Mathematical Biosciences and Engineering

MBE, 20 (6): 11328–11352.
DOI: 10.3934/mbe.2023502
Received: 21 December 2022
Revised: 07 March 2023
Accepted: 23 March 2023
Published: 27 April 2023

http://www.aimspress.com/journal/MBE

Research article

A comparative study on daily evapotranspiration estimation by using various artificial intelligence techniques and traditional regression calculations

Hasan Güzel[1,*], Fatih Üneş[1], Merve Erginer[1], Yunus Ziya Kaya[2], Bestami Taşar[1], İbrahim Erginer[1] and Mustafa Demirci[1]

[1] Department of Civil Engineering, Iskenderun Technical University, Turkey
[2] Department of Civil Engineering, Osmaniye Korkut Ata University, Turkey

*** Correspondence:** Email: hasan.guzel@iste.edu.tr.

Abstract: Evapotranspiration is an important parameter to be considered in hydrology. In the design of water structures, accurate estimation of the amount of evapotranspiration allows for safer designs. Thus, maximum efficiency can be obtained from the structure. In order to accurately estimate evapotranspiration, the parameters affecting evapotranspiration should be well known. There are many factors that affect evapotranspiration. Some of these can be listed as temperature, humidity in the atmosphere, wind speed, pressure and water depth. In this study, models were created for the estimation of the daily evapotranspiration amount by using the simple membership functions and fuzzy rules generation technique (fuzzy-SMRGT), multivariate regression (MR), artificial neural networks (ANNs), adaptive neuro-fuzzy inference system (ANFIS) and support vector regression (SMOReg) methods. Model results were compared with each other and traditional regression calculations. The ET amount was calculated empirically using the Penman-Monteith (PM) method which was taken as a reference equation. In the created models, daily air temperature (T), wind speed (WS), solar radiation (SR), relative humidity (H) and evapotranspiration (ET) data were obtained from the station near Lake Lewisville (Texas, USA). The coefficient of determination (R^2), root mean square error (RMSE) and average percentage error (APE) were used to compare the model results. According to the performance criteria, the best model was obtained by Q-MR (quadratic-MR), ANFIS and ANN methods. The R^2, RMSE, APE values of the best models were 0,991, 0,213, 18,881% for Q-MR; 0,996; 0,103; 4,340% for ANFIS and 0,998; 0,075; 3,361% for ANN, respectively. The Q-MR, ANFIS and ANN models had slightly better performance than the MLR, P-MR and SMOReg models.

Keywords: Evapotranspiration; Fuzzy-SMRGT; ANN; MR; ANFIS; SMOReg

1. Introduction

Water is a vital natural resource for the existence of all species and the natural balance of the universe. It is critical to be conscious of how much water is used in order to avoid depleting this natural resource. Drought is becoming more frequent as a result of factors such as global warming and climate change, and water resources are depleting. In order to manage water resources, it is critical to reduce the effects of such adverse situations and use existing water resources as efficiently as possible [1]. Evapotranspiration (ET), defined as the loss of water by evaporation from the soil surface and by transpiration from the leaves, is one of the basic components of the hydrological cycle, and it is a very important parameter for the accurate estimation of water resources. ET, like other hydrological events, depends on many factors, and it is very difficult to accurately determine all of these factors. As too many parameters affect hydrological events and the relationship between these events and these parameters is nonlinear, it is very difficult to create real-size models of hydrological events with classical methods. For this reason, researchers are trying to develop easy and practical methods for solving nonlinear problems such as ET [2–4].

Artificial intelligence methods, which can be modeled more easily than classical methods and are more suitable for real-world situations, are being increasingly used for nonlinear problems. Researchers applied artificial intelligence techniques in modeling groundwater level [5], the predictions of suspended sediment [6] and dam inflow [7], modeling and prediction of water level in a dam [8, 9], solar radiation forecasting [10–12] and in various engineering fields [13–15]. Demirci and Baltaci [16] used fuzzy logic, multiple linear regression and traditional sediment rating curves to determine the amount of suspended sediment according to the 5-year continuous flow data of the Sacramento Freeport Station operated by the United States Geological Survey. As a result of this study, it was determined that the fuzzy logic method gave better results than other classical methods. Currently, there are numerous studies using artificial neural network [17,18], hybridized fuzzy [19], support vector machines [20], least-squares boost [21] and long short-term memory network [22] methods to estimate the amount of ET.

In the study by Kaya et al [23], the M5T method used in nonlinear physical events was used for ET estimation. The model results created with the M5T method were compared with the results obtained with the Turc empirical formula, and it was determined that the M5T method performed better. In the study by Yildirim et al [24], data from Samsun, Bafra and Carsamba meteorological stations were used. They used artificial neural network (ANN) and adaptive network-based fuzzy inference system (ANFIS) estimation models to create daily ET estimations. As a result of their study, it has been determined that ANN and ANFIS methods are useful methods for estimating ET. Ozel and Buyukyildiz [25] estimated the monthly evaporation amount of the Karaman meteorology station located in Konya, Turkey with ANN, epsilon-support vector regression (ε-SVR) and ANFIS methods. In their study, ε-SVR was the most successful method in estimating evaporation at the Karaman station.

Soft computing methods (fuzzy-SMRGT, ANN, MR, ANFIS, SMOReg) were used in this study to reduce the need for large data sets. The easily customizable simple membership function and fuzzy rule generation technique (fuzzy-SMRGT) based on experience is one of the models open to development to reduce the need for large data sets. Fuzzy-SMRGT, first proposed by Toprak [26]. was applied in modeling the relationship between any open canal flow (natural or artificial) and its hydraulics and geometrical properties. Altaş et. al [27] used fuzzy-SMRGT method to determine water surface profiles in open channels under different hydraulic conditions. It was shown that the results

were in good agreement with experimental measurements.

Indeed, there are not many regressive and empirical studies, and the fuzzy-SMRGT model has not been applied in any study for the estimation of ET in literature so far; thus, it was seen by the authors as an opportunity for ET studies. Besides, being a novelty of the study, the purpose of the study was making a comprehensive comparison between traditional and relatively new soft computing techniques. The calculations and models of selected soft computing techniques were created by using various meteorological parameters. In addition, ANFIS, another method used in the study is a flexible method that allows users to add prior knowledge to a neural network as a rule and has the ability to capture the nonlinear structure of a process with very fast convergence. However, the multi-linear regression (MLR) method is one of the classical comparison methods frequently used in the literature. MR is a multi-linear and nonlinear technique that can be adapted to new problems relatively easily, such as images, time series, language and long short-term memory, etc. In this study, prediction models were developed for ET, which has an important effect on water resources management. In estimation models, temperature, humidity in the atmosphere, wind speed and solar radiation were used as independent variables. ANN and fuzzy models from artificial intelligence techniques, which exhibit superior performance in nonlinear problems, were used. Also, the classical statistical MLR method and MR methods such as interaction (I-MR), pure quadratic (P-MR) and quadratic (Q-MR) methods were used and compared to new models. In this scope of the study, artificial intelligence techniques such as ANN, ANFIS and fuzzy-SMRGT and MR algorithms were used, and the results of the obtained estimation models were compared according to the reference ET amount.

2. Materials and methods

2.1. Case study and data used

In this study, data between 2008–2012 were taken from the station near Lewisville Lake (USGS, [28]), which has a surface area of 29592 decares. Models were created for the amount of ET by using the data obtained from Lewisville Lake, the image of which was given in Figure 1. In the models created, temperature (T), wind speed (WS), solar radiation (SR) and atmospheric humidity (H) given in the graphics in Figure 2 were used as independent variables. By using these independent variables, the ET amount was calculated empirically via the Penman-Monteith (PM) method, and these results were accepted as a reference.

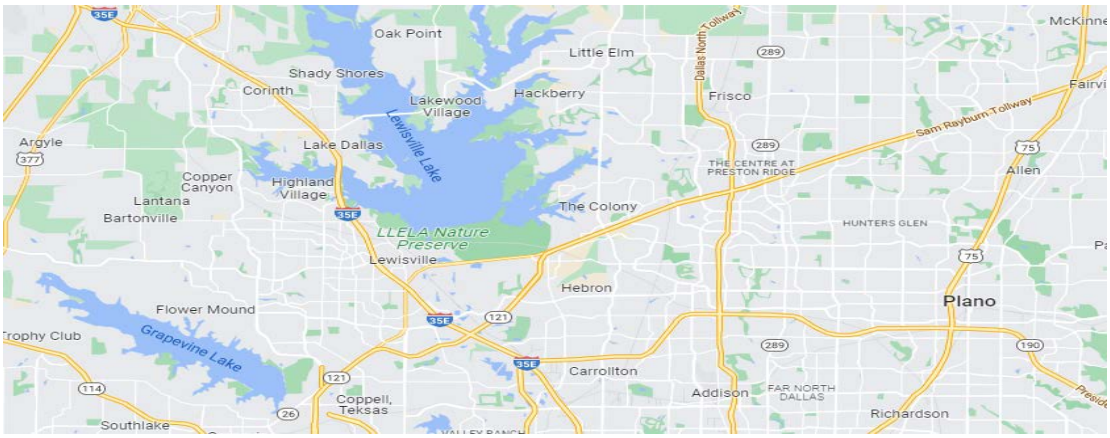

Figure 1. Satellite Image of Lewisville Lake (USGS).

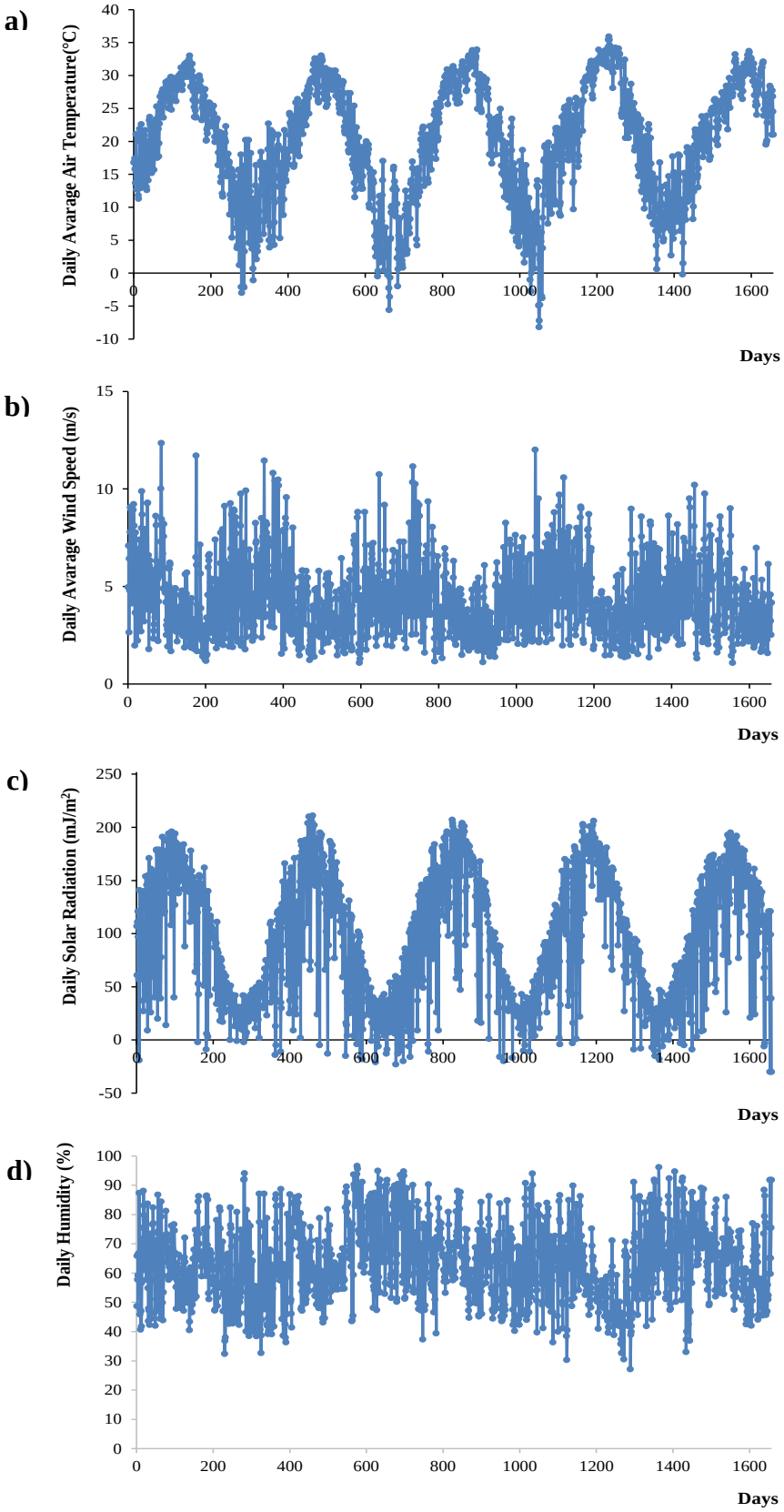

Figure 2. Variation of independent variables, T-H-WS-SR: **a)** Daily Average Air Temperature (T, °C), **b)** Daily Average Wind Speed (WS,m/s), **c)** Daily Solar Radiation (SR, mJ/m^2) and d) Daily Humidity (H,%).

