Kütahya Dumlupınar University
Institute of Graduate Studies

Journal of Scientific
Reports-A
E-ISSN: 2687-6167
Number 46, June 2021

RESEARCH ARTICLE

THE EFFECT OF THE DATA TYPE ON ANFIS RESULTS, CASE STUDY TEMPERATURE AND RELATIVE HUMIDITY

Pinar Bakhtiyar Abdulkareem SALIHI[1], Nadire UCLER[2]*

[1]Van Yüzüncü Yıl University, Faculty od Engineering, Department of Civil Engineering, 65090, Van, TUKEY, bnarbaxtyar77@gmail.com, Tel: +964 771 013 7007, ORCID: 0000-0003-0521-9274
[2]Van Yüzüncü Yıl University, Faculty od Engineering, Department of Civil Engineering, 65090, Van, TUKEY, nadireucler@yyu.edu.tr, Tel: +90 530 881 37 11, ORCID: 0000-0001-6407-121X

Recieved Date:05.10.2020 *Accepted Date: 12.04.2021*

ABSTRACT

In this study, the Adaptive Neuro-Fuzzy Inference System (ANFIS) was used to create models to predict mean relative humidity and temperature with the most suitable inputs. To find the most appropriate data type for these meteorological parameters both hourly-daily and raw-normalized data sets were used, and results were compared. The models were trained with 2014-2017 data observed at Kirkuk city station in Iraq, were checked with both 2018 data of Kirkuk and Sanliurfa city station in Turkey to investigate whether a model set with the data of a country could be used for another country data set. The execution of models was evaluated by using root mean square error (RMSE), mean absolute error (MAE), and determination coefficient R^2. Among the two parameters, the temperature achieved the best performance using relative humidity and dew point as input variables. According to the results, daily normalized data had lower error values and higher R^2 than hourly un-normalized data. Additionally, the results showed that the model performed successfully at the Sanliurfa city station on the temperature parameter because of similar climate conditions to Kirkuk city.

Keywords: *ANFIS, Normalization, Relative humidity, Temperature*

1. INTRODUCTION

Many academics have suggested many different methods or models to overcome the forecasting problems [1]. ANFIS is one of the most popular approaches to create prediction models. Generally, the results of this approach have been compared with the results of the other prediction models such as Artificial Neural Network (ANN), Multiple Linear Regression (MLR), and Multiple Nonlinear Regression (MNLR) [2] and ANFIS has gotten acceptable results for the prediction of parameters such as dew point temperature [3], evaporation [4-7], evapotranspiration [8, 9], rainfall [10-12], rainfall-runoff [13], groundwater level [14, 15], wind speed [16, 17], water usage [18], soil variable [19].

Scientists are trying to find the most suitable input combinations and the type of the method to generate fis to set the most accurate models. Additionally, various time scales such as minute [20, 21], hourly [22], daily [23-25], and monthly [26-29] are chosen to forecast different variables [30].

JSR A
Journal of Scientific Reports

Salihi, P. B. A. and Ucler, N., Journal of Scientific Reports-A, Number 46, 14-33, June 2021.

Because of the major differences among values of the data used in the model can negatively affect the performance of the models, researchers generally use normalization procedure on data to improve the training speed and the accuracy of models. D_Min_Max [31], Z-score [32], sigmoid [33], linear, logarithmic, and square root methods [34] were used to investigate the importance of normalization in the forecasting by ANFIS.In this study, the most suitable input version is tried to be found by trying different input combinations for the prediction of temperature and relative humidity, which are the main meteorological parameters. Also, the results were compared using both hourly and daily data to determine which time series is more suitable for these two parameters. Besides, to investigate the effect of the normalization process on the outcome of the ANFIS method, the ANFIS method was applied to both raw data and data normalized with the max-min normalization method which is the most common normalization method.It is not easy to create models for each city, sometimes it is very difficult to obtain specific city data due to policy interference or the economic aspect or the lack of provision of the equipment and tools required to observe the weather condition. So, to investigate whether the model that is created based on a city data can be used for other city with similar climatic conditions, the model set with the data of Kirkuk station was checked with data of Sanliurfa station. May). The city is located at 35.4666° N latitude and 44.3799° E longitude and it is situated at elevation 346 meters above sea level in north-central of Iraq with a 9,679 km^2area of the city.

2. MATERIAL AND METHOD

2.1. Case Study

In this study, 5 years climate data of dew point (Dp), solar radiation (Sr), relative humidity (H), temperature (T), wind speed (Ws) and pressure (P) which their units Centigrade, Watt/m2, %, C°, m/sec, mm, and bar respectively have been used. Hourly data of parameters that were recorded by a meteorological station in Kirkuk city (Figure 1) were transformed into daily data by calculating their averages. Data of (2014-2017) have been used as training data and the data of 2018 used to be checking data.

Kirkuk climate is generally characterized by hot summers and rate rain in the winters, the area receives an average annual rainfall of 361.3 mm which occurs during the monsoon period (October to

JSR A
Journal of Scientific Reports

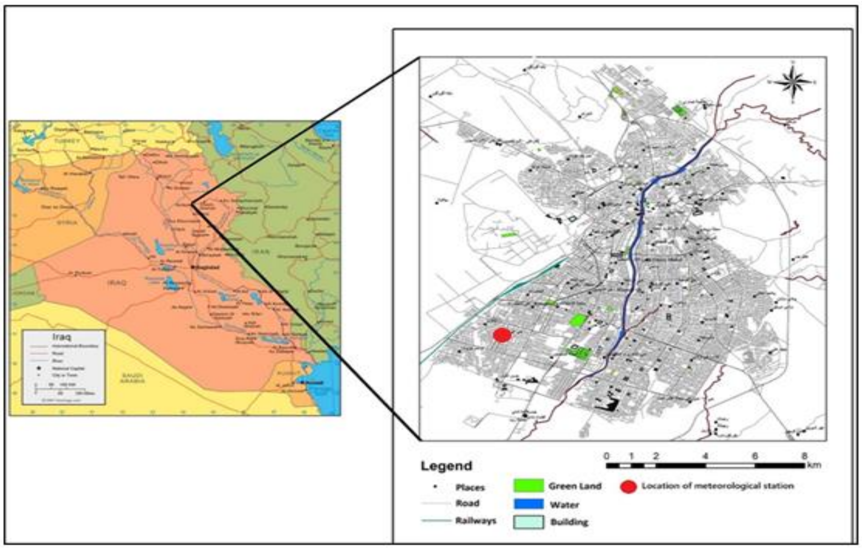

Figure 1. The location of Kirkuk station.

