combination of two input parameters (T_{min} and V_P) are shown in Fig. 5 (a) and (b), respectively for Kerman and Tabass. The training errors and the checking errors are represented by the dashed curve and solid curve, respectively.

Figure 6 (a) and (b) show the input–output ANFIS surface for the most relevant combination of the two input parameters (T_{min} and V_P). As can be seen, this surface is nonlinear and monotonic and illustrates that how the T_{dew} values vary with the values of T_{min} and V_P.

3.1. A comparative study on evaluating the ANFIS's predictions

After determining the most influential input parameters and the best combinations of 2 input parameters for both stations, it is worthwhile to evaluate the proficiency of the ANFIS model to predict dew point using selected inputs. To achieve this, the predictions' performance of the ANFIS is compared with other well-known soft computing approaches including support vector machine (SVM) [47], artificial neural network (ANN) [48] and genetic programming (GP) [49] considered as the benchmark models. On this account, a comparison between the predicted values of daily dew point temperature by the models and the measured ones was performed based on RMSE. Tables 5 and 6 display the attained RMSE for all developed models using the most significant combination of two inputs for both training and checking phases, respectively for Kerman and Tabass stations. It should be mentioned again that the smaller values of RMSE represent further precision of the predicted dew point temperature and in an ideal case they are zero. Therefore, according to the results, it is apparently found that ANFIS approach provides higher precision to predict daily dew point temperature than the SVM, ANN and GP models. Finally, the ANFIS results are compared with already published models as presented in Table 7. It is clear that ANFIS outperformed other models.

4. Conclusions

In this study, a systematic approach was carried out based upon the ANFIS methodology to identify the most relevant parameters for prediction of dew point temperature (T_{dew}) at two Iranian locations of Kerman and Tabass, located in the central and south central parts of the country. For the first time, the ANFIS network was used to perform a variable search for determining that how eight variables of T_{min}, T_{max}, T_{avg}, R_h, P, V_P, n and H influence T_{dew}. The motivation behind the selection of these eight parameters was on the basis of physical factors influencing dew formation. The simulations of this research work were employed in MATLAB, and the outcomes were checked on the corresponding output blocks.

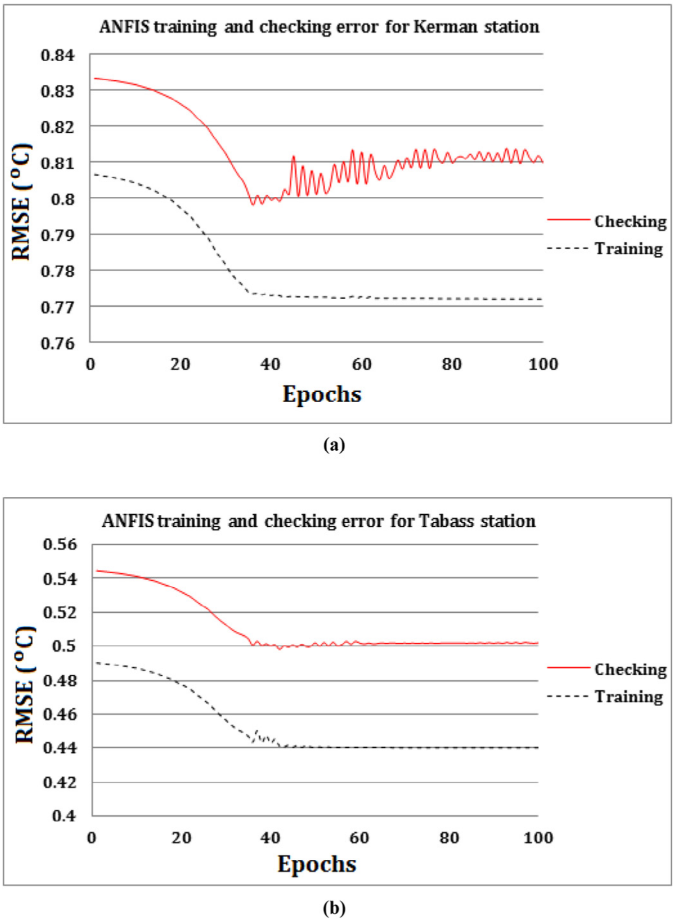

Fig. 5. ANFIS training and checking errors for 2 selected inputs for dew point temperature prediction for (a) Kerman and (b) Tabass stations.