2.2. Methods

In this study, the reference evapotranspiration (ET) amount obtained by the PM method was compared with the results of the models created by the MR, ANFIS, ANN, and Fuzzy-SMRGT methods.

2.2.1 Penman-Monteith (PM) Method

Penman argued in 1948 that the aerodynamic approach was not sufficient for the determination of the amount of evapotranspiration (ET) and that more accurate results could be obtained by using empirical expressions [29]**.** As a result, Penman devised the linear expression for estimating ET. Jensen et al [30] proposed that equation, which became known as 'Penman's linear equation' and was later developed. The expression of the PM method developed by Jensen et al [30] is given in Equation 1.

$$ET_0 = \left(\frac{\Delta}{\Delta+\gamma}\right) * R_n + \left(\frac{\gamma}{\gamma+\Delta}\right) * \left[\frac{15.36*(1+0.0062*U_2)*(e_w-e_a)*1}{\lambda}\right] \quad (1)$$

In the expression given in Equation 1, the notations ET_0, Δ, γ, R_n, U_2, e_w, e_a, λ represent the reference evapotranspiration, the slope of the vapor pressure curve, the physicometric constant, the clear radiation, the wind speed, the saturated vapor pressure, the actual vapor pressure and the potential heat of vaporization, respectively. In this study, amount of the evapotranspiration was calculated according to the PM method. The amount of ET (Fig 3) was calculated according to PM method which was considered as a reference for comparison of the model results.

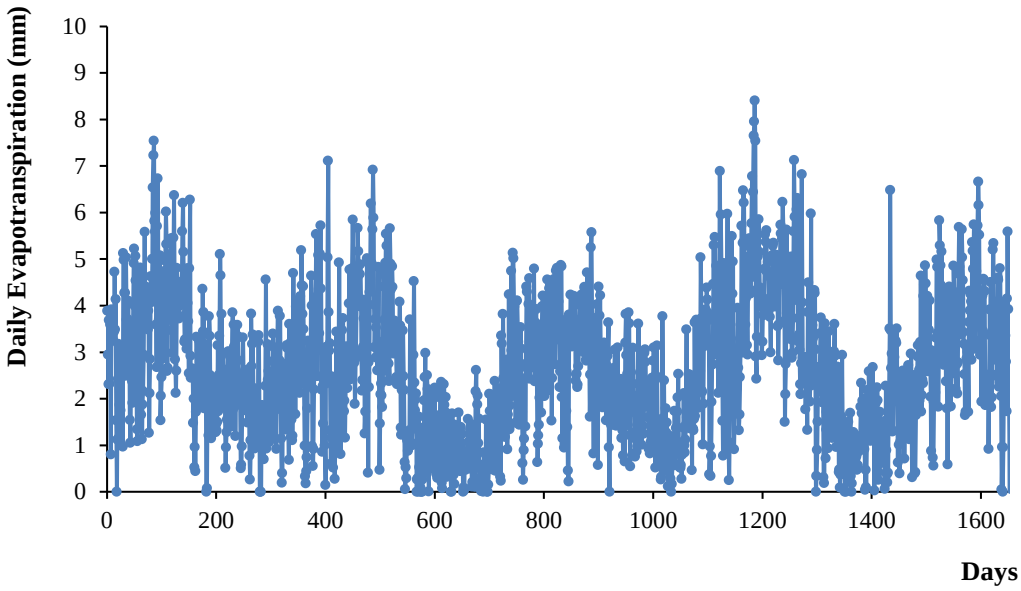

Figure 3. Daily reference ET (mm) change.

2.2.2. Multivariate Regression (MR)

a. Multi-Linear Regression (MLR)

It is accepted that there is a relationship between the variables in problems expressed with two or more variables. The general equation of the multi-linear regression method used to determine the effect of the independent variables on the dependent variables was given in Equation 2,

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \quad (2)$$

In the expression given in Equation 2 for the multi-linear regression method and in the subsequent equations given followingly, y represents the dependent variable, X_1, X_2,…, X_n represent the independent variables, β_0, β_1,..., β_n represent the regression coefficients, ε represents the error between the real value and the model.

b. Interaction Multivariate Regression (I-MR)

Multivariate Regression is a statistical method, used to estimate the dependent variables using independent variables and to determine the effect of the independent variables on the dependent variables. The models created by this method were expressed mathematically based on the regression coefficients obtained as a result of the regression analysis [31]. In this study, an interaction model from multivariate regression methods was created and the general form of the equation was given in Equation 3.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ,,, + \beta_n X_n + \beta_{n+1} X_1 X_2 + \beta_{n+2} X_1 X_3 + ,,, + \varepsilon \quad (3)$$

In the interaction expression given in Equation 3, there were terms in which constant, linear and interaction of the independent variables were included.

c. Pure Quadratic Multivariate Regression (P-MR)

In the pure quadratic regression method, which is one of the multivariate regression methods, the general expression of which was given in Equation 4; the relationship between independent variables and dependent variables was expressed in linear, constant and squared terms.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ,,, + \beta_n X_n + \beta_{n+1} X_1^2 + \beta_{n+2} X_2^2 + ,,, + \beta_m X_n^2 + \varepsilon \quad (4)$$

d. Quadratic Multivariate Regression (Q-MR)

It is a multivariate regression method in which the relationship between independent variables and dependent variables is expressed in linear, constant, interaction and squared terms. The general form of the expression representing the relationship between the variables was given in Equation 5.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ,,, + \beta_n X_n + \beta_{n+1} X_1 X_2 + \beta_{n+2} X_1 X_3 + ,,, + \beta_{n+2} X_1^2 + \beta_{n+3} X_2^2 + ,,, + \beta_m X_n^2 + \varepsilon \quad (5)$$

2.2.3. Artificial Neural Network (ANN)

The artificial neural network (ANN) method, which is defined as the black box model, works similarly to the working mechanism of nerve cells in humans [32]. In the ANN method, which is developed by taking the structure of the nervous system as a model, models are created without any pre-acceptance as in statistical methods [33]. In this method, which establishes a relationship between dependent and independent variables through mathematical algorithms, each independent variable is multiplied by different weights and transferred to the neurons in the next layer. These values distributed to the neurons are collected and passed through the transfer function.

In the study, while ANN models were created in Matlab software, Levenberg-Marquardt, Gradient Discent, Resilient Back-Propagation Learning Algorithms, Gradient Descent with Momentum training algorithms were used. A total of 12000 models were created using these training algorithms. While

creating the models, purelin, tagsig and logarithmic-sig were used as transfer functions for the hidden layer and output layer. The best of these models was determined according to the coefficient of determination (R^2), average percentage error (APE) and the mean of the squares of the errors (MSE). The general view of a loop of the created ANN model is given in Figure 4. This cycle was repeated based on the reference evapotranspiration amount and was limited to 10 epochs. When training on ANN, the first 100 models were used. Then re-changing the starting weights were changed 2900 times. A total of 3000 different models are obtained for each algorithm. Since a total of 12000 different models were created with 4 algorithms, 10 epochs were considered sufficient for each.

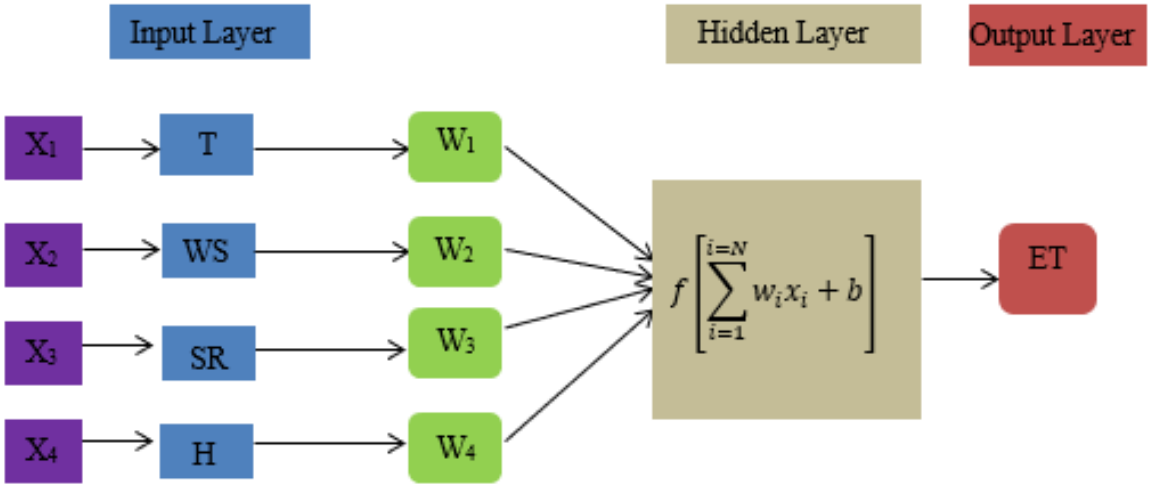

Figure 4. Loop of the ANN.

In the ANN cycle given in Figure 4, T represents temperature, WS wind speed, SR solar radiation, H represents atmospheric humidity, W weights, f transfer function, b bias value of the model and ET amount.

2.2.4. Adaptive Neural Fuzzy Inference System (ANFIS)

Fuzzy inference is a process of matching from a given input to an output dataset using the theory of fuzzy clusters [34]. Mamdani and Takagi-Sugeno (ANFIS) approaches are widely used in the literature. ANFIS, known as the Sugeno type fuzzy system, is a hybrid method in which ANN and fuzzy inference methods (FIS) are used together [35]. ANN proposed by Jang is used to solve nonlinear problems and complex time series [36,37]. The general structure of ANFIS is given in Figure 5. In this structure "T, WS, SR, H", "ET", "A, B, C, D", "∏", "N", "w" represent input parameters (independent variable), output parameter (dependent variable), membership functions, fuzzy rules, normalization, and 'w' weights, respectively.

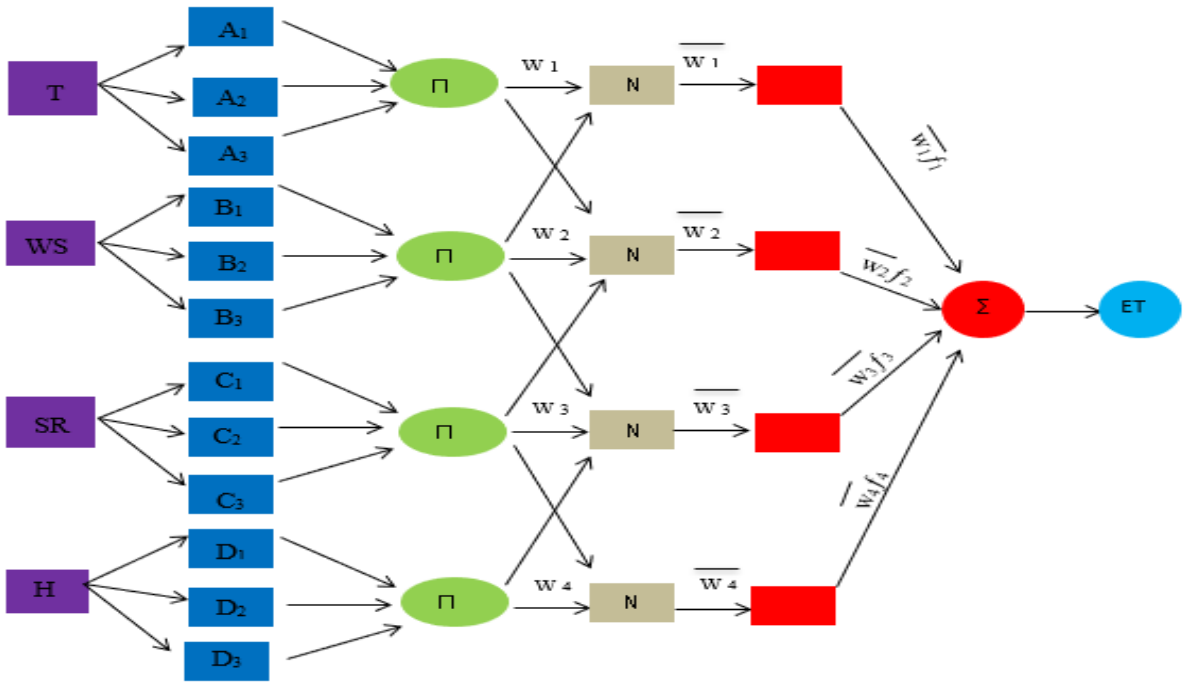

Figure 5. ANFIS structure used in this study.