The results have been compared with the data of Sanliurfa city in Turkey, which have similar features with Kirkuk city. Sanliurfa city is at the southeast part of the country, the area of the city is 18,584 km^2. Its coordinates are 37°10′01″ N latitude and 38°47′38″ E longitude and it lies on 543m above sea level the climate here is mild, and generally warm and temperate. The rain in Şanliurfa falls mostly in the winter, with relatively little rain in the summer. The rainfall here is around 477 mm per year. The location of Sanliurfa station in Turkey is shown in Figure 2.

Figure 2. The location of Sanliurfa station.

Some statistical data of parameters are listed in the following tables. The tables show the mean, minimum and maximum values of parameters, standard deviation, skewness, and correlations. The tables consist of the data of Kirkuk and Sanliurfa for the years (2014-2017) and 2018 separately.

The statistical parameters and the features of the hourly dataset of 2014-2017 of Kirkuk city is presented in Table 1. The best value of correlation is between temperature and relative humidity 0.862, which means that they have a stronger linear relationship comparing with the other values of correlation coefficients. Pressure has the lowest skewness value -0.042 and it can be called symmetrical. Wind speed has the highest value of skewness 1.837 which is highly skewed.

Table 1. The hourly statistical parameters of the climatic data of Kirkuk (2014-2017).

Data	**Units**	X_{mean}	X_{min}	X_{max}	**Standard Deviation**	**Skewness**	**Correlation with relative humidity**
Relative Humidity	%	40.75	4	99	24.20	0.605	1
Temperature	C°	24.04	-1.7	49.3	11.24	0.119	-0.862
Pressure	hpa	760.299	745.1	776.1	6.49	-0.042	0.687
Solar radiation	Watt/m^2	178.675	0	917	244.55	1.129	-0.383
Wind speed	m/s	0.676	0	6.7	0.81	1.837	-0.151
							Correlation with temperature
Relative humidity	%	40.75	4	99	24.20	0.605	-0.862
Dew point	Centigrade	6.687	-26.9	20.4	4.26	-1.022	0.309

Table 2 shows the statistical parameters of each variable for the dataset of 2018 for Kirkuk city. Among the data of 2018 of Kirkuk city, the best linear relationship is between temperature and relative humidity -0.836. Additionally, pressure has the less skewness that is 0.172.

Table 2. The hourly statistical parameters of the climatic data of Kirkuk (2018).

Data	**Units**	X_{mean}	X_{min}	X_{max}	**Standard Deviation**	**Skewness**	Correlation with relative humidity
Relative Humidity	%	37.436	8	97	20.59	0.961	1
Temperature	C°	27.059	4.6	49.7	9.84	-0.188	-0.836
Pressure	hpa	757.971	746.3	773	5.86	0.172	0.546
Solar radiation	Watt/m^2	194.488	0	880	259.36	1.06	-0.371
Wind speed	m/s	0.895	0	5.8	0.81	1.391	-0.124
							Correlation with temperature
Relative humidity	%	37.436	8	97	20.59	0.961	-0.836
Dew point	Centigrade	8.901	-9.6	19.2	4.11	-0.746	0.393

Sanliurfa city data gives a weak positive linear relationship by giving 0.105 as a correlation coefficient between relative humidity and dew point, the temperature has the less skewness -0.126, it indicates as symmetrical. The highest value of the standard deviation is 21.799 the result of relative humidity. The highest correlation value is -0.798 and its the relationship between relative humidity and temperature (Table 3).

Table 3. The hourly statistical parameters of the climatic data of Sanliurfa (2018).

Parameters	Units	X_{mean}	X_{min}	X_{max}	Standard Deviation	Skewness	Correlation with	
							H	T
Temperature	C°	22.292	2.3	43	9.432	-0.126	-0.769	1
Relative humidity	%	45.02	9	98	21.799	0.637	1	-0.769
Wind speed	m/s	1.451	0	5.5	0.839	0.903	-0.349	0.38
Pressure	hap	947.61	933.7	964.2	5.578	0.224	0.203	-0.581
Dew point	C°	7.533	-10.2	22.7	5.443	-0.206	0.105	0.496

Among the daily dataset of Kirkuk relative humidity and temperature have the most influential correlation coefficient -0.871 which makes temperature the most effective parameter in that model. A symmetrical skewness is temperature, and it is 0.031 (Table 4).

Table 4. The daily statistical parameters of the climatic data of Kirkuk (2014-2017).

Data	Units	X_{mean}	X_{min}	X_{max}	Standard Deviation	Skewness	Correlation with Relative humidity
Relative Humidity	%	41.253	8.875	97.083	22.54	0.515	1
Temperature	C°	23.824	1.647	43.308	10.58	0.031	-0.871
Pressure	hpa	760.385	746.212	775.045	6.42	-0.056	0.728
Solar radiation	Watt/m^2	171.038	0	382.4	74.77	-0.132	-0.733
Wind speed	m/s	0.636	0	4.708	0.53	2.222	-0.189
							Correlation with temperature
Relative humidity	%	41.253	8.857	97.983	22.54	0.515	-0.871
Dew point	Centigrade	6.666	-21.962	17.808	3.96	-1.186	0.354

Despite the similarities in the characteristics of the two cities, there are some differences between their features. It is discovered that Kirkuk's average daily data can give better results than hourly data because taking average or arithmetic means give us a rough estimate about the common values in that set so that the calculations on all the values will be more or less the same. The average is a good measure of the dataset when a dataset contains values that are relatively evenly spread with no exceptionally high or low values (Table 5).

Table 5. The daily statistical parameters of the climatic data of Kirkuk (2018).

Data	Units	X_{mean}	X_{min}	X_{max}	Standard Deviation	Skewness	Correlation with Relative humidity
Relative Humidity	%	38.266	11.958	92.146	19.58	0.887	1
Temperature	C°	26.726	8.512	43.062	9.23	-0.356	-0.848
Pressure	hpa	758.076	747.237	771.862	5.81	0.136	0.606
Solar radiation	Watt/m^2	187.331	14.458	308.521	77.12	-0.413	-0.687
Wind speed	m/s	0.855	0	2.545	0.49	0.991	-0.183
							Correlation with temperature
Relative humidity	%	38.266	11.958	92.146	19.58	0.887	-0.849
Dew point	Centigrade	8.887	-5.883	17.4	3.84	-0.842	0.427

Sanliurfa daily dataset has almost the same results as hourly data there is not an obvious difference between the two performs of the datasets (Table 6).