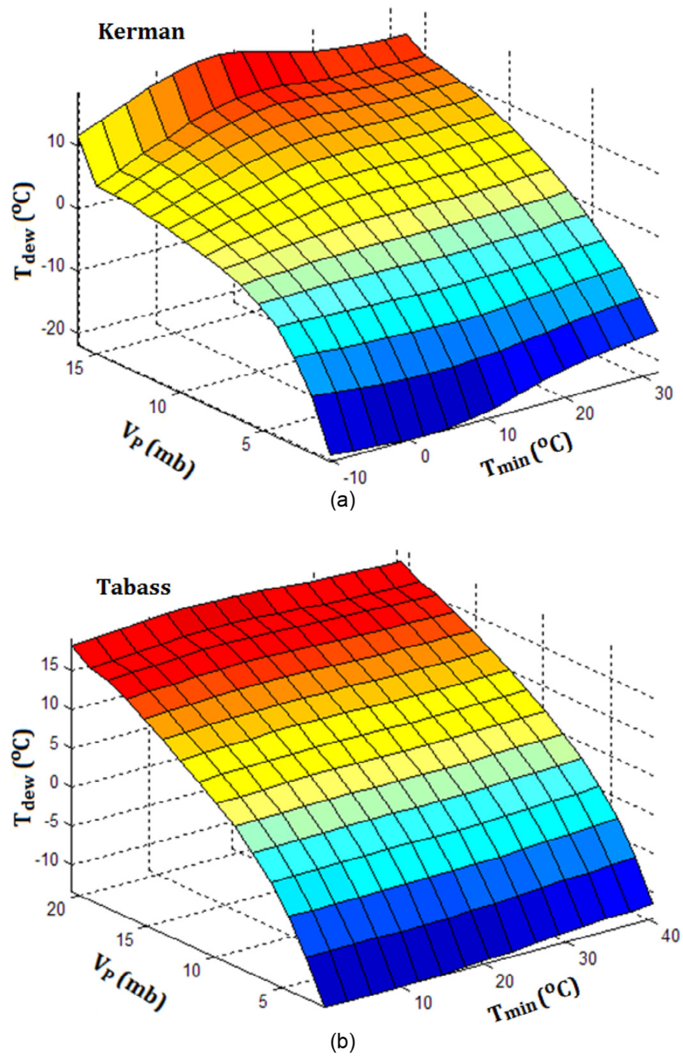

Fig. 6. ANFIS predicted relationship between the most influential parameters for dew point temperature prediction for (a) Kerman and (b) Tabass stations.

The ANFIS process for variable selection was implemented which included a number of ways to recognize the parameters offering favorable prediction. RMSE was used as a reliable benchmark to show the accuracy level of T_{dew} prediction using considered input and subsequently determine the rank of each input parameter from the most relevant to least relevant set. The input or set of inputs which provides the lowest errors was considered most relevant and vice versa.

The achieved results indicated that V_P is the most significant parameter for prediction of T_{dew} at both nominated stations. Also, R_h is the less relevant input parameter. The results clearly showed that there is a remarkable difference between the RMSE values for the most and less relevant input parameters that highlight the significance of appropriate selection of input element to precisely predict T_{dew}. It was also found that combination of T_{min} and V_P is the most influential set to predict T_{dew}. However, the utilization of more than two inputs may not be advisable and appropriate; consequently, considering the most relevant combination of 2 parameters would be more proper to attain higher precision and lower complexity in predictions.

In the final analysis, after determining the most relevant sets of inputs, the predictions of ANFIS were compared with the other soft computing approaches. It was found that using the selected inputs, ANFIS enjoys superiority over other intelligent methods established in this study to predict daily dew point temperature in the studied locations. Moreover, as a further verification, by comparing the performance of the ANFIS model with the selected two inputs with some existing models from the literature it was found that the ANFIS offers more accurate results.