In this method, which is modeled in 5 layers, the first layer is the fuzzification layer where membership functions are created for each variable. Fuzzy rules are created in the second layer, followed by the third layer, where normalization is performed to determine the appropriate weights. In the fourth layer, the signals from the third layer are multiplied in the polynomial function and processed with fuzzy rules. The total output value obtained from overall outputs is calculated. There is one node and it is represented by Σ.

2.2.5. Simple Membership Functions and Fuzzy Rules Generation Technique (Fuzzy-SMRGT)

Fuzzy-SMRGT is a rule-based modeling method. Determining the fuzzy rules (FRs) base and membership functions (MFs) together and simply is an important advantage of the method [38]. It is very important to understand how to generate FRs and MFs in the SMRGT method, which was developed to obtain the best model results in fuzzy systems where expert opinion is not sufficient.

In this study, the maximum and minimum values of the independent variables T, WS, SR and H were determined. Fuzzy clusters were formed based on these values. Then the shapes of the membership functions (MFs) were determined. Triangular membership functions (MFs) were used in the model. After deciding on the shape of the MFs, the membership function unit width, core value and key values were determined.Following this stage, the education process started with the creation of fuzzy rules (FRs) according to fuzzy clusters. In the Fuzzy-SMRGT flowchart given in Figure 6, the notations expressed with V_R, UW, Ci, Euw, n_u, K_1 and K_2 are parameters of MFs. V_R represents the difference between the maximum and minimum values of the variables, UW unit width, Ci core value, n_u represents the number of right triangles, K_1 and K_2 key values, Euw base width and were calculated with Equations (6) – (11).

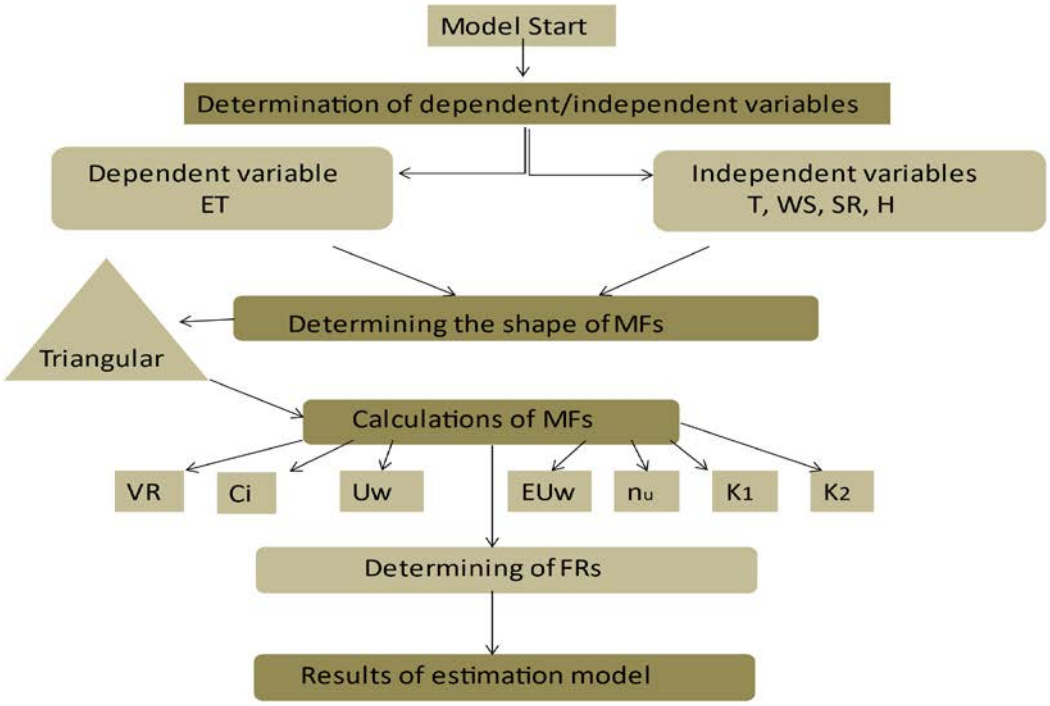

Figure 6. Fuzzy-SMRGT Flowchart.

$$V_R = (T, WS, SR, H, ET)_{max} - (T, WS, SR, H)_{min} \tag{6}$$

$$Ci = \frac{V_R}{2} - (T, WS, SR, H, ET)_{min} \tag{7}$$

$$UW = \frac{V_R}{n_U} \tag{8}$$

$$Euw = \frac{V_R}{n_U} + \frac{UW}{2} \tag{9}$$

$$K_1 = (T, WS, SR, H, ET)_{min} + \frac{Euw}{3} \tag{10}$$

$$K_2 = (T, WS, SR, H, ET)_{max} - \frac{Euw}{3} \tag{11}$$

Fuzzy-SMRGT model was created by calculating the parameters given in the Equations (6) – (11). The plots of the MFs for each variable were given in Figures (7) – (12).

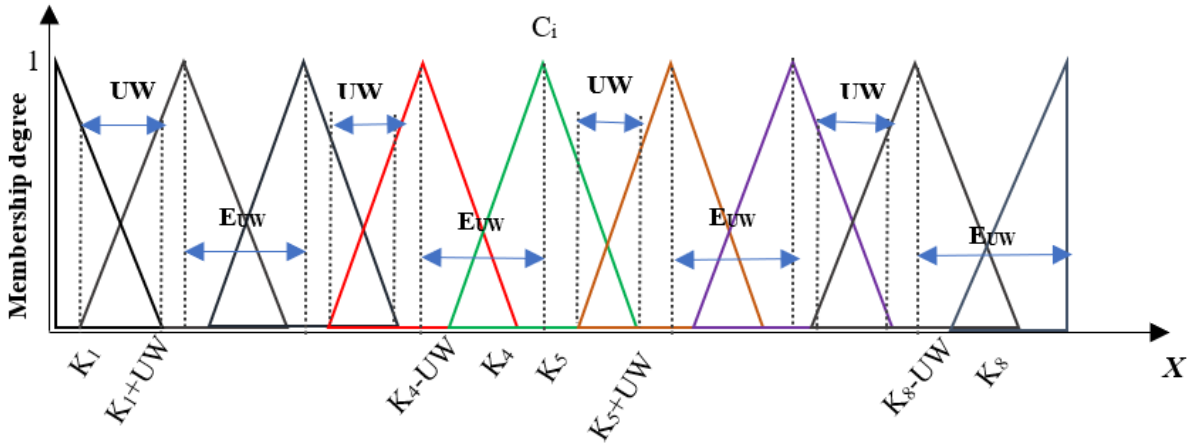

Figure 7. General SMRGT MFs plot for variables.

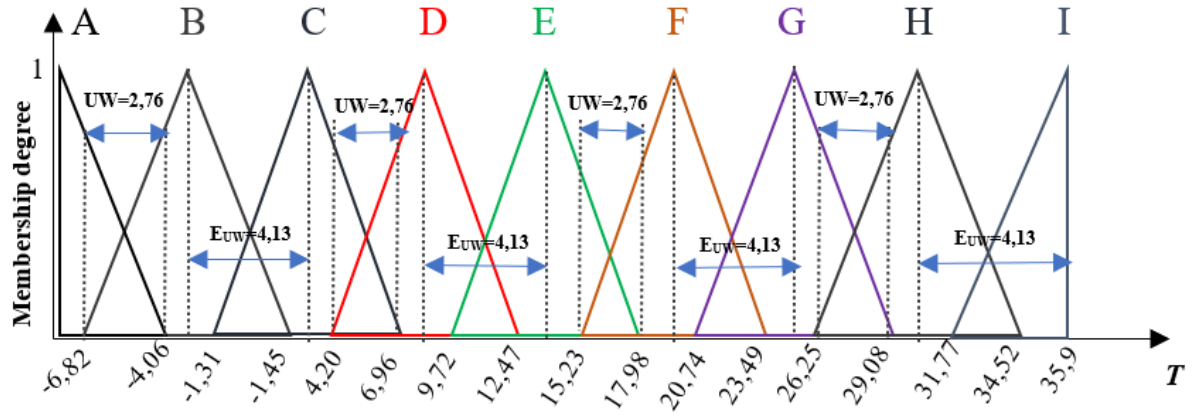

Figure 8. The SMRGT MFs plot for temperature variables.

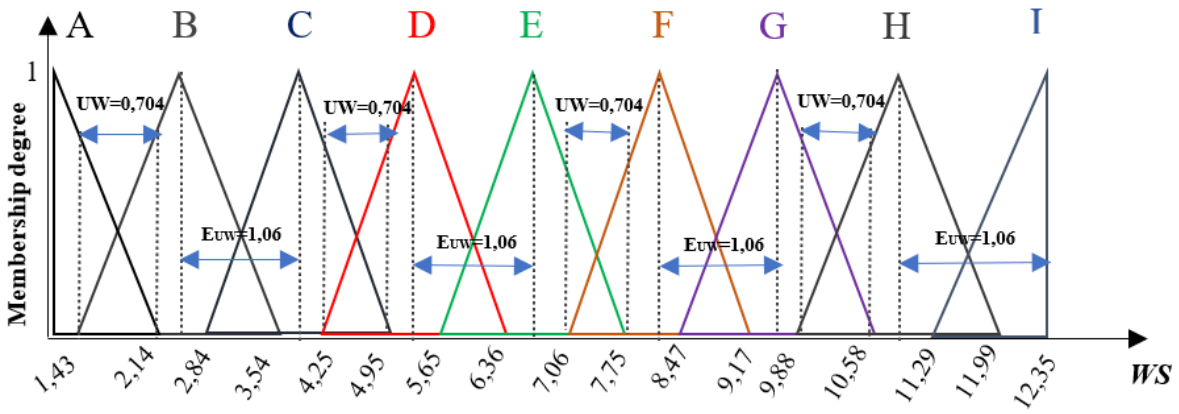

Figure 9. The SMRGT MFs plot for wind speed variable.

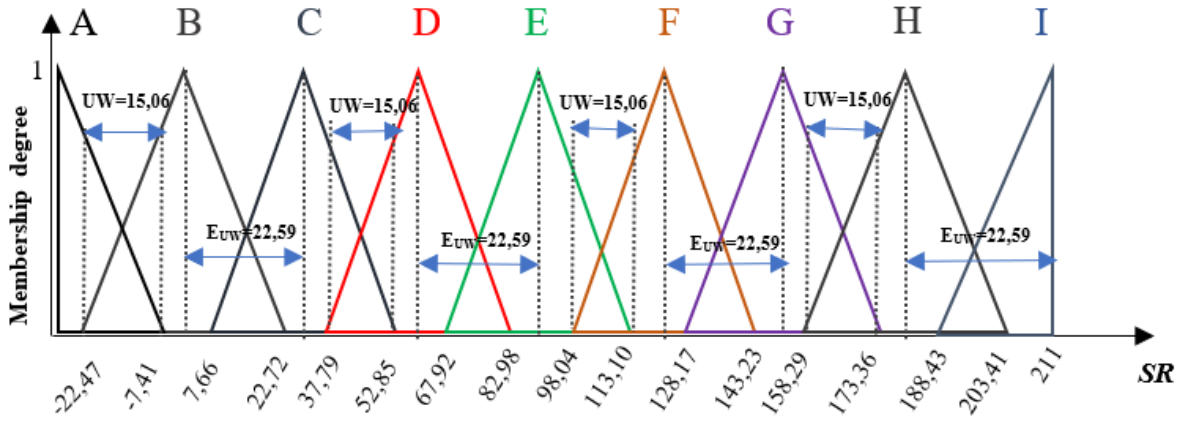

Figure 10. The SMRGT MFs plot for solar radiation variable.