Table 6. The daily statistical parameters of the climatic data of Sanliurfa (2018).

Parameters	Units	X_{mean}	X_{min}	X_{max}	Standard Deviation	Skewness	Correlation with	
							H	**T**
Temperature	C°	22.093	5.575	34.861	8.544	-0.305	-0.754	1
Humidity	%	45.671	18.21	96.166	18.967	0.805	1	-0.754
Wind speed	m/s	1.433	0.3	3.5	0.533	0.633	-0.481	0.546
Pressure	hpa	947.87	935.204	962.22	5.607	0.268	0.204	-0.623
Dew point	C°	7.556	19.487	-7.179	5.072	-0.327	0.046	0.594

2.2 Adaptive Neuro Fuzzy Inference System

The adaptive network-based fuzzy inference system (ANFIS) that was proposed by Roger Jang [35] is one of the systems that most widely used among fuzzy inference systems. ANFIS returns a Takagi–Sugeno FIS and has a network of five layers of feed forwards. The first hidden layer is used for fuzzification the input variables and in the second hidden layer T-norm operators are used to define the preceding part of the rule.

The third hidden layer normalizes the strength of the rule and the fourth hidden layer follows where the relevant rule parameters are calculated. The output layer measures all the inputs as the number of all incoming signals. ANFIS uses the learning of back propagation to test hypothesis parameters (to know about membership function parameters) and least square estimation to evaluate the correct parameters. There are two sections of a method of learning: In the first step, the patterns of input are propagated, and optimal consequent parameters are determined by the least mean square method,

JSR A
Journal of Scientific Reports

although the parameters of the assumptions are believed to be defined for the current period through the training set. Within the second section, the patterns are transmitted. During this epoch, back propagation is used to change the parameters of the assumptions, although the related parameters remain fixed [14].

This system is a fuzzy Sugeno network configuration. Usually, this form of model is built and put within the context of a neural network model to facilitate adaptation. Figure 3 displays an ANFIS framework of two inputs, one output and two input rules. This system has two inputs x and y and one output, where its rule is shown in equations 1 and 2:

If x is A_1 and y is B_1 then $f = p_1x + q_1y + r_1$ (1)

If x is A_2 and y is B_2 then $f = p_2x + q_2y + r_2$ (2)

A_i and B_i are fuzzy sets, fi is the output of the fuzzy area defined by the fuzzy rule. p_i, $q_i$and r_i layout criteria that are determined during the training phase. Figure 3 indicates that each node in this layer is a fuzzy set and the output of any node in this layer belongs to the membership level of the input variable in this fuzzy set. Within this layer, the form parameters decide the structure of the fuzzy set membership function [2].

ANFIS has five layers and each layer contains several node function and node features. Nodes are classified into two groups: adaptive nodes and fixed nodes. The layers are defined as follows:

Layer 1: The nodes that exist in this layer are adaptive nodes.

$$Q_{i.1=}\mu A_i(x) \quad (3)$$
$$Q_{i.1=}\mu B_i(x) \quad (4)$$
Layer 2: The nodes are set and represented by a circle and identified by $\prod$. While $\mathrm{W_i}$ denotes the force of the rules. The output is determined on the basis on equation 5:

$$Q_{2.i} = w_i = \mu A_i(y)\pi B_i(y) \quad with\ i = 1.2 \quad (5)$$

Layer 3: All nodes are set and represented by a circle that called N. The name of the output of this layer is normalized firing strength. The output is determined by the i-th firing strength of the rule by summing up all of them.

$$Q_{3.i} = \bar{w}_i = \frac{w_i}{w_1+w_2} \quad with\ \ i = 1.2 \quad (6)$$

Layer 4: The nodes are adaptive nodes which are shown as follows: Consecutive criteria are p_i, q_i, r_i.

$$Q_{4.i} = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \quad (7)$$

Layer 5: The last layer in which the node is single and marked by $\sum$ and seen by the circle [36].

$$Q_{5.i} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum w_i} \quad (8)$$

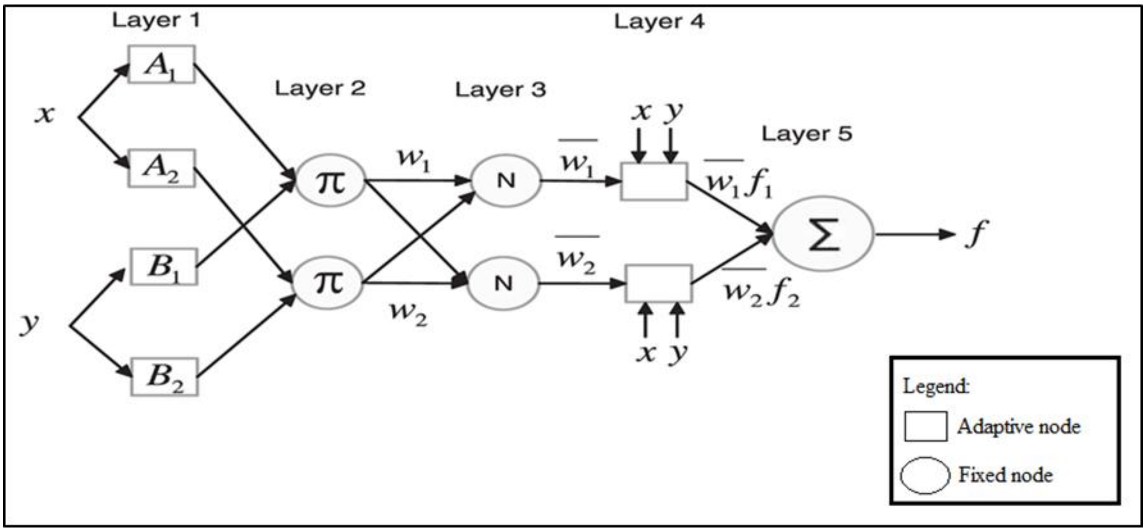

Figure 3. ANFIS structure with two inputs, one output and two rules.