Generally, there are many advantages in the use of the ANFIS scheme such as being adaptable for optimization and adaptive methods as well as being computationally efficient. ANFIS can be integrated with professional systems and rough sets for use in other applications. Systems that handle more complex parameters can also employ the use of ANFIS, as it is much faster compared to other control strategies.

Table 5
The achieved RMSE for ANFIS model with the two selected inputs and other benchmark models for Kerman station.

Model	Training	Checking
	RMSE (°C)	RMSE (°C)
ANFIS	0.8068	0.8335
SVM	1.1123	1.1482
ANN	1.1509	1.1125
GP	1.2360	1.2723

Table 6
The achieved RMSE for ANFIS model with the two selected inputs and other benchmark models for Tabass station.

Model	Training	Checking
	RMSE (°C)	RMSE (°C)
ANFIS	0.4903	0.5445
SVM	0.6117	0.6471
ANN	0.6709	0.6927
GP	0.8760	0.8926

Table 7
Comparison between the ANFIS model of this study using two best inputs and other published models.

References	Model type	Number of inputs	Location	RMSE (°C)
Present study	ANFIS	2	Iran (Kerman)	0.8335
Present study	ANFIS	2	Iran (Tabass)	0.5445
Hubbard et al. [7]	Regression based	2	USA (six stations)	3.22
Zounemat-Kermani [17]	MLR	4	Canada (Geraldton)	0.931
Zounemat-Kermani [17]	ANN	4	Canada (Geraldton)	0.904
Kim et al. [19]	GRNN	2	USA (U.C. Riverside)	1.20
Kim et al. [19]	GRNN	2	USA (Durham)	1.84
Kim et al. [19]	MLP	2	USA (U.C. Riverside)	1.29
Kim et al. [19]	MLP	2	USA (Durham)	1.89

Acknowledgements

The first author, Kasra Mohammadi, would like to offer his deep appreciation to Mr. Kenneth J. Reade (Director of International Student and Scholar Services at the University of Massachusetts Amherst) for all encouraging supports during the time we performed this research work.

This project was supported by the High Impact Research Grant (UM.C/625/1/HIR/MOHE/FCSIT/15) and Fundamental Research Grant Scheme (FRGS) – FP071-2015A from the University of Malaya and the Ministry of Higher Education, Malaysia.

Appendix

The RMSE determines the precision of the model by comparing the deviation between the predicted and real data. The RMSE has always a positive value and is calculated by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_{i,pred} - X_{i,meas})^2}$$

where $X_{i,pred}$ and $X_{i,meas}$ are the ith predicted and measured values, respectively.