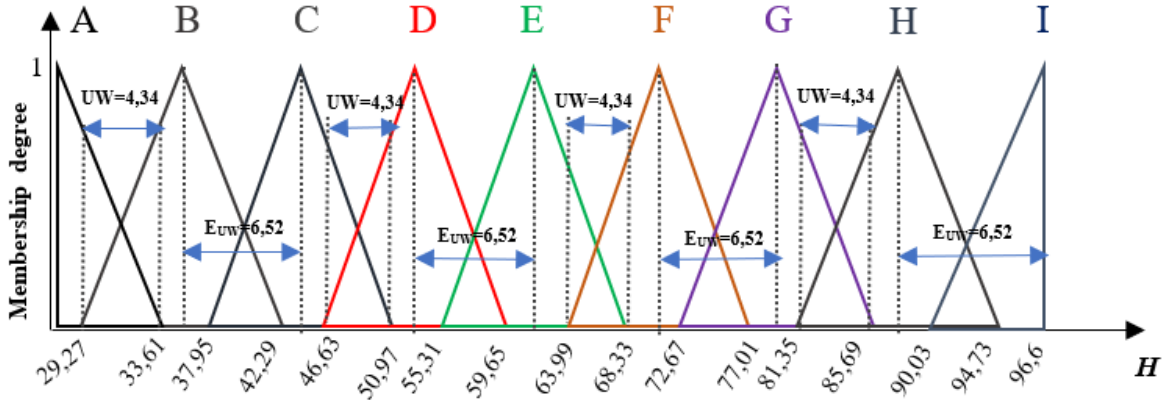

Figure 11. The SMRGT MFs plot for humidity in the atmosphere variable.

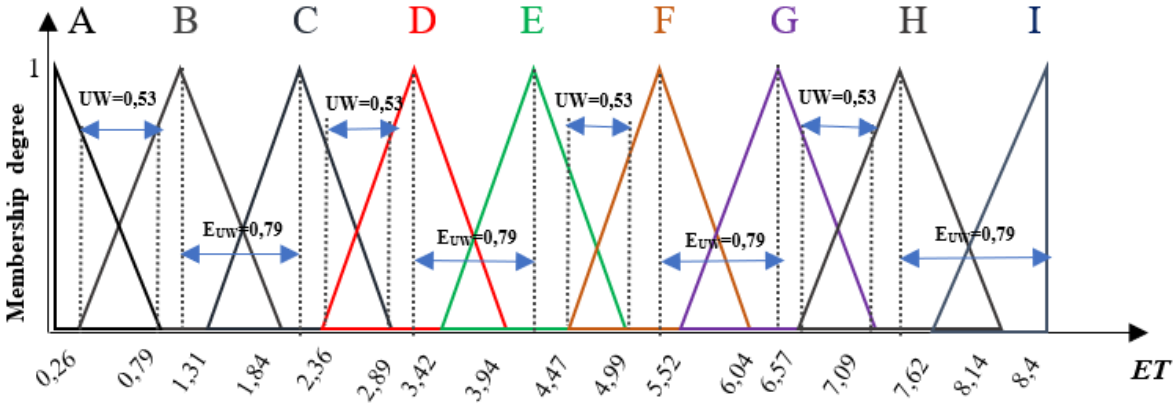

Figure 12. The SMRGT MFs plot for evapotranspiration dependent variable.

2.2.6. SMOReg Approach

The SMOReg method was developed based on the Support Vector Machines to create estimation models. In this study, the SMOReg method was applied by using WEKA open-source software. The Kernel type was selected as PolyKernel. The filter type was selected as Normalize training data and the regOptimizer was chosen as RegSMOImproved. The RegSMOImproved is one of the most popular algorithms and it is the default option of the SMOReg. More information about the used regOptimizer can be found in Shevade, Keerthi et al [39].

In this approach data set was divided into training and test sets exactly as it was done for other methods. Air temperature, wind speed, humidity, and solar radiation parameters were taken as inputs while evapotranspiration was considered as output. Further information about the base of the application of Support Vector Machines can be found in Smola and. Schoelkopf [40].

3. Analysis of the results and discussion

In the data set used in this study, 80% of the data was used for training and 20% for testing. The results of the estimation models, created for the amount of evapotranspiration, were compared according to the coefficient of determination (R^2), average percentage error (APE), and root mean square error (RMSE). The test performance results of the estimation models created by the analysis are presented in Table 1.

Table 1. Test performance results of the models.

Model	Model Inputs	RMSE (mm)	R^2	APE (%)
MLR	T, WS, SR, H	0,390	0,929	34,120
I-MR	T, WS, SR, H	0,164	0,988	10,537
P-MR	T, WS, SR, H	0,342	0,947	35,823
Q-MR	T, WS, SR, H	0,213	0,991	18,880
ANN	T, WS, SR, H	0,075	0,998	3,361
ANFIS	T, WS, SR, H	0,103	0,996	4,340
Fuzzy- SMRGT	T, WS, SR, H	0,361	0,946	40,335
SMOReg	T, WS, SR, H	0,399	0,927	35,860

3.1. ANN Results

Among the 12000 models created with ANN, the model using the Levenberg-Marquardt training algorithm, which consists of a hidden layer and an output layer, gave the best results. In this model, there are 16 neurons in the first layer and 1 neuron in the output layer. Tansig was used in the hidden layer and purelin was used in the output layer of the model as the transfer function. The appearance of this model was shown in Figure 13a. The mean squared error (MSE), training phase and correlation graphs of the ANN model were given in Figure 13b and Figure 13c, respectively.

When the results obtained with this method were compared with the reference values, the R^2, RMSE and APE values were determined as 0,998, 0,075 and 3,361%, respectively. The distribution and scatter plots obtained with this estimation model were given in Figure 14 and Figure 15.

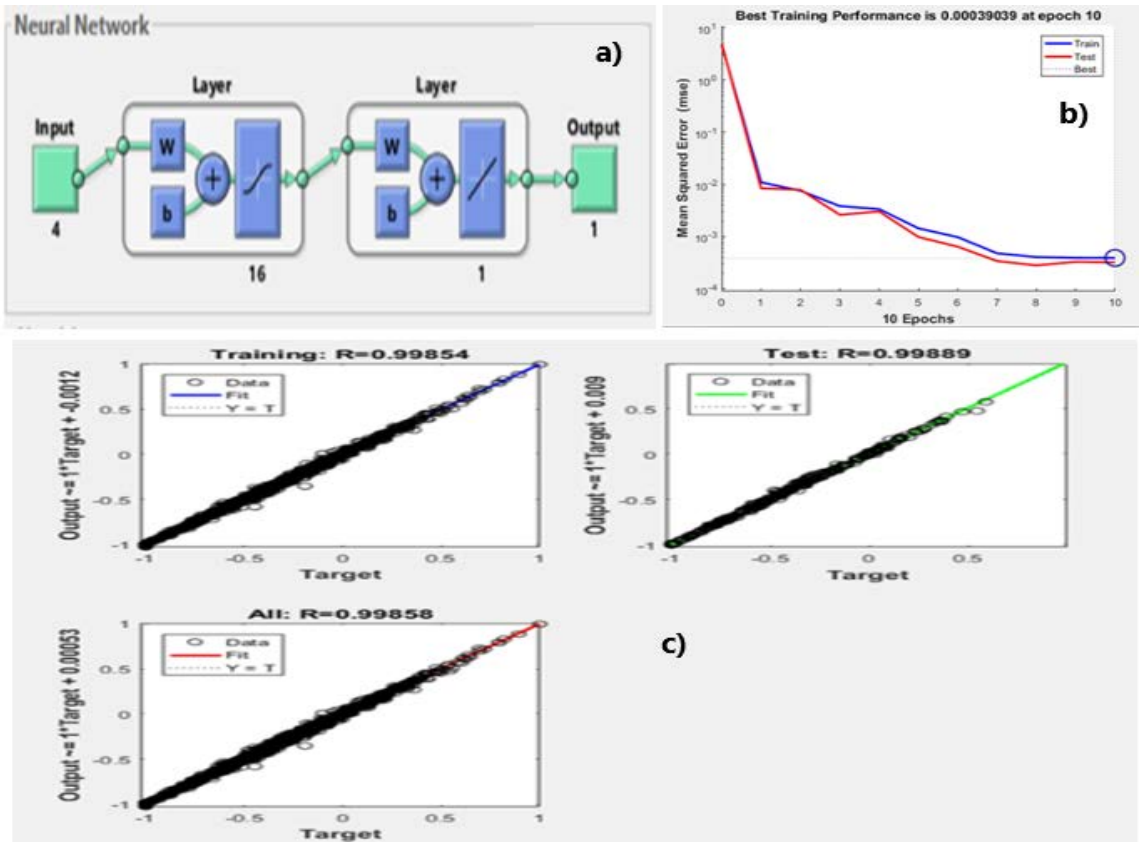

Figure 13. a) The best performing ANN network model **b)** Change of MSE Value of ANN Model **c)** Result of ANN Model.

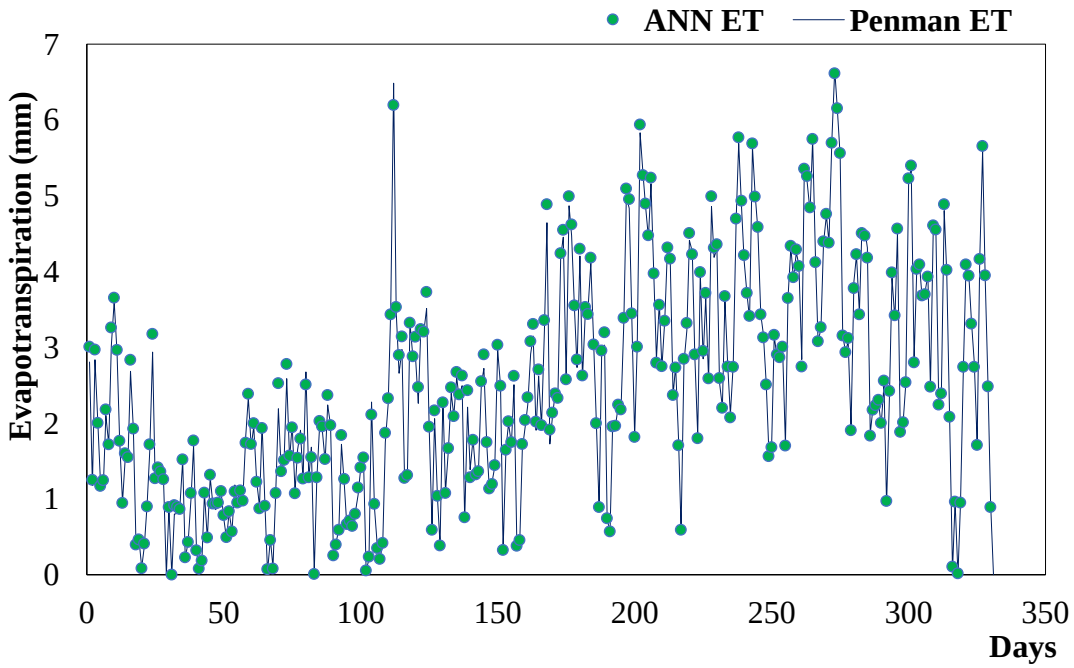

Figure 14. ANN Method Distribution Plot.

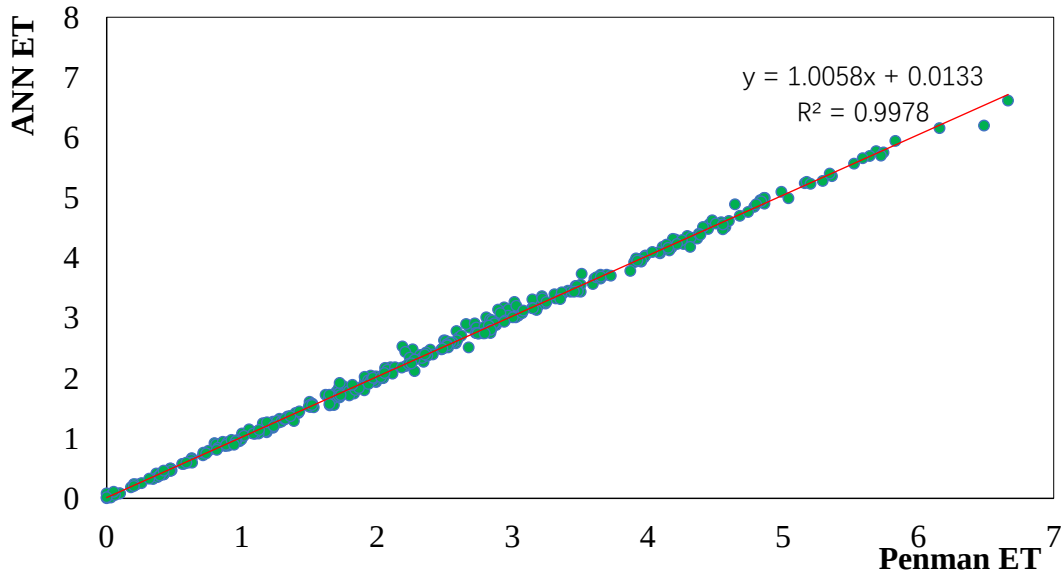

Figure 15. ANN Method Scatter Plot.