3. APPLICATION AND RESULTS

3.1. Normalization and Evaluation

Equation (9) was used to normalize data before the procedure was implemented. Normalization is a database design technique that organizes data in a manner that reduces redundancy and dependency of data. Normalization divides larger data into smaller data and links them using relationships. The purpose of normalization is to eliminate redundant data and ensure data is stored logically and comparing the result of the un-normalized data to discover how much does the normalization effect on reducing the error value of the program. Here, X_{norm}, X_{min} and X_{max} signify the normalized, minimum, and maximum values of the data set, respectively.

$$x_{norm} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{9}$$

Various variations of meteorological variables were used as inputs of the ANFIS model to analyze the magnitude of each parameter's impact on the prediction of relative humidity and temperature. As assessment criteria, Root mean square errors (RMSE) (Equation 10), mean absolute relative error (MAE) (Equation 11) and determination coefficient (R^2) (Equation 12) statistics were chosen.

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(f_{i(observed)}-f_{i(predicted)}\right)^2} \tag{10}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|f_{i\,(observed)} - f_{i\,(predicted)}\right| \tag{11}$$

JSR A
Journal of Scientific Reports

$$R^2 = \frac{(f_i observed - f_i predicted)^2}{(f_{i\,observed} - average\ of\ dataset)^2} \quad (12)$$

Where; n, $f_{iobserved}$, $\overline{fi}_{observed}$; $f_{ipredicted}$ and $\overline{fi}_{predicted}$ predicted symbolizes number of data, observed, average of observed, predicted and average of predicted values, respectively. To achive the minimum error and max R^2 values between observed and predicted values ideal model parameters were chosen. Sugeno type was used to set the ANFIS model. The input membership functions were 'gaussmf' and the output membership functions were 'linear', defuzzification method was 'wtaver', generate fis type was 'subtractive clustering' method and optimization method was 'Hybrid Optimization Method'.

ANFIS model has been built many times with different inputs each time to find the best effect of each parameter.

The first model is forecasting relative humidity depending on 15 various inputs that include:
(T), (P), (Sr), (Ws),
(T and P), (T and Sr), (T and Ws), (P and Sr), (P and Ws), (Sr and Ws),
(T, P and Sr), (T, P and Ws), (T, Sr, Ws), (P, Sr and Ws) and
(T, P, Sr and Ws) respectively.

The second model predicts temperature based on 2 different input:
(H), (Dp),
(H and Dp) respectively.

Figure 4 illustrates a gaussian membership function and the rules of a model which has two inputs.

JSR A
Journal of Scientific Reports

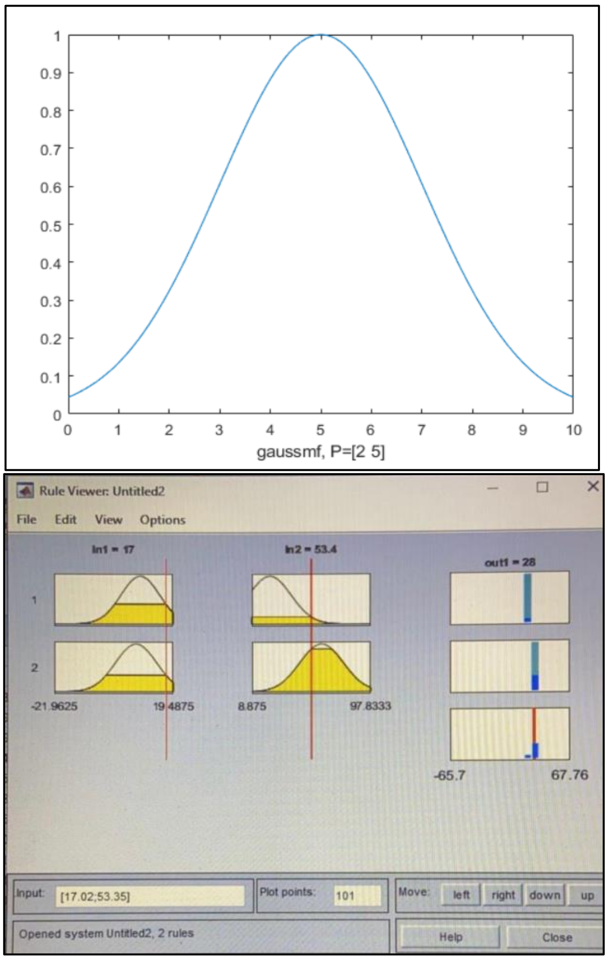

Figure 4. Agaussian membership function and rule view of a model with two inputs.

ANFIS has been applied for each of these models on daily and hourly data plus normalized and un-normalized data. The relative humidity and temperature parameters are defined to be the outputs of the models, these model's inputs have chosen according to the correlation between parameters. All data converted to normalized data to reduce the dependency and make them more accurate, so the second comparison is between normalized and un-normalized data, using the same two models. Average daily and hourly data of (2014-2017) used for training, and daily and hourly data of 2018 of Kirkuk and Sanliurfa were for checking. These comparisons were done by calculating RMSE, MAE and R^2.

3.2. The Results of Relative Humidity Model

Table 7 represents the result of daily data of the relative humidity model. Although results were so close to each other, the best result was obtained using un-normalized T, P, and Ws parameters as inputs. As well as the models that depended on one and two inputs, (T) and (T, P) had the best results with normalized data.

JSR A
Journal of Scientific Reports

Salihi, P. B. A. and Ucler, N., Journal of Scientific Reports-A, Number 46, 14-33, June 2021.

Table 7. RMSE, MAE and R^2statistics of daily relative humidity model.

Output: Relative humidity	**Un-normalized**			**Normalized**		
Inputs	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2
T	10.804	7.631	0.695	10.525	7.292	0.711
P	15.936	11.533	0.337	15.567	11.164	0.367
Sr	14.752	11.106	0.432	14.422	11.004	0.457
Ws	18.535	15.711	0.103	18.865	16.045	0.071
T, P	9.852	6.931	0.746	9.631	6.573	0.757
T, Sr	10.991	7.758	0.684	10.531	7.240	0.711
T, Ws	10.984	7.611	0.685	10.526	7.178	0.711
P, Sr	14.520	10.977	0.229	13.934	10.187	0.493
P, Ws	16.171	11.692	0.317	15.487	11.278	0.374
Sr, Ws	14.441	10.770	0.455	14.163	10.723	0.476
T, P and Sr	9.966	7.199	0.741	9.789	6.725	0.749
T, P and Ws	9.455	6.551	0.766	9.479	6.430	0.765
T, Sr and Ws	10.528	7.539	0.711	10.884	7.528	0.691
P, Sr and Ws	14.509	10.717	0.451	15.875	11.329	0.342
T, P, Sr and Ws	10.023	7.031	0.737	9.745	6.737	0.752

Figure 5 shows the scattering of the ANFIS model of daily un-normalized relative humidity model for three inputs which had the highest R^2 value. In addition, the comparison of the observed and predicted data is presented in Figure 6.