References

[1] D.B. Shank, Dew point temperature prediction using artificial neural networks, (MS thesis), Harding University, 2006.
[2] A.J. Atzema, A.F.G. Jacobs, L. Wartena, Moisture distribution within a maize crop due to dew, Neth. J. Agric. Sci. 38 (1990) 117–129.
[3] J. Shiri, S. Kim, O. Kisi, Estimation of daily dew point temperature using genetic programming and neural networks approaches, Hydrol. Res. 45 (2) (2014) 165–181.
[4] R.L. Snyder, J.P.D. Melo-Abreu, Frost Protection: Fundamentals, Practice and Economics, vol. 1, Food and Agricultural Organization of the United Nations, Rome, 2005.
[5] J.S. Kimball, S.W. Running, R. Nemani, An improved method for estimating surface humidity from daily minimum temperature, Agric. Forest Meteorol. 85 (1997) 87–98.
[6] N. Agam, P.R. Berliner, Dew formation and water vapor absorption in semi-arid environments – a review, J. Arid Environ. 65 (2006) 572–590.
[7] K.G. Hubbard, R. Mahmood, C. Carlson, Estimating daily dew point temperature for the northern Great Plains using maximum and minimum temperature, Agron. J. 95 (2) (2003) 323–328.
[8] C. Jeong, J.Y. Shin, T. Kim, J.H. Heo, Monthly precipitation forecasting with a neuro-fuzzy model, Water Resour. Manage. 26 (2012) 4467–4483.
[9] O. Kisi, H. Sanikhani, Prediction of long-term monthly precipitation using several soft computing methods without climatic data, Int. J. Climatol. (2015) doi:10.1002/joc.4273.
[10] O. Hamidi, J. Poorolajal, M. Sadeghifar, H. Abbasi, Z. Maryanaji, H.R. Faridi, et al., A comparative study of support vector machines and artificial neural networks for predicting precipitation in Iran, Theor. Appl. Climatol. 119 (2015) 723–731, doi:10.1007/s00704-014-1141-z.
[11] H. Citakoglu, M. Cobaner, T. Haktanir, O. Kisi, Estimation of monthly mean reference evapotranspiration in Turkey, Water Resour. Manage. 28 (2014) 99–113.
[12] H. Tabari, P.H. Talaee, P. Willems, Short-term forecasting of soil temperature using artificial neural network, Meteorol. Appl. (2014) doi:10.1002/met.1489.
[13] O. Kisi, M. Tombul, M. Zounemat-Kermani, Modeling soil temperatures at different depths by using three different neural computing techniques, Theor. Appl. Climatol. (2014) doi:10.1007/s00704-014-1232-x.
[14] P.H. Talaee, Daily soil temperature modeling using neuro-fuzzy approach, Theor. Appl. Climatol. 118 (2014) 481–489.
[15] M. Cobaner, H. Citakoglu, O. Kisi, T. Haktanir, Estimation of mean monthly air temperatures in Turkey, Comput. Electr. Agric. 109 (2014) 71–79.
[16] D.B. Shank, G. Hoogenboom, R.W. McClendon, Dew point temperature prediction using artificial neural networks, Appl. Meteorol. Climatol. 47 (2008) 1757–1769.
[17] M. Zounemat-Kermani, Hourly predictive Levenberg–Marquardt ANN and multi linear regression models for predicting of dew point temperature, Meteorol. Atmos Phys. 117 (2012) 181–192.
[18] K. Nadig, W. Potter, G. Hoogenboom, R. McClendon, Comparison of individual and combined ANN models for prediction of air and dew point temperature, Appl. Intell. 39 (2) (2013) 354–366.
[19] S. Kim, V.P. Singh, C.J. Lee, Y. Seo, Modeling the physical dynamics of daily dew point temperature using soft computing techniques, KSCE J. Civil Eng. (2014) doi:10.1007/s12205-014-1197-4.
[20] Wikipedia. Kerman. <http://en.wikipedia.org/wiki/Kerman>, 2015 (accessed 30.02.15).
[21] M. Kottek, J. Grieser, C. Beck, B. Rudolf, F. Rubel, World map of the Koppen–Geiger climate classification updated, Meteorol. Z. 15 (3) (2006) 259–263.
[22] K. Mohammadi, S. Shamshirband, S. Motamedi, D. Petković, R. Hashim, M. Gocic, Extreme learning machine based prediction of daily dew point Temperature, Computers and Electronics in Agriculture 117 (2015) 214–225.
[23] T. Cibas, F.F. Soulie, Variable selection with neural networks, Neurocomputing 12 (1996) 223–248.
[24] F.O. Anderson, M. Aberg, S.P. Jacobsson, Algorithmic approaches for studies of variable influence, contribution and selection in neural networks, Chemometr. Intell. Lab. Syst. 51 (2000) 61–72.
[25] G. Castellano, A.M. Fanelli, Variable selection using neural-network models, Neurocomputing 31 (2000) 1–13.
[26] F. Dieterle, S. Busche, G. Gauglitz, Growing neural networks for a multivariate calibration and variable selection of time-resolved measurements, Anal. Chim. Acta 490 (2003) 71–83.
[27] K. Mohammadi, S. Shamshirband, D. Petković, H. Khorasanizadeh, Determining the most important variables for diffuse solar radiation prediction using adaptive neuro-fuzzy methodology; case study: City of Kerman, Iran. Renew. Sustain. Energy Rev. 53 (2016) 1570–1579.
[28] D. Petkovic, M. Gocic, S. Trajkovic, S. Shamshirband, S. Motamedi, R. Hashim, et al., Determination of the most influential weather parameters on reference evapotranspiration by adaptive neuro-fuzzy methodology, Computers and Electronics in Agriculture 114 (2015) 277–284.
[29] J.S.R. Jang, ANFIS: adaptive-network-based fuzzy inference systems, IEEE Trans. Syst. Man Cybern. 23 (1993) 665–685.
[30] A. Al-Ghandoor, M. Samhouri, Electricity consumption in the industrial sector of Jordan: application of multivariate linear regression and adaptive neuro-fuzzy techniques, Jordan J. Mech. Ind. Eng. 3 (2009) 69–76.
[31] D. Petković, M. Protić, S. Shamshirband, S. Akib, M. Raos, D. Marković, Evaluation of the most influential parameters of heat load in district heating systems, Energy Build. 104 (2015) 264–274.
[32] R. Singh, A. Kianthola, T.N. Singh, Estimation of elastic constant of rocks using an ANFIS approach, Appl. Soft Comput. 12 (2012) 40–45.
[33] S. Kariminia, S. Motamedi, S. Shamshirband, D. Petković, C. Roy, R. Hashim, Adaptation of ANFIS model to assess thermal comfort of an urban square in moderate and dry climate, Stochastic Environmental Research and Risk Assessment. (2015) 1–15, doi:10.1007/s00477-015-1116-3.
[34] V. Nikolic, S. Shamshirband, D. Petkovic, K. Mohammadi, Z. Cojbasic, T.A. Altameem, et al., Wind wake influence estimation on energy production of wind farm by adaptive neuro-fuzzy methodology, Energy 80 (2015) 361–372.
[35] R. Sivakumar, K. Balu, ANFIS based distillation column control, IJCA Spec. Issue Evol. Comput. Optim. Tech. (2010) 67–73.
[36] M. Hosoz, H.M. Ertunc, H. Bulgurcu, An adaptive neuro-fuzzy inference system model for predicting the performance of a refrigeration system with a cooling tower, Expert Syst. Appl. 38 (2011) 14148–14155.
[37] L. Tian, C. Collins, Adaptive neuro-fuzzy control of a flexible manipulator, Mechatronics 15 (2005) 1305–1320.