3.2. ANFIS Results

As in other models, temperature (T), humidity in the atmosphere (H), wind speed (WS) and solar radiation (SR) variables were also used in the ANFIS method and the result of the ANFIS model was shown Figure 16 and Figure 17.

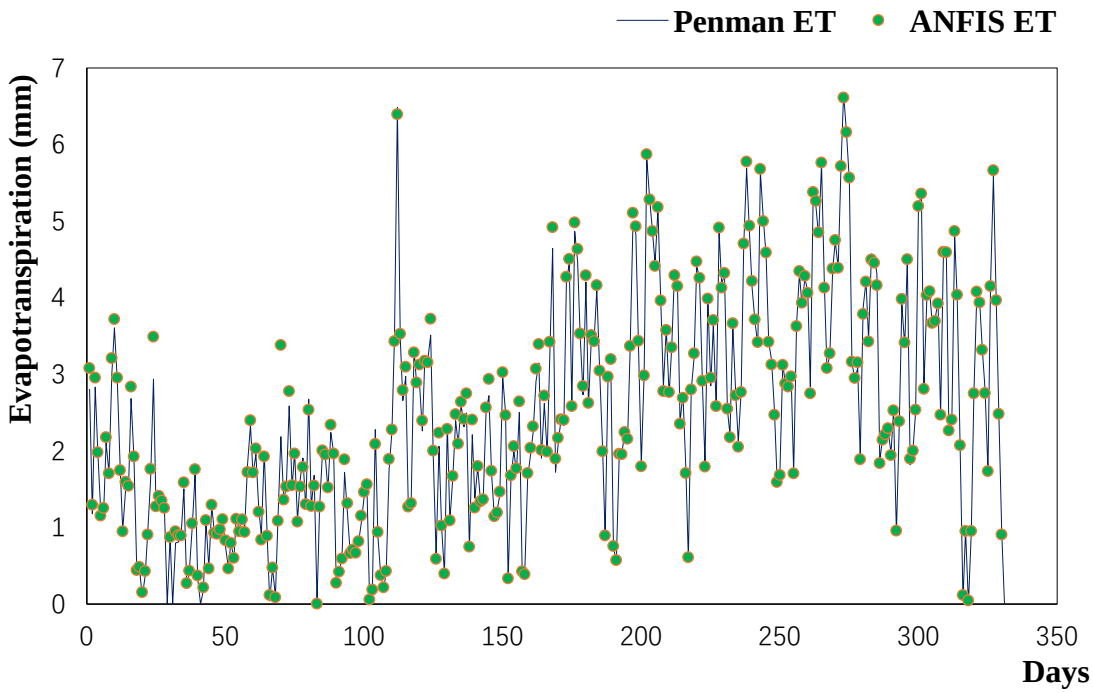

Figure 16. ANFIS Method Distribution Plot

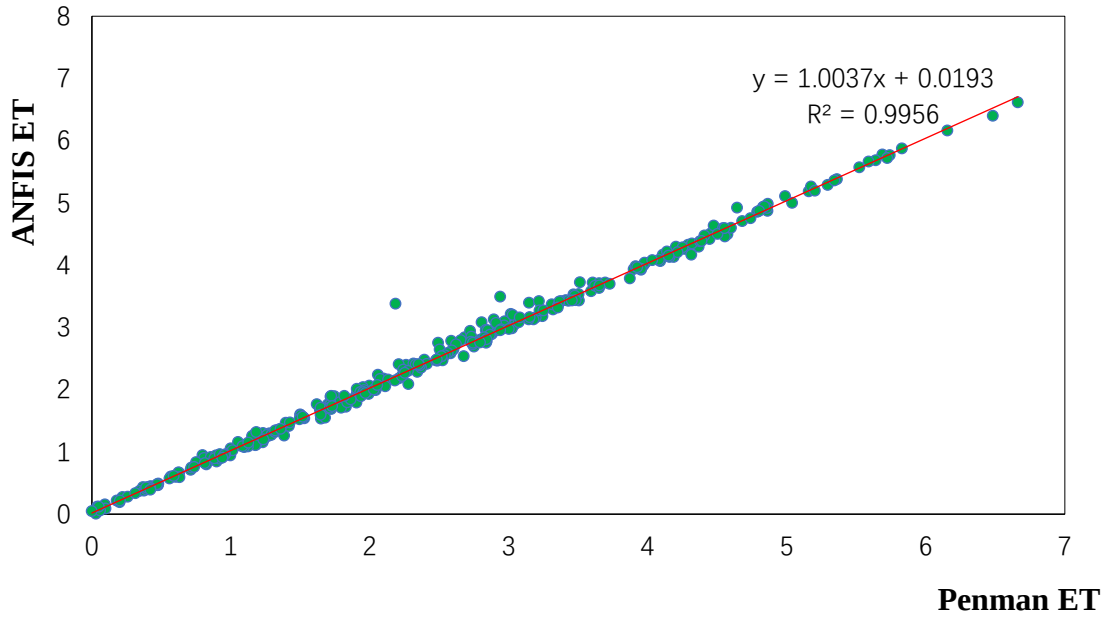

Figure 17. ANFIS Method Scatter Plot

It was seen that the results of the ANFIS estimation model, whose distribution and scatter plots were given in Figure 16 and Figure 17, were quite compatible with the reference values. The coefficient of determination (R^2) of the model created by this method was obtained as 0,996. In addition, with the calculations made, correlation, the root mean square error (RMSE), and average percentage error (APE) values of the model were determined as 0,996 0,103 and 4,340 %, respectively.

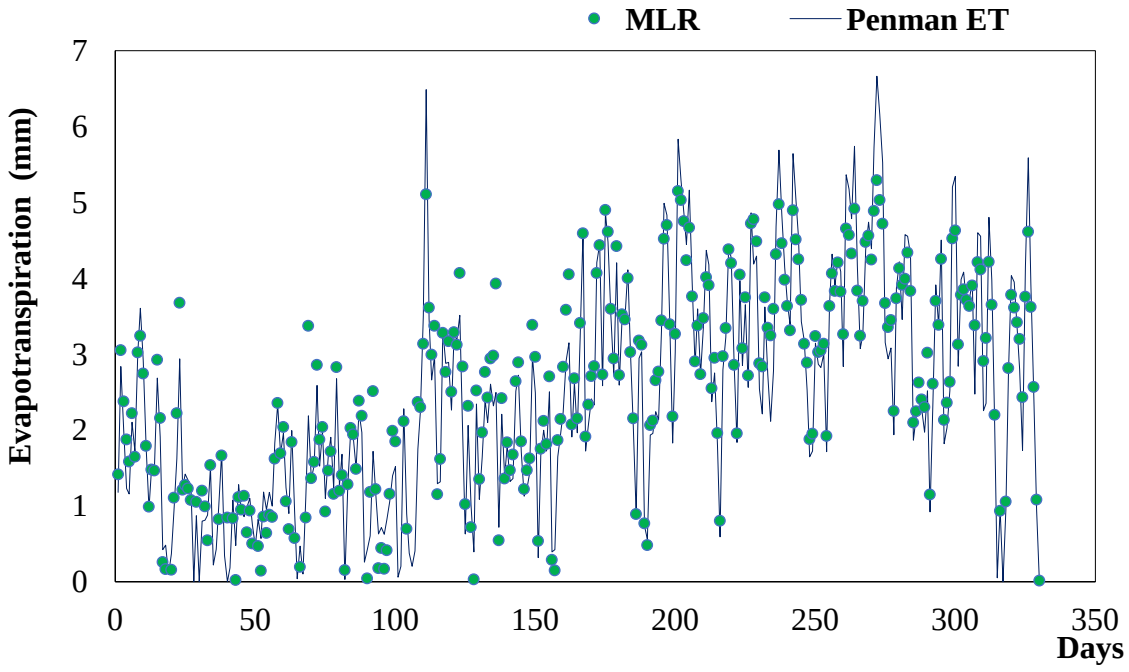

Figure 18. MLR Method Distribution Plot.

3.3. MLR Results

For MLR models like in ANFIS and the other models, temperature (T), humidity in the atmosphere (H), wind speed (WS), and solar radiation (SR) were used for estimation of amount of evapotranspiration (ET) as independent variables. As a result of the analysis made with this method, a linear equation given in Equation 12 was obtained. The distribution and scatter plots of this model were given in Figure 18 and Figure 19.

$$ET = 2{,}332785 + 0{,}70282 * T + 0{,}007782 * WS + 0{,}382118 * H - 0{,}05613 * SR \quad (12)$$

As a result of the linear regression analysis, it was seen that there was a good agreement between the estimation model and the observed results. The determination coefficient value of the model created as a result of the regression analysis was 0,929, the root mean squares error was 0,390 and the average percentage error was 34,120 %.

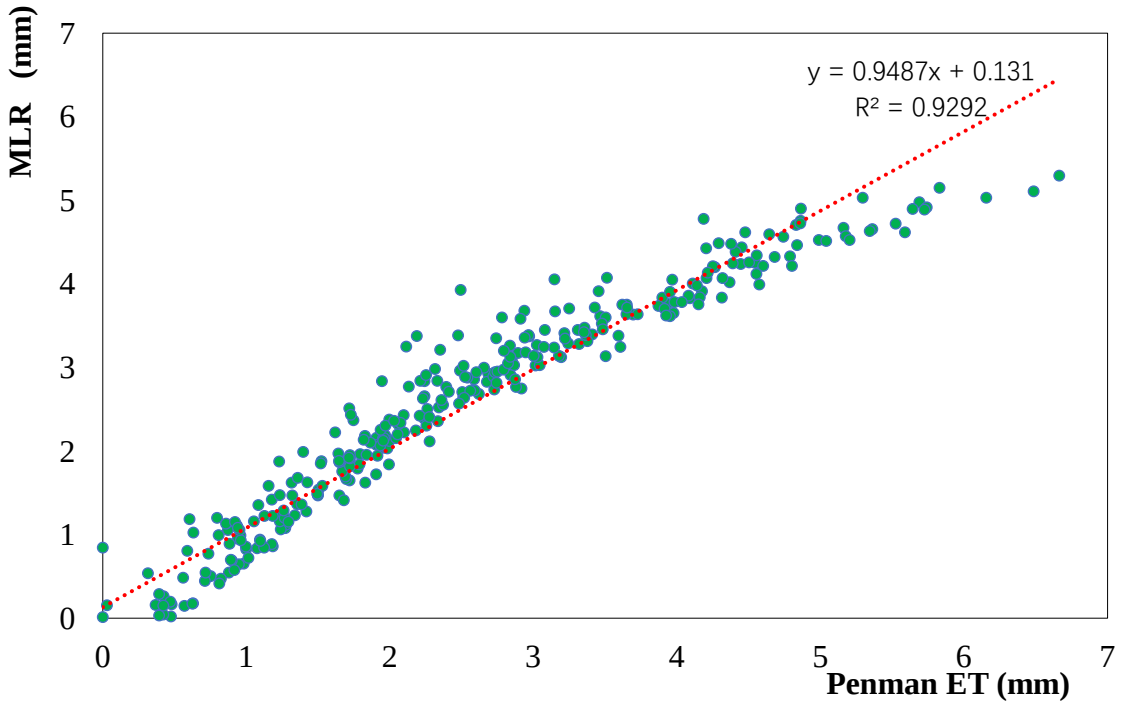

Figure 19. MLR Method Scatter Plot.

3.4. I-MR Results

The expression obtained by the interaction method in multivariate regression methods was given in Equation 13. The distribution and scatter plots of this model were given in Figure 20 and Figure 21.

$$ET = -0{,}371505445 + 0{,}111745468 * T + 0{,}670952906 * WS + 0{,}000642904 * SR + 0{,}012219651 * H + 0{,}016117455 * T * WS + 0{,}000235816 * T * SR - 0{,}002020822 * T * H - 0{,}000149775 * WS * SR - 0{,}008955872 * WS * H + 2{,}55e^{-5} * SR * H \quad (13)$$

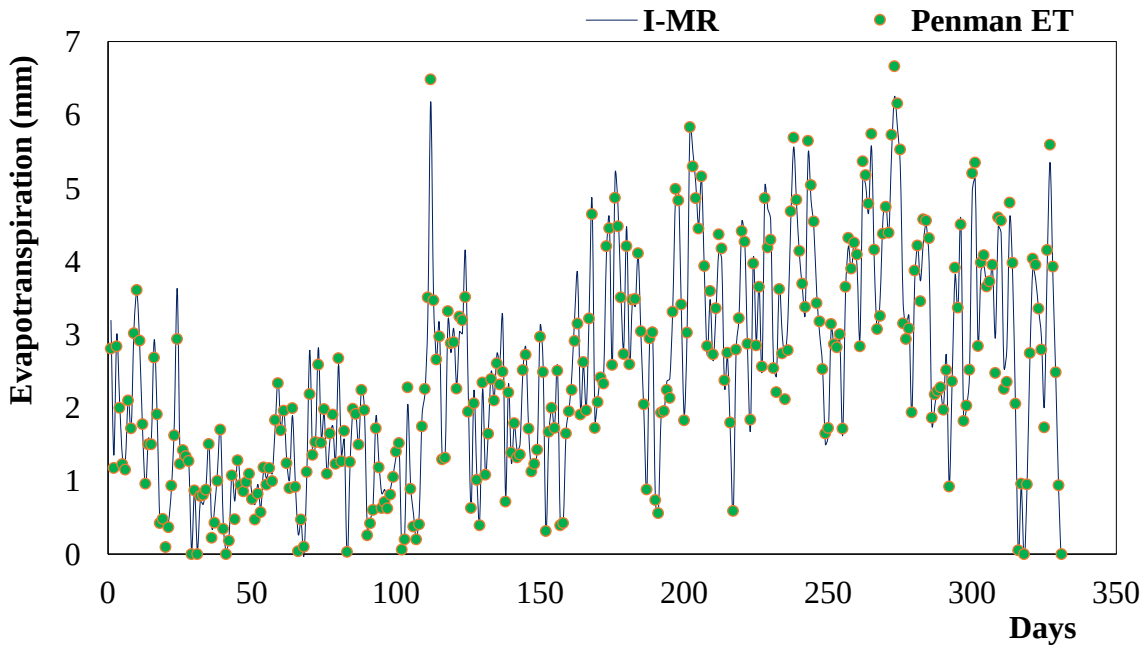

Figure 20. I-MR Method Distribution Plot.

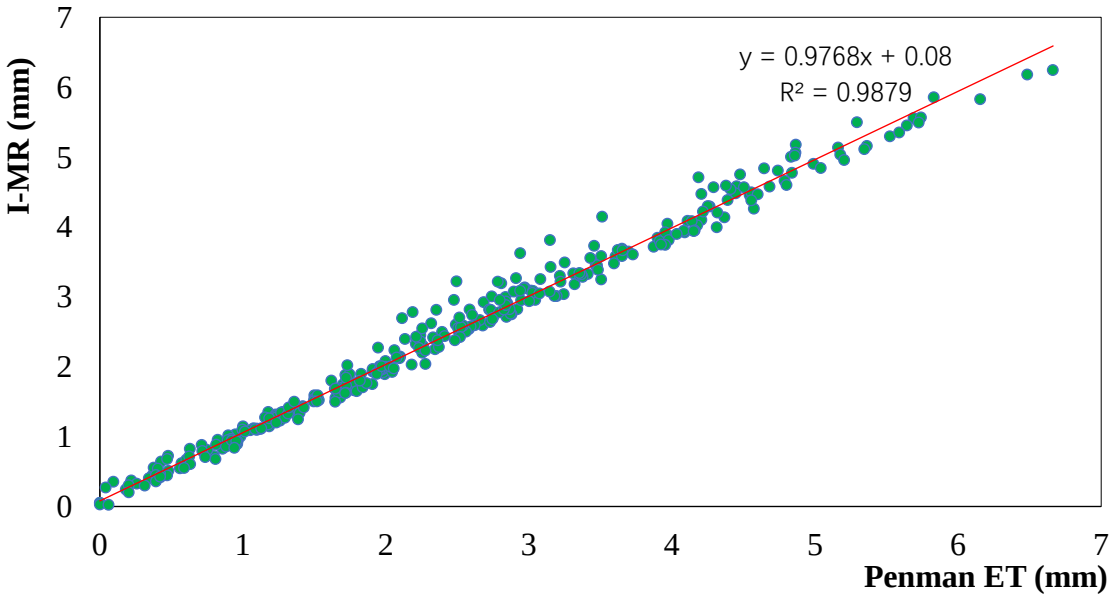

Figure 21. I-MR Method Scatter Plot.

As can be seen from the graphs, the results obtained in the model created by the interaction regression method were found to be quite compatible with the reference values. The coefficient of determination (R^2) of this model was obtained as 0,988, the root mean square error (RMSE) was 0,164 and the average percent error (APE) of the model was 10,537 %.

3.5. P-MR Results

The expression obtained for the pure-quadratic regression model, whose distribution and scatter plots are shown in Figures 22 and 23, is given in Equation 14.

$$ET = 2{,}888431 + 0{,}048029 * T + 0{,}751878144 * WS + 0{,}004537 * SR - 0{,}093264589 * H + 0{,}00067 * T^2 - 0{,}03475 * (WS)^2 + 1{,}75 * 10^{-5}(SR)^2 + 0{,}000284 * H^2 \quad (14)$$

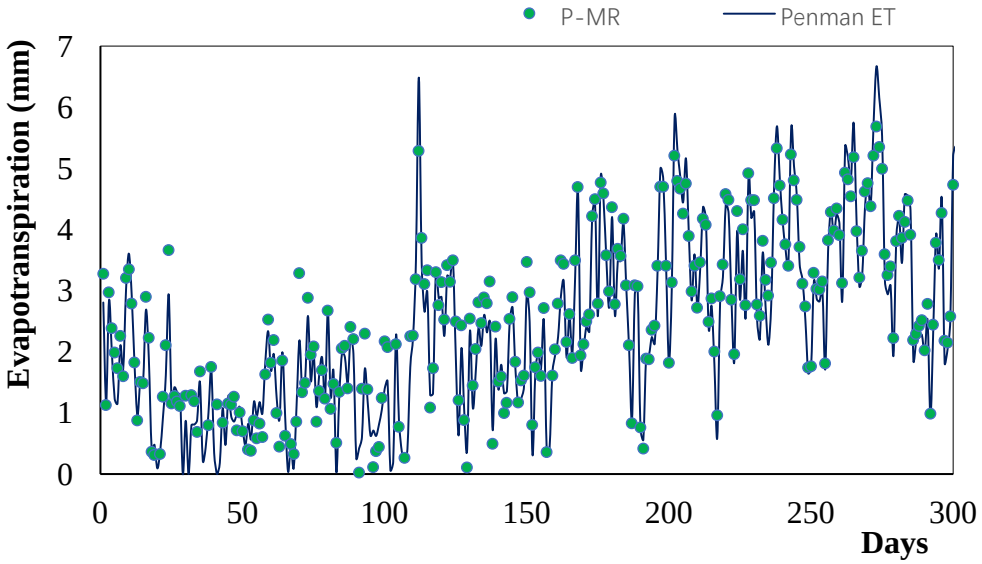

Figure 22. P-MR Method Distribution Plot.

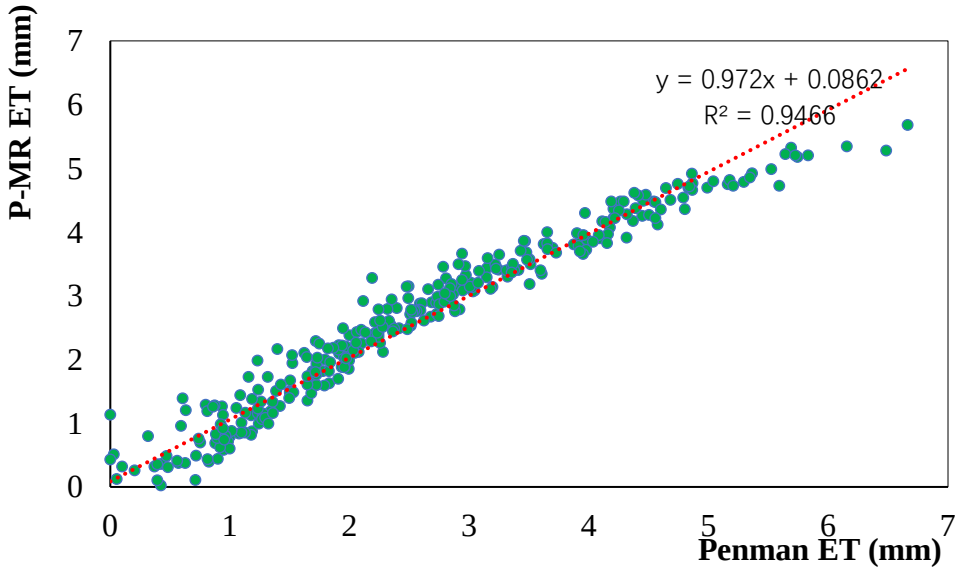

Figure 23. P-MR Method Scatter Plot.

According to the distribution and scatter plots given in Figure 22 and Figure 23, it was understood that the results obtained with the pure-quadratic estimation model were compatible with the reference values. With the estimation model created, R^2, RMSE and APE values were obtained as 0,947, 0,342, 35,823 % respectively.

3.6. Q-MR Results

The expression obtained by the estimation model created by the quadratic regression method was given in Equation 15. The distribution and scatter plots of this method were given in Figure 24 and Figure 25.

$$ET = -0{,}72999 + 0{,}116906 * T + 0{,}890184873 * WS + 0{,}000338 * SR + 0{,}006055103 * H + 0{,}01422 * T * WS + 0{,}000123 * T * SR - 0{,}00208 * T * H - 3{,}37 * 10^{-5} * WS * SR - 0{,}0086 * WS * H + 4{,}41 * 10^{-5} * SR * H + 0{,}000464 * T^2 - 0{,}02071 * (WS)^2 + 5{,}12 * 10^{-6} * (SR)^2 + 3{,}39 * 10^{-5} * H^2 \quad (15)$$

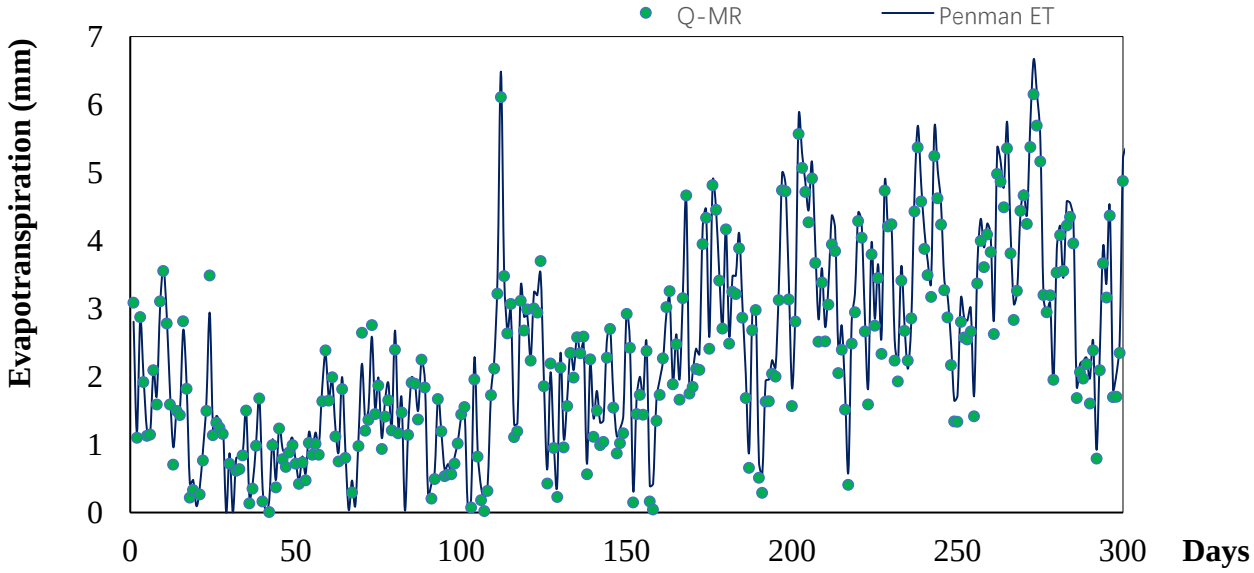

Figure 24. Q-MR Method Distribution Plot.

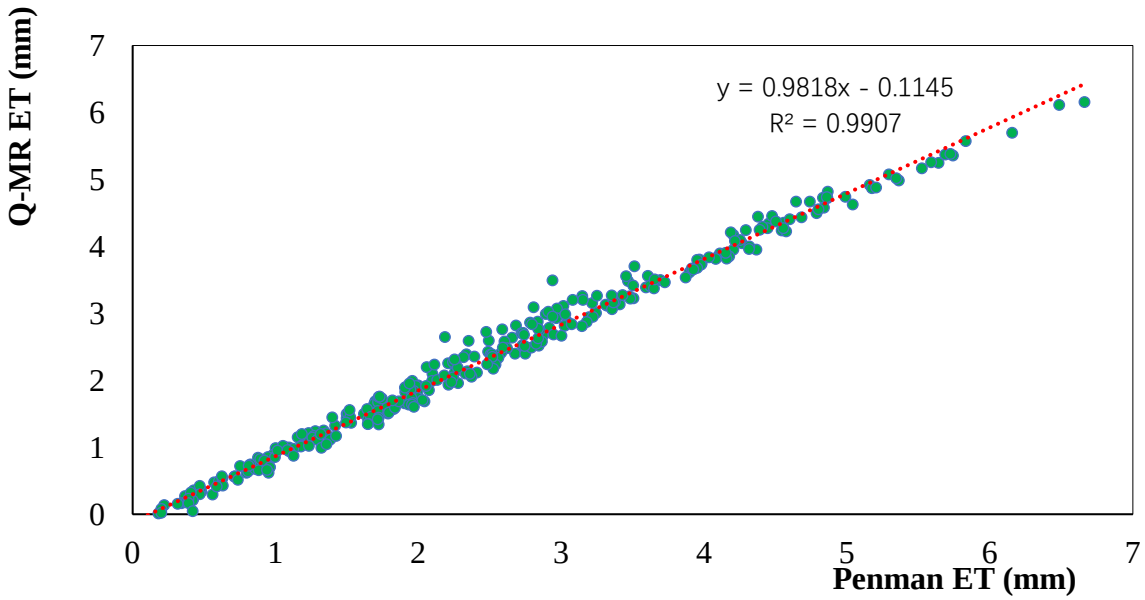

Figure 25. Q-MR Method Scatter Plot.

It was observed that the evapotranspiration values obtained with the estimation model created using the quadratic regression method and the reference values were compatible with each other. The coefficient of determination value of the model created by the quadratic regression method is 0,991, the root mean squares error is 0,213 and the mean percent error is 18,881 %.

3.7. Fuzzy-SMRGT Results

Fuzzy-SMRGT estimation model distribution and scatter plots, which were created by dividing each of the independent variables used in all other methods into five fuzzy clusters, were given in Figures 26 and Figure 27.

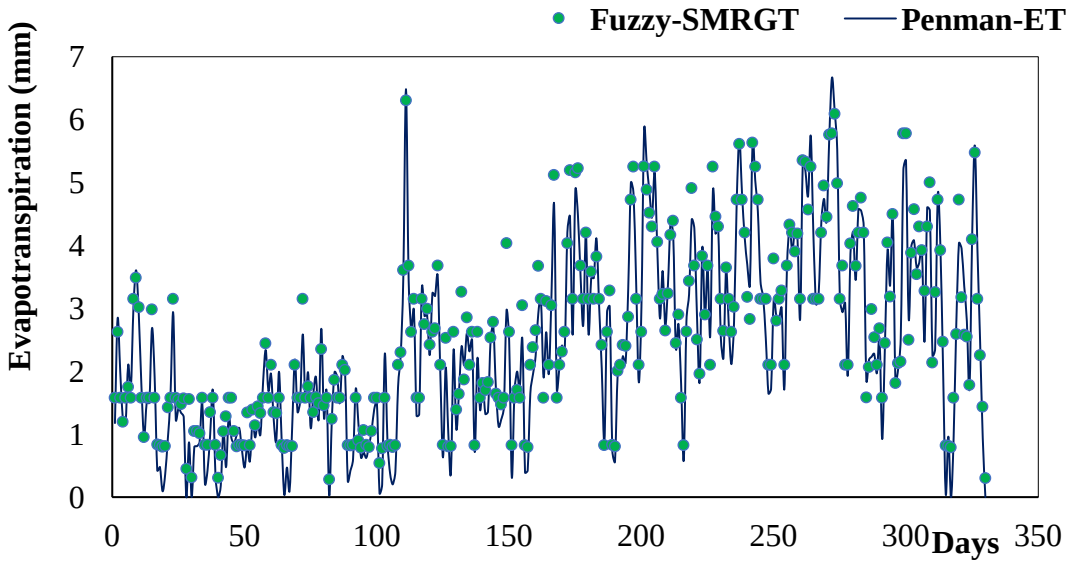

Figure 26. Fuzzy-SMRGT Method Distribution Plot.

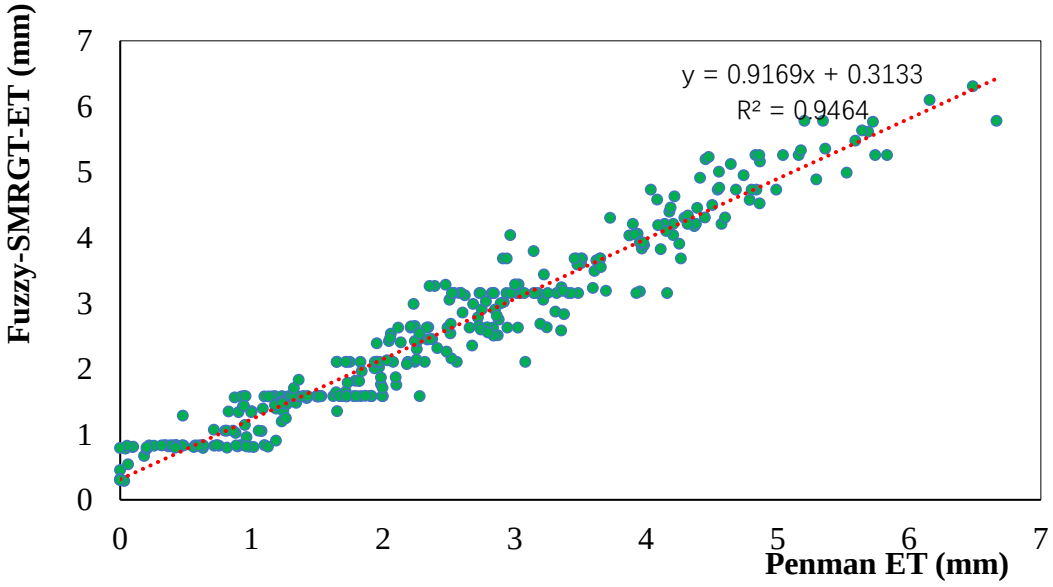

Figure 27. Fuzzy-SMRGT Method Scatter Plot.

As a result of the analysis using this method, the R^2, RMSE and APE values of the model were obtained as 0,973, 0,361, 40,335 % respectively. When these values were examined, it was seen that the model results were compatible with the reference values according to the R^2, RMSE and APE values.

3.8. SMOReg results

SMOReg approach which is based on Support Vector Machines was applied to the same data set. ET results were estimated slightly close to the simple MLR method. The determination coefficient was calculated as 0,927. The distribution graph of the SMOReg was given in Figure 28.

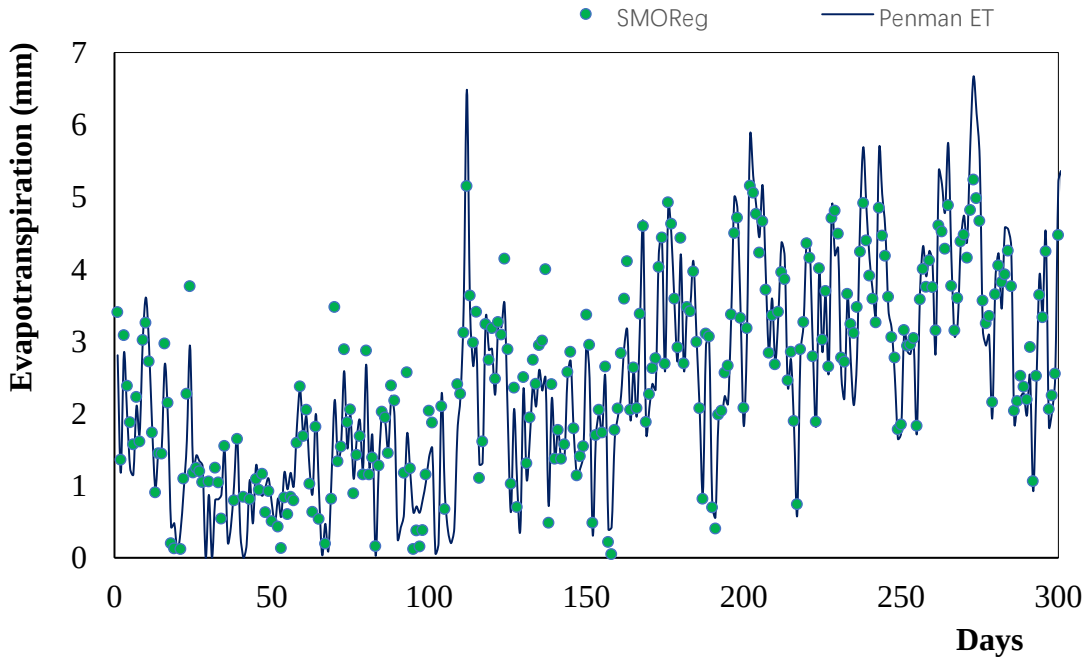

Figure 28. SMOReg Method Distribution Plot.

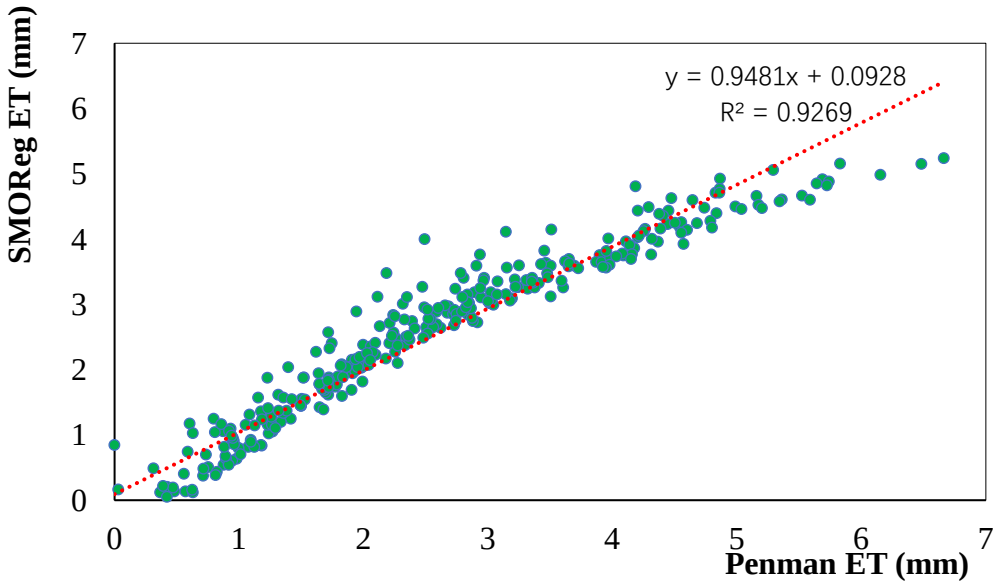

Figure 29. SMOReg Method Scatter Plot.

The correlation between the Penman ET results and the SMOReg results was given in Figure 28. Tthe scatter plot belonging to the SMOReg results was shared in Figure 29.

Figures 28 and 29 showed that the SMOReg approach gave also acceptable results. However, considering the other approaches, the SMOReg approach was one of the worst approaches among the used methods in the case of ET estimation of the mentioned study area.

4. Conclusions

In this study, using the T, WS, SR and H independent variables, evapotranspiration (ET) estimation models were developed. Models were created by ANN, Fuzzy-SMRGT, ANFIS and multivariate regression (MLR, I-MR, P-MR and Q-MR) methods. While creating the models, 80% of the data set was used as training and 20% as a test. The estimation models were compared with the ET amount obtained by the Penman method which was taken as a reference. Model results were compared according to RMSE, R^2 and APE statistical values. Based on the results of all estimation models, the reference ET amount and the model outputs were found to be compatible. It was determined that all the models were usable for ET estimation. Because the evapotranspiration values obtained from all the estimation models created were quite compatible with the reference ET.

ANN gave the best results when compared in terms of R^2, APE and RMSE. When other methods were compared in terms of R^2, APE and RMSE, it was determined that the ANFIS estimation model gave the best results following ANN. When the other estimation models were compared according to R^2, the best results were obtained with Q-MR, I-MR, P-MR, Fuzzy-SMRGT, MLR, and SMOReg methods, respectively. However, when compared in terms of RMSE, this ranking was formed as I-MR, Q-MR, P-MR, Fuzzy-SMRT, MLR, and SMOReg from best to worst, respectively. When the methods were compared in terms of APE, the best results were determined as I-MR, Q-MR, MLR, P-MR, SMOReg, and Fuzzy-SMRGT, respectively.

Although the estimation models created within the scope of this study have comparison parameters that have advantages and disadvantages when compared to each other, it has been determined that all the models created are quite useful in estimating the ET. In future studies, the effect of each input on the ET parameter, or the ET event with 2-3 different input variations can be investigated. Also, with the development of Fuzzy-SMRGT, ANFIS and ANN methods, the performance of different models in ET prediction can be researched.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. C. İnal, P. Fakioğlu, S. Bülbül, Determination of sediment volumes in dams with hydrographic surveys, *Selcuk University J. Eng. Sci. Technol.*, **3** **(**2015), 1–12. https://dergipark.org.tr/en/pub/sujest/issue/23197/247765 (accessed March, 2023)
2. B. Taşar, F. Üneş, M. Demirci, Y. Z. Kaya, Forecasting of daily evaporation amounts using artificial neural networks technique, *Dicle University J. Eng.*, **9** (2018), 543–551. https://doi.org/10.1002/CLEN.200900238

3. Y. Z. Kaya, M. Zelenakova, F. Üneş, M. Demirci, H. Hlavata, P. Mesaros, Estimation of daily evapotranspiration in Košice City (Slovakia) using several soft computing techniques, *Theor. Appl. Climatol.,* **144** (2021), 287–298. https://doi.org/10.1007/S00704-021-03525-Z
4. F. Üneş, Y. Z. Kaya, M. Mamak, Daily reference evapotranspiration prediction based on climatic conditions applying different data mining techniques and empirical equations, *Theor. Appl. Climatol*., **141** (2020), 763–773. https://doi.org/10.1007/S00704-020-03225-0/TABLES/4
5. M. Demirci, F. Üneş, S. Körlü, Modeling of groundwater level using artificial intelligence techniques: A case study of Reyhanli region in Turkey, *Appl. Ecol. Environ. Res*., **17** (2019), 2651–2663. https://doi.org/10.15666/AEER/1702_26512663
6. F. Üneş, B. Taşar, M. Demirci, M. Zelenakova, Y. Z. Kaya, H. Varçin, Daily suspended sediment prediction using seasonal time series and artificial intelligence techniques, *Rocznik Ochrona Środowiska*, **23** (2021). https://doi.org/10.54740/ros.2021.008
7. H. S. Choi, J. H. Kim, E. H. Lee, S. K. Yoon, Development of a revised multi-layer perceptron model for dam inflow prediction, *Water*, **14** (2022), 1878. https://doi.org/10.3390/W14121878
8. H. Leyla, S. Nadia, R. Bouchrit, Modeling and predictive analyses related to piezometric level in an earth dam using a back propagation neural network in comparison on non-linear regression, *Model. Earth Syst. Environ.,* **9** (2022), 1169–1180. https://doi.org/10.1007/S40808-022-01558-5
9. Y. Ouma, D. Moalafhi, G. Anderson, B. Nkwae, P. Odirile, B. P. Parida, et al., Dam water level prediction using vector autoregression, random forest regression and MLP-ANN models based on Land-use and climate factors, *Sustainability*, **14** (2022), 14934. https://doi.org/10.3390/su142214934
10. M. Guermoui, F. Melgani, K. Gairaa, M. L. Mekhalfi, A comprehensive review of hybrid models for solar radiation forecasting, *J. Clean. Prod.,* **258** (2020), 120357. https://doi.org/10.1016/J.JCLEPRO.2020.120357
11. M. Guermoui, S. Benkaciali, K. Gairaa, K. Bouchouicha, T. Boulmaiz, J. W. Boland, A novel ensemble learning approach for hourly global solar radiation forecasting, *Neural Comput. Appl.*, **34** (2022), 2983–3005. https://doi.org/10.1007/S00521-021-06421-9
12. M. Guermoui, K. Gairaa, K. Ferkous, D. S. O. Santos, T. Arrif, A. Belaid, Potential assessment of the TVF-EMD algorithm in forecasting hourly global solar radiation: Review and case studies, *J. Clean. Prod.,* **385** (2023), 135680. https://doi.org/10.1016/j.jclepro.2022.135680
13. I. Karatas, A. Budak, Prediction of labor activity recognition in construction with machine learning algorithms, *Icontech Int. J.*, **5** (2021), 38–47. https://doi.org/10.46291/ICONTECHvol5iss3pp38-47
14. C. Kayadelen, G. Altay, S. Önal, Y. Önal, Sequential minimal optimization for local scour around bridge piers, *Mar. Georesour. Geotec.,* **40** (2021), 462–472. https://doi.org/10.1080/1064119X.2021.1907635
15. C. Kayadelen, G. Altay, Y. Önal, Numerical simulation and novel methodology on resilient modulus for traffic loading on road embankment, *Int. J. Pavem. Eng.,* **23** (2021), 3212–3221. https://doi.org/10.1080/10298436.2021.1886296
16. M. Demirci, A. Baltaci, Prediction of suspended sediment in river using fuzzy logic and multilinear regression approaches, *Neural Comput. Appl.,* **23** (2013), 145–151. https://doi.org/10.1007/S00521-012-1280-Z
17. M. Achite, M. Jehanzaib, M. T. Sattari, A. K. Toubal, N. Elshaboury, A. Wałęga, et al., Modern techniques to modeling reference evapotranspiration in a semiarid area based on ANN and GEP models, *Water,* **14** (2022), 1210. https://doi.org/10.3390/W14081210

18. F. Üneş, S. Doğan, B. Taşar, Y. Z. Kaya, M. Demirci, The evaluation and comparison of daily reference evapotranspiration with ANN and empirical methods, *Nat. Eng. Sci.*, **3** (2018), 54–64.
19. H. Tao, L. Diop, A. Bodian, K. Djaman, P. M. Ndiaye, Z. M. Yaseen, Reference evapotranspiration prediction using hybridized fuzzy model with firefly algorithm: Regional case study in Burkina Faso, *Agr. Water Manag*., **208** (2018), 140–151. https://doi.org/10.1016/J.AGWAT.2018.06.018
20. G. Huang, L. Wu, X. Ma, W. Zhang, J. Fan, X. Yu, et al., Evaluation of CatBoost method for prediction of reference evapotranspiration in humid regions, *J. Hydrol.*, **574** (2019), 1029–1041. https://doi.org/10.1016/J.JHYDROL.2019.04.085
21. M. Kadkhodazadeh, M. V. Anaraki, A. Morshed-Bozorgdel, S. Farzin, A new methodology for reference evapotranspiration prediction and uncertainty analysis under climate change conditions based on machine learning, multi criteria decision making and Monte Carlo methods, *Sustainability*, **14** (2022), 2601. https://doi.org/10.3390/SU14052601
22. M. Chia, Y. Huang, C. Koo, J. Ng, A. Ahmed, A. El-Shafie, Long-term forecasting of monthly mean reference evapotranspiration using deep neural network: A comparison of training strategies and approaches, *Appl. Soft Comput.*, **126** (2022), 109221. https://doi.org/10.1016/j.asoc.2022.109221
23. Y. Z. Kaya, M. Mamak, F. Üneş, Evapotranspiration prediction using M5T data mining method, *Ijaers J.*, **3** (2016), 2456–1908. https://doi.org/10.22161/ijaers/3.12.40
24. D. Yildirim, B. Cemek, E. Küçüktopcu, Estimation of daily evaporation using fuzzy artificial neural network (ANFIS) and multilayer artificial neural network system (ANN), *Toprak Su J.*, (2019), 24–31, (in Turkish). https://doi.org/10.21657/TOPRAKSU.654778
25. A. Ozel, M. Buyukyildiz, Usability of artificial intelligence methods for estimation of monthly evaporation, *Omer Halisdemir University J. Eng. Sci.*, **8** (2019), 244–254. https://doi.org/10.28948/NGUMUH.516891
26. Z. F. Toprak, Flow discharge modeling in open canals using a new fuzzy modeling technique (SMRGT), *CLEAN–Soil Air Water*, **37** (2009), 742–752. https://doi.org/10.1002/CLEN.200900146
27. E. Altaş, M. C. Aydın, Z. F. Toprak, Modeling Water surface profile in open channel flows using fuzzy SMRGT method, *Dicle University J. Eng.*, **9** (2018), 975–981.
28. USGS.gov | Science for a changing world, (n.d.). https://www.usgs.gov/.
29. H. L. Penman, Natural evaporation from open water, hare soil and grass, *Proc. R Soc. Lond. A Math. Phys. Sci.*, **193** (1948), 120–145. https://doi.org/10.1098/RSPA.1948.0037
30. M. Jensen, R. Burman, R. Allen, Evapotranspiration and irrigation water requirements, *ASCE*, New York, (1990). https://cedb.asce.org/CEDBsearch/record.jsp?dockey=0067841
31. F. Cansiz, F. Üneş, I. Erginer, B. Taşar, Modeling of highways energy consumption with artificial intelligence and regression methods, *Int. J. Environ. Sci. Technol.*, **19** (2022), 9741–9756. https://doi.org/10.1007/S13762-021-03813-1
32. Ö. F. Cansiz, İ. Erginer, E. Doğru, Estimation number of traffic accidents and number of injured by artificial neural networks and regression methods, *Osmaniye Korkut Ata University J. Inst. Sci. Technol.*, **3** (2020), 29–35. https://doi.org/10.47495/OKUFBED.844250
33. C. Riviere, P. Lauret, J. F. M. Ramsamy, Y. Page, A Bayesian Neural Network approach to estimating the Energy Equivalent Speed, *Accid Anal. Prev.*, **38** (2006), 248–259. https://doi.org/10.1016/J.AAP.2005.08.008

34. N. Walia, H. Singh, A. Sharma, ANFIS: Adaptive neuro-fuzzy inference system-a survey, *Int. J. Comput. Appl.*, **123** (2015), 32–38.
35. F. Üneş, M. Demirci, M. Zelenakova, M. Çalişici, B. Taşar, F. Vranay, et al., river flow estimation using artificial intelligence and fuzzy techniques, *Water*, **12** (2020), 2427. https://doi.org/10.3390/W12092427
36. J. S. R. Jang, Fuzzy modeling using generalized neural networks and Kalman filter algorithm, In *AAAI-91 Proceedings*, (1991), 762–767.
37. J. S. R. Jang, ANFIS: Adaptive-network-based fuzzy inference system, *IEEE Transact. Syst. Man. Cybernet.*, **23** (1993), 665–685.
38. Z. F. Toprak, A. Toprak, Z. Aykac, Practical applications of Fuzzy SMRGT method, *Dicle University J. Eng.*, **8** (2017), 123–132.
39. S. K. Shevade, S. S. Keerthi, C. Bhattacharyya, K. R. K. Murthy, Improvements to the SMO algorithm for SVM regression, *IEEE Trans. Neural Netw.*, **11** (2000), 1188–1193. https://doi.org/10.1109/72.870050
40. A. J. Smola, B. Schoelkopf, A tutorial on support vector regression, 1998.

AIMS Press