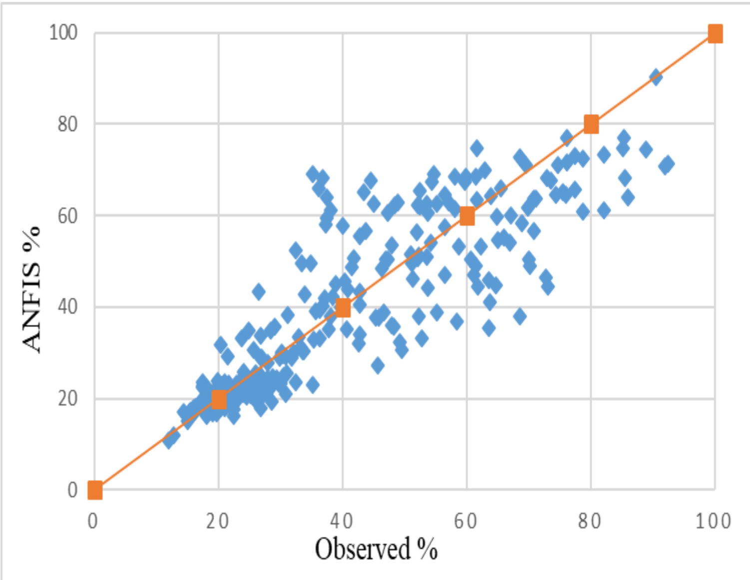

Figure 5. The scattering of ANFIS of daily un-normalized relative humidity model for three inputs.

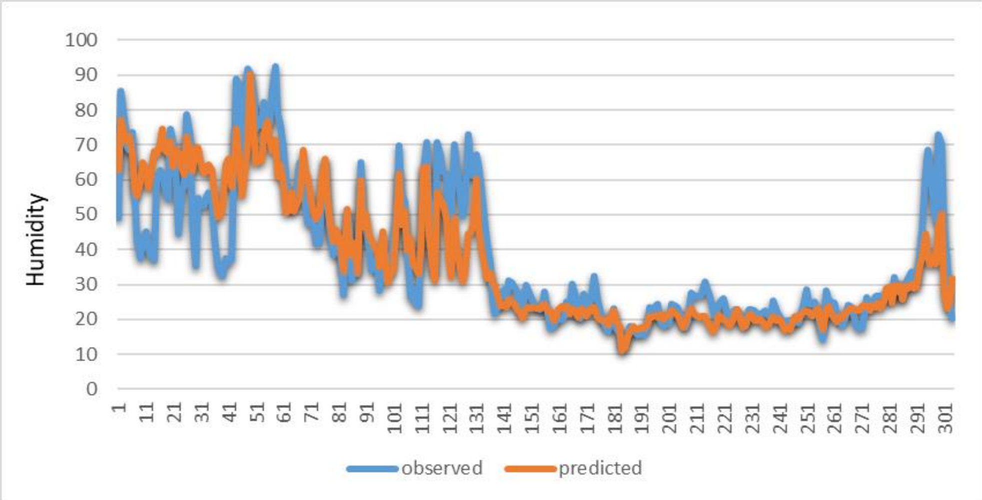

Figure 6. The observed and predicted data of ANFIS method for daily un-normalized humidity model.

Table 8 shows that hourly data got the best results depending on normalized T, P, Sr, and Ws as inputs. RMSE, MAE, and R^2 were 10.185, 7.025 and 0.756, respectively. The daily data of the relative humidity model had better results than hourly data, this means that the central tendency of a dataset may have better performance than a single data. (T) and (T, P) got the best results as one and two input models using normalized data.

Table 8. RMSE, MAE and R^2statistics of hourly relative humidity model.

Output: Relative humidity	Un-normalized			Normalized		
Inputs	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2
T	11.664	8.289	0.679	11.237	7.771	0.702
P	17.437	12.885	0.282	17.084	12.760	0.311
Sr	19.421	16.184	0.110	19.665	16.665	0.087
Ws	20.587	17.392	0.0001	20.844	17.916	-0.025
T, P	10.691	7.565	0.730	10.326	7.175	0.748
T, Sr	11.565	8.214	0.684	11.123	7.710	0.708
T, Ws	11.569	8.248	0.684	11.164	7.758	0.705
P, Sr	16.379	12.078	0.367	16.012	11.837	0.395
P, Ws	17.314	12.761	0.292	17.154	11.558	0.291
Sr, Ws	19.291	16.019	0.122	19.471	16.409	0.105
T, P and Sr	10.763	7.612	0.726	10.328	7.132	0.748
T, P and Ws	10.597	7.486	0.735	10.277	7.139	0.751
T, Sr and Ws	11.456	8.144	0.690	11.111	7.726	0.708
P, Sr and Ws	16.324	12.073	0.371	15.951	11.836	0.399

T, P, Sr and Ws	10.563	7.512	0.736	10.185	7.025	0.755

Figure 7 shows the scattering chart of the ANFIS model of hourly normalized relative humidity model for four inputs which had the highest R^2 value.

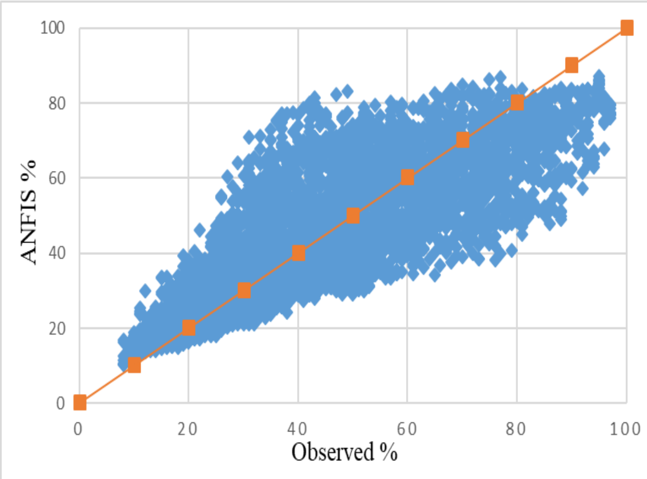

Figure 7. The scattering of ANFIS of hourly normalized relative humidity model for four inputs.