[38] F.G. Areed, A.Y. Haikal, R.H. Mohammed, Adaptive neuro-fuzzy control of an induction motor, Ain Shams Eng. J. 1 (2010) 71–78.
[39] S. Kurnaz, O. Cetin, O. Kaynak, Adaptive neuro-fuzzy inference system based autonomous flight control of unmanned air vehicles, Expert Syst. Appl. 37 (2010) 1229–1234.
[40] S. Ravi, M. Sudha, P.A. Balakrishnan, Design of intelligent self-tuning GA ANFIS temperature controller for plastic extrusion system, Model. Simul. Eng. (2011) 1–8.
[41] A.A. Aldair, W.J. Wang, Design an intelligent controller for full vehicle nonlinear active suspension systems, Int. J. Smart Sens. Intell. Syst. 4 (2011) 224–243.
[42] M.R. Dastranj, E. Ebroahimi, N. Changizi, E. Sameni, D.C. Control, Motorspeed with adaptive neuro-fuzzy control (ANFIS), Aus. J. Basic Appl. Sci. 5 (2011) 1499–1504.
[43] T.L. Grigorie, R.M. Botez, Adaptive neuro-fuzzy inference system-based controllers for smart material actuator modeling, Aerosp. Eng. 223 (2009) 655–668.
[44] R.S.D. Wahida Banu, A. Shakila Banu, D. Manoj, Identification and control of nonlinear systems using soft computing techniques, Int. J. Model. Optim. 1 (2011) 24–28.
[45] H. Moustapha, R. Krishnamurti, Arabic Calligraphy: A Computational Exploration. School of Architecture, Carnegie Mellon University, Pittsburgh, USA, 2001.
[46] M.A. Akcayol, Application of adaptive neuro-fuzzy controller for SRM, Adv. Eng. Softw. 35 (2004) 129–137.
[47] V.N. Vapnik, V. Vapnik, Statistical Learning Theory, vol. 2, Wiley, New York, 1998.
[48] R.J. Schalkoff, Artificial Neural Networks, McGraw-Hill Higher Education, 1997.
[49] J.R. Koza, Genetic Programming: On the Programming of Computers by Means of Natural Selection, MIT press, 1992.