International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Mathematical Modelling of Engineering Problems
Vol. 9, No. 5, October, 2022, pp. 1261-1270
Journal homepage: http://iieta.org/journals/mmep

Prediction of Monthly Evaporation Model Using Artificial Intelligent Techniques in the Western Desert of Iraq-Al-Ghadaf Valley

Ahmed Saud Mohammed, Atheer Saleem Almawla*, Salah Sabbar Thameel

Upper Euphrates Basin Developing Center, University of Anbar, Ramadi 00964, Iraq

Corresponding Author Email: eng.atheer84@uoanbar.edu.iq

https://doi.org/10.18280/mmep.090513

Received: 1 June 2022
Accepted: 12 October 2022

Keywords:
artificial intelligent, evaporation, MLP, RBF, SVM model

ABSTRACT

The use of traditional methods to predict evaporation may face many obstacles due to the influence of many factors on the pattern of evaporation's shape. Therefore, the use of existing methods of artificial intelligence is a reliable prediction model in many applications in engineering. Monthly measurements were employed in the present work to predict for duration eighteen years, from beginning of January 2000 until December 2017. The best model was chosen using ANNs (MLP, RBF) and AI (SVM) techniques. The best evaporation model prediction was made using ANNs (MLP, RBF) and AI (SVM) technologies, with temperature, wind speed, relative humidity, and sunshine hours used as independent variables. Several statistical metrics have been used to evaluate the effectiveness of the proposed model to other popular artificial intelligence models. The obtained result denotes the superiority of the MLP models over the RBF and SVM models. It is concluded that the MLP model is better than RBF and SVM for evaporation prediction for both groups. A comparison of the model performance between MLP, RBF, and SVM models indicated that the MLP-ANN method presents the best estimates of monthly evaporation rate with minimum RMSE 0.033, minimum MAE 0.026, and maximum determination coefficient 0.967.

1. INTRODUCTION

Various meteorological factors can generate evaporation, therefore precise pan evaporation modelling can be regarded a significant step, especially in the locations found in the desert where rainfall has a low rate. The accuracy of evaporation prediction is important and critical for various applications of water resources engineering. This importance appears in countries such as Iraq, where meteorological stations are still not well maintained and operated in such situ estimations [1]. Evaporation has an extremely high rate in arid and semi-arid environments that are found in Iraq [2]. The interrelationships that clarify the evaporation with the factors that affect its occurrence are non-linear and in most cases are not clear [3]. In addition, dry areas contain a low rate of data that is sometimes inaccurate for the researchers to estimate evaporation, this motivated and prompted them to look for other ways or means to be able to deal with the problem that is reflected or manifested by artificial intelligence techniques [4]. Evaporation is a phenomenon that is various and it is manifested by the effects of elements of the atmosphere which are various in most times, such as temperature, rainfall, humidity, wind speed solar and radiation [5]. Low humidity, high wind speed and max temperatures lead to an increase in the rates of evaporation [1, 6, 7]. These factors are the main cause that makes evaporation estimation complex or it can be regarded as the most complex calculation in hydrology and meteorology sciences. How measure evaporation is a complicated process. Class pan evaporation that is used in meteorological stations is exploited to measure the evaporation. It has its limitations and it is an index for the evaporation of a region. Sammen [8] stated that traditional methodologies are different to a great extent in their ability to talk about or define the magnitude and various forms of evaporation from the reservoirs in an arid area. Therefore, there is a persistent need to develop or find out alternative approaches for estimating the evaporation rates depending on metrology variables because of their easiness in measuring and estimating. The use of soft computing modelling techniques which is also called Artificial Intelligence (AI) is regarded one of the recent alternate approaches. They have better modelling flexibility and capability rather compared with the previous empirical approaches wherein this modelling of each of the metrological parameters has its share proportionally. The power of AI to handle and process complicated problems generates from the ability to imitate the behavior of the human brain [9-11]. Artificial neural networks (ANN) seem very promising for regression and classification, especially for large covariate spaces. The artificial neural network is applied in many of study and different sciences. The behavior of ANN is the same manner as biological neural networks. It is a simulation of the input and output data. Estimation of evaporation losses is one of these applied [5].

The use of this form of evaporation estimation has its major problem that is reflected by the dynamic nature of the used meteorological variables because of their non-stationary, nonlinearity and stochastic features. These models are efficiently able to present a mimication and solution to the stochasticity of the processes of complex hydro-climatological. The studies of evaporation prediction that have done recently have clarified a noticeable achievement toward the best, more reliable, predictive and generalized models [1]. It is thus necessary to make development or improvement reliable and robust intelligent predictive models of this process. The major concentrate in engineering and water resources management is reflected in the development of such models [12]. The artificial neural network is applied in many of study and different sciences. The behavior of ANN is the same manner as biological neural networks. It is a simulation of the input and output data. Estimation of evaporation losses is one of these applied [5].

Moghaddamnia et al. [13] and Keskin et al. [14] they appointed air temperature, solar radiation, relative humidity and wind speed as input data for ANN models, they proved the ANN can perform reliable estimation of evaporation. Semman used ANN method to estimate the evaporation of Hemren reservoir in Iraq depending on the use of daily temperature, wind velocity, relative humidity, sunshine hours, and data of evaporation. The Lev. Marqn. Back Prog. (LMBP) has been utilized to form or construct the ANN models. The results affirm that ANN can be regarded as the best model for the estimation (4-10-1), it has a correlation coefficient (R^2) was 0.999 and MSE was 0.112. Pallavi and Rajeev [15] studied the Predicting Reservoir Evaporation using Artificial Neural Network using four ANN models to predict the evaporation losses from NathSagar Reservoir of Maharashtra with change input and number of hidden neurons. THE study shows that the second model (ANN-2 (4-9-1)) was the best model for evaporation estimation using four inputs (Max temp, Relative Humidity, Sunshine Hours and Min. Temp). The correlation coefficient (R^2) for this model was 0.974 and the root means square value (RMSE) was 1.276.

Yaseen et al. [1] applied the (SVM) in the two of famous meteorological stations (Mosul and Baghdad) depending on the use of various combinations of meteorological variables that are manifested in wind speed, sunshine hours, rainfall, humidity and the low and high results of temperatures. The results of SVM are good one with the speed of wind, the amounts of rainfall, and humidity as variable data at Mosul Station ($R^2 = 0.92$), and with the expected variables as data used at Baghdad Station ($R^2 = 0.97$). Zounemat-Kermani et al. [16] estimated the evaporation the climate of dry area that can be hot and dry by the use of (ANN) and (ANFIS) for daily prediction of pan evaporation. The parameters of daily climate that include the average air temperature (T avg), relative humidity (RH), sunshine hours (S), pan evaporation (E) and wind speed (W) are the variable factors. The results indicate that accuracy can generally be got depending on ANN model. This make ANN model can be the suitable model in the prediction of daily evaporation.

Allawi et al. [17] presented their own modern model that is named as Coactive Neuro-Fuzzy Inference System (CANFIS). The efficiency of this new model compared with the other models of artificial intelligence depending on statistical indicators. The study ended with the fact that the modified GA-CANFIS model has a higher rank compared with GA-ANFIS, GA-SVR, and GA-RBFNN for prediction of the evaporation. GA-CANFIS could get minimum RMSE that equals (15.22 mm in the month−1 for AHD, 8.78 mm in the month−1 for TTD), minimum MAE (12.48 mm in the month−1 for AHD, 5.11 mm monthly −1 for TTD), and maximum of determination coefficient R^2 were 0.98 and 0.95 for AHD, TTD) respectively. Guan et al. [18] resorted to the use of the optimization algorithms which is also called shark algorithm (SA) and it is also called frefy algorithms (FFAs), to presents the train of "the adaptive neuro-fuzzy interface system (ANFIS), multilayer perceptron (MLP) model and radial basis function (RBF) model" for the process of the prediction of evaporation during a period of one month. The results affirmed that the new ANFIS models integrated with shark algorithms could form a powerful tool for the prediction of evaporation. The use of Emadi et al. [19] to adaptive neuro-fuzzy inference system (ANFIS), artificial neural networks (ANNs), wavelet-hybrids (WANN, WANFIS, and WGEP) as well as gene expression programming (GEP) to estimate of monthly Epan in Iran. The results indicated that WGEP and ANN are suitable methods for an estimate of monthly Epan. The R^2 ranged between 0.973 to 0.911 and RMSE ranged from 38.4 to 15.8.

The current study presents a mathematical equation for the estimation of monthly Epan that has a significant impact on the process by which water resources policymakers can plan and manage their projects in the future. The results of the current study's use of artificial intelligence (AI) produced evaporation model have demonstrated the model's effectiveness, capacity to forecast evaporation in arid regions, and superiority over the most significant models used for estimated evaporation. Estimation of evaporation for the ungauged basin in an arid area environment is a major challenge. The main problem is there are no precise equations to estimate evaporation in the arid area due to a lack of data in these regions, especially the study area, therefore this study aims to investigate the potential and usefulness of ANN depending on modelling for evaporation prediction for a period of a month by using MLP, RBF and SVM models at Al-Ghadaf valley whose location is in Anbar governorate in Iraq by the use of monthly temperature, relative humidity, wind speed, sunshine hours as well as evaporation data in several meteorological stations available in the area of the study. In addition, the weight connections of the ANN models were used to get an evaluation of the effects of the input factors on the variable output in the current study.

2. STUDY AREA AND DATA COLLECTED

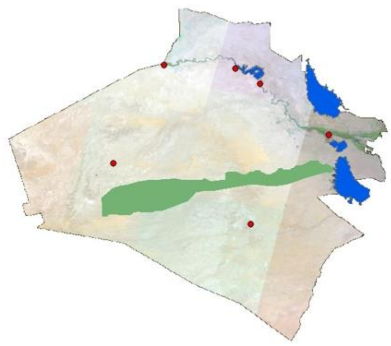

Figure 1. Locations of Al-Ghadaf valley and the climatic stations

The studied area lies within Anbar Province southwestern Ramadi city and is bounded by long. (41◦ 35′ - 43◦ 30′) E. and lat. (32◦ 40′ - 33◦ 10′). The area reaches 2558 Km^2 of the southern part of Ghadaf valley. Al-Ghadf watershed was selected based on previous studies that indicated that the Al-Ghadaf basin is the best area for water harvesting projects in Anbar Governorate from the number of times that surface runoff occurs relative to the area of the water basin. The water movement is generally from west and southwest toward East and Northeast. The climate is of desert type where it is characterized by low rainfall and hot dry summer. For the study of climatic conditions, monthly climate data for six weather stations was provided by the "Iraqi Meteorological Organization Seismology (IMOS)" in Baghdad (see Figure 1). The weather stations are Anah, Qaim, Rutba, Ramadi, Haditha, and Nakheb. The data were evaporation(E), temperature (T), relative humidity (RH), sunshine hours (SS) and wind speed (WS) for the period 2000-2017.

3. ARTIFICIAL INTELLIGENCE (AI) TECHNIQUES

The power of AI to handle and process the problems results from the power for the mimication of the behavior of the human brain behavior [9-11]. AI was exploited successfully and extensively in many complex hydrological studies that are characterised by the reliable performance that has a less cost with a relative rate, the required effort and data. It has a superior rank compared with the process-based models. AI techniques have a performance of many advantages compared with other techniques when the phenomenon in question is complex and there is a shortage in obtainable data, which reflects the situation in most hydrological variables in the dry regions. In many cases, AI models outperformed the process-based models [4, 20]. To this moment, several artificial intelligence techniques that have many common points with the artificial neural networks ANNs such as support vector machines (SVM) and (MLP and RBF) that serve for determining the relationship with rainfall-runoff. Their efficiency has been proven as a means of new modelling compared with the traditional models that were used in this area [11].

3.1 The support vector machines model

The support vector machines (VSM) is regarded as a new data-driven way that applies the theory of mathematics learning whose use is manifested in unravelling regression and classification issues [21]. The SVM that is classified as a new generation of learning machine that depends on use of a hyperplane for the process of separating the data that have one dimension to those that have higher dimensional space. The next step is to present solutions for the regression problems by restoration to the following equation [1, 22].

$$y = f(x) = \sum_{i=0}^{n} w * k\,(xi\,, x) + b \qquad (1)$$

where, w is the weight vector, b is the bias value and $k\,(x_i, x)$ is the kernel function. The internal parameters with their values are determined by use of the least-squares method by minimizing the sum of the squared deviations.

SVM is built on the statistical learning theory (SLT). It is regarded as a new theory in the machine learning technique that can be characterized by its modernity and its dependence on small data samples to estimate and predict nowadays [23, 24]. SVM