3.3. Results of Temperature Model

During this study and using different inputs and outputs, it was found out that temperature could have the best performance depending on H and Dp as input parameters. Un-normalized model got the lowest error values with RMSE= 0.629 and MAE= 0.727 also highest R^2 value. As one input parameter, relative humidity had better performance using normalized data (Table 9).

Table 9. RMSE, MAE and R^2statistics of daily dataset of temperature model.

Output: Temperature	**Un-normalized**			**Normalized**		
Inputs	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2
H	4.596	3.757	0.752	4.369	3.195	0.775
Dp	8.188	7.159	0.219	8.224	7.069	0.205
H and Dp	0.629	0.727	0.993	3.184	2.803	0.881

Figure 8 and Figure 9 represents the scattering and distribution of the ANFIS model of daily un-normalized temperature model for two inputs which had the highest R^2 value, respectively.

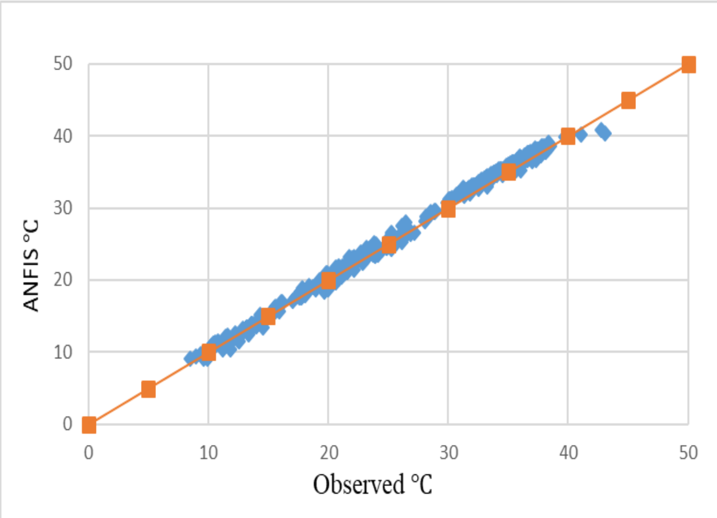

Figure 8. The scattering of ANFIS of daily un-normalized temperature model for two inputs.

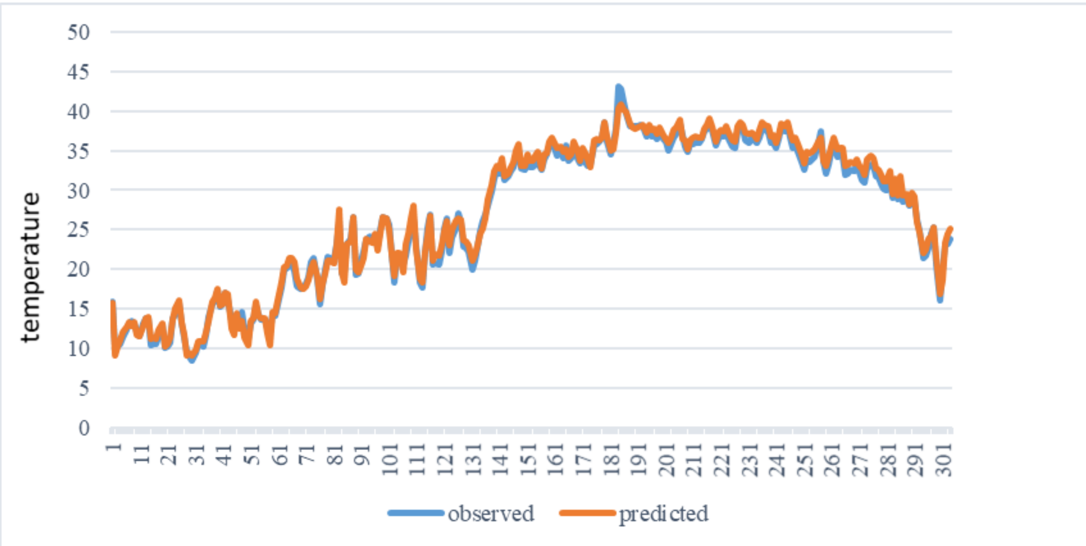

Figure 9. The observed and predicted data of ANFIS method for daily un-normalized temperature model.

Table 10 represents that there is not a big difference between R^2 of daily and hourly datasets but according to the other statistics it was clear that daily data had better performance.

JSR A
Journal of Scientific Reports

Salihi, P. B. A. and Ucler, N., Journal of Scientific Reports-A, Number 46, 14-33, June 2021.

Table 10. RMSE, MAE and R^2 statistics of hourly temperature model.

Output: Temperature	**Un-normalized**			**Normalized**		
Inputs	**RMSE**	**MAE**	R^2	**RMSE**	**MAE**	R^2
H	5.086	4.224	0.732	17.087	14.831	0.708
Dp	9.022	7.616	0.159	11.463	9.631	-0.356
H and Dp	0.927	0.788	0.991	18.699	16.067	0.843

Figure 10 represents the scattering chart of the ANFIS model of hourly un-normalized temperature model for two inputs which had the highest R^2 value.

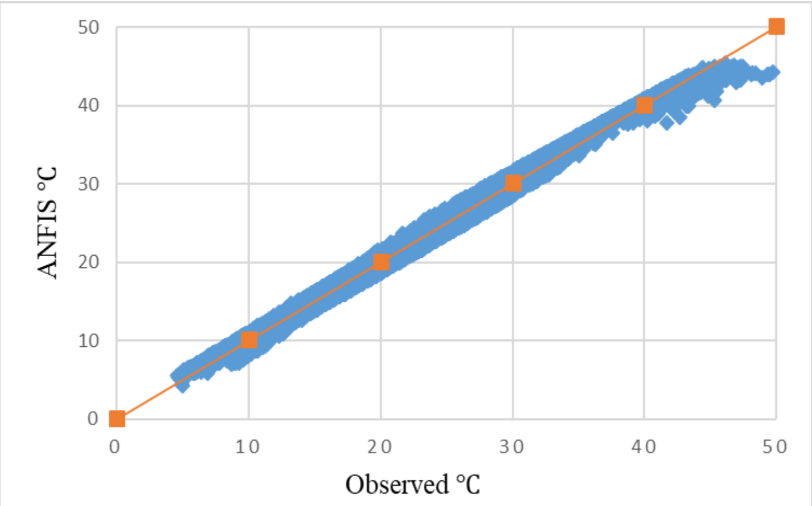

Figure 10. The scattering of ANFIS of hourly un-normalized temperature model for two inputs.

3.4. The Results of Sanliurfa Data Set

Best models of the relative humidity and temperature parameters were chosen to check with data that were observed in Sanliurfa station. Table 11 shows the results of this checking.

The relative humidity model had 14.453 as RMSE, 11.076 as MAE and 0.432 as R^2 depending on T, P and Ws as inputs for daily normalized data. But R^2 value is not sufficient. The temperature model had the best performance for a daily normalized dataset using both inputs H and Dp with R^2=0.876. Also, the error values were quite reasonable for this model.

Table 11. Results of RMSE, MAE and R^2 statistics of Sanliurfa daily dataset.

Type of data	**Inputs**	**Output**	**RMSE**	**MAE**	$\mathbf{R^2}$
Un-normalized	T, P and Ws	H	17.244	13.926	0.173
Normalized	T, P and Ws	H	14.290	11.076	0.432
Un-normalized	H and Dp	T	5.548	5.044	0.578
Normalized	H and Dp	T	3.005	2.389	0.876

JSR A
Journal of Scientific Reports

Salihi, P. B. A. and Ucler, N., Journal of Scientific Reports-A, Number 46, 14-33, June 2021.

Figure 11 and 12 shows the result of temperature model using ANFIS analysis for prediction, depending on the daily normalized dataset.

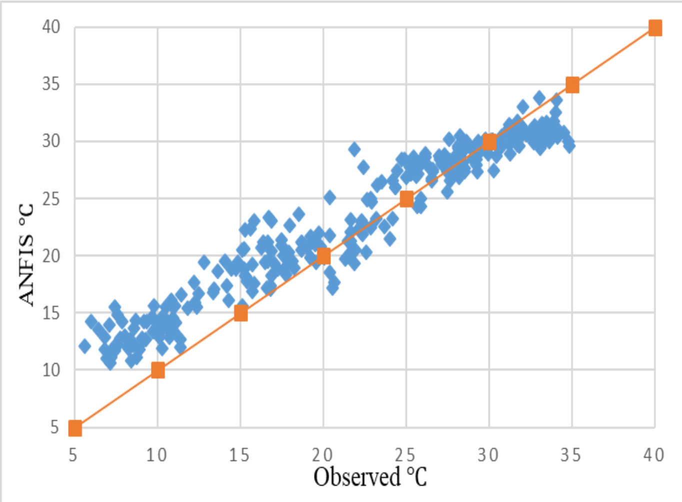

Figure 11. The scattering of ANFIS of daily normalized temperature model for two inputs at Sanliurfa station.

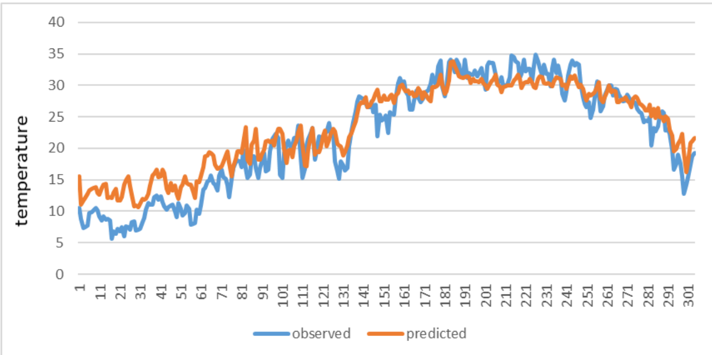

Figure 12. The observed and predicted data of ANFIS method for temperature model of Sanliurfa station.

4. CONCLUSION

In this study, the max-min normalization procedure was applied to the dataset to investigate the effect of the normalization on the results of the ANFIS method. Additionally, the results of the models which were set with hourly and daily data were compared to see the impact of the data type on the results of the models. The results of temperature and relative humidity were compared between Sanliurfa that have similar weather conditions with Kirkuk.

JSR A
Journal of Scientific Reports

It was obtained that the calculation of the average daily parameters from the original dataset could affect positively the results and get much better outcomes than hourly data. Although normalized data got better results than the un-normalized data for almost all models, there were no big differences between the results. This situation indicates that the normalization process cannot help to increase the accuracy of the model without choosing the most appropriate input parameters. A better correlation between the input and output is the reason for the temperature had a perfect performance and got the highest value of R^2 and the most acceptable error values of RMSE and MAE among the two parameters. According to the Sanliurfa results, the temperature model got acceptable results due to the similar weather conditions with Kirkuks's climate. Therefore, it can be said that if it is necessary for some reason, the model obtained using a different data set may be used in another country with a similar climate.

ACKNOWLEDGEMENTS

This article has been prepared using the results obtained in the Master's Thesis supervised by Nadire UCLER and written by Pinar Bakhtiyar Abdulkareem SALIHI.

REFERENCES

[1] Chen, S.-M. and Hwang, J.-R., (2000), Temperature prediction using fuzzy time series. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 30(2), p. 263-275.

[2] Rezaeianzadeh, M., Tabari, H. T., Yazdi, A.A., Isik, S. I. and Kalin, L., (2014), Flood flow forecasting using ANN, ANFIS and regression models. Neural Computing and Applications, 25(1), p. 25-37.

[3] Baghban, A., Bahadori, M., Rozyn, J., Lee, M., Abbas, A., Bahadori, A. and Rahimali, A. , (2016), Estimation of air dew point temperature using computational intelligence schemes. Applied thermal engineering, 93, p. 1043-1052.

[4] Moghaddamnia, A., Gousheh, M.G., Piri, J., Amin, S., and Han, D., (2009), Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. Advances in Water Resources, 32(1), p. 88-97.

[5] Shiri, J., Dierickx, W., Pour-Ali Baba, A., Neamati, S., and Ghorbani, M., (2011), Estimating daily pan evaporation from climatic data of the State of Illinois, USA using adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). Hydrology Research, 42(6), p. 491-502.

[6] Eslamian, S. and Amiri, M.J., (2011), Estimation of daily pan evaporation using adaptive neural-based fuzzy inference system. International Journal of Hydrology Science and Technology, 1(3-4), p. 164-175.

[7] Goyal, M.K., Bharti, B., Quilty, J., Adamowski, J., and Pandey, A., (2014), Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. Expert systems with applications, 41(11), p. 5267-5276.

[8] Kisi, O.,(2013), Applicability of Mamdani and Sugeno fuzzy genetic approaches for modeling reference evapotranspiration. Journal of hydrology, 504, p. 160-170.

[9] Pour-Ali Baba, A., Shiri, J., Kisi, O., Fard, A.F., Kim, S., and Amini, R.,(2013), Estimating daily reference evapotranspiration using available and estimated climatic data by adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). Hydrology Research, 44(1), p. 131-146.

[10] Agboola, A., Gabriel, A., Aliyu, E., and Alese, B.,(2013), Development of a fuzzy logic based rainfall prediction model. International Journal of Engineering and Technology, 3(4), p. 427-435.

[11] Sharifi, S.S., Delirhasannia, R., Nourani, V., Sadraddini, A.A. and Ghorbani, A., (2013), Using artificial neural networks (ANNs) and adaptive neuro-fuzzy inference system (ANFIS) for modeling and sensitivity analysis of effective rainfall. Recent Advances in Continuum Mechanics, Hydrology and Ecology, Mladenov V (eds), p. 133-139.

[12] Hashim, R., Roy, C., Motamedi, S., Shamshirband, S., Petković, D., Gocic, M. and Lee, S.C. ,(2016), Selection of meteorological parameters affecting rainfall estimation using neuro-fuzzy computing methodology. Atmospheric Research, 171, p. 21-30.

[13] Remesan, R., Shamim, M.A., Han, D., and Mathew. J.,(2008), *ANFIS and NNARX based rainfall-runoff modeling*. in *2008 IEEE International Conference on Systems, Man and Cybernetics*. IEEE.

[14] Sreekanth, P., Sreedevi, P., Ahmed, S. and Geethanjali, N., (2011), Comparison of FFNN and ANFIS models for estimating groundwater level. Environmental Earth Sciences, 62(6), p. 1301-1310.

[15] Alipour, Z., Ali, A.M.A., Radmanesh, F. and Joorabyan, M., (2014), Comparison of three methods of ANN, ANFIS and time series models to predict ground water level:(case study: North Mahyar plain). Bulletin of Environment, Pharmacology and Life Sciences, 3, p. 128-134.

[16] Pérez, E.C., Algredo-Badillo, I. and Rodríguez, V.H.G., (2012), Performance analysis of ANFIS in short term wind speed prediction. arXiv preprint arXiv:1212.2671.

[17] Maroufpoor, S., Sanikhani, H., Kisi, O., Deo, R.C. and Yaseen, Z.M., (2019), Long-term modelling of wind speeds using six different heuristic artificial intelligence approaches. International Journal of Climatology, 39(8), p. 3543-3557.

[18] Altunkaynak, A., Özger, M. and Çakmakci, M., (2005), Water consumption prediction of Istanbul city by using fuzzy logic approach. Water Resources Management, 19(5), p. 641-654.

[19] Kashi, H., Emamgholizadeh, S. and Ghorbani, H., (2014), Estimation of soil infiltration and cation exchange capacity based on multiple regression, ANN (RBF, MLP), and ANFIS models. Communications in soil science and plant analysis, 45(9), p. 1195-1213.

[20] Kuo, C.-C., Gan, T.Y. and Yu, P.-S., (2010), Seasonal streamflow prediction by a combined climate-hydrologic system for river basins of Taiwan. Journal of hydrology, 387(3-4), p. 292-303.

[21] Talei, A., Chua, L.H.C., Quek, C. and Jansson, P.-E.,(2013), Runoff forecasting using a Takagi–Sugeno neuro-fuzzy model with online learning. Journal of Hydrology, 488, p. 17-32.

[22] Bazartseren, B., Hildebrandt, G. and Holz, K.-P., (2003), Short-term water level prediction using neural networks and neuro-fuzzy approach. Neurocomputing, 55(3-4), p. 439-450.

[23] Kişi, Ö.,(2013), Evolutionary neural networks for monthly pan evaporation modeling. Journal of hydrology, 498, p. 36-45.

[24] Nourani, V., Kisi, Ö. and Komasi, M., (2011), Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. Journal of Hydrology, 402(1-2), p. 41-59.

[25] Valizadeh, N. and El-Shafie, A., (2013), Forecasting the level of reservoirs using multiple input fuzzification in ANFIS. Water resources management, 27(9), p. 3319-3331.

[26] Dorum, A., Yarar, A., Sevimli, M.F. and Onüçyildiz, M., (2010), Modelling the rainfall–runoff data of susurluk basin. Expert Systems with Applications, 37(9), p. 6587-6593.

[27] Hayashi, S., Murakami, S., Xu, K.-Q., and Watanabe, M., (2008), Effect of the Three Gorges Dam Project on flood control in the Dongting Lake area, China, in a 1998-type flood. Journal of Hydro-environment Research, 2(3), p. 148-163.

[28] El-Shafie, A., Taha, M.R., and Noureldin, A.,(2007), A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam. Water resources management, 21(3), p. 533-556.

[29] Wang, W.-C., Chau, K.-W., Cheng, C.-T., and Qiu, L.,(2009), A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. Journal of hydrology, 374(3-4), p. 294-306.

[30] Fahimi, F., Yaseen, Z.M. and El-shafie, A.,(2017), Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review. Theoretical and applied climatology, 128(3-4), p. 875-903.

[31] Viswavandya, M., Sarangi, B., Mohanty, S. and Mohanty, A.,(2020), *Short Term Solar Energy Forecasting by Using Fuzzy Logic and ANFIS*, in *Computational Intelligence in Data Mining*, Springer. p. 751-765.

[32] Heddam, S.,(2014), Modeling hourly dissolved oxygen concentration (DO) using two different adaptive neuro-fuzzy inference systems (ANFIS): a comparative study. Environmental Monitoring and Assessment, 186(1), p. 597-619.

[33] Azadeh, A., Saberi, M. and Ghorbani, S., (2010), An ANFIS algorithm for improved forecasting of oil consumption: a case study of USA, Russia, India and Brazil.

[34] Ghassemzadeh, S., Shafflie, M., Sarrafi, A. and Ranjbar, M.,(2013), The importance of normalization in predicting dew point pressure by ANFIS. Petroleum science and technology, 31(10), p. 1040-1047.

[35] Jang, J.-S.,(1993), ANFIS: adaptive-network-based fuzzy inference system. IEEE transactions on systems, man, and cybernetics, 23(3), p. 665-685.

[36] Varzandeh, M.H.M., Rahbari, O., Vafaeipour, M., Raahemifar, K. and Heidarzade. F., (2014), *Performance of wavelet neural network and ANFIS algorithms for short-term prediction of solar radiation and wind velocities*. in *The 4th World Sustainability Forum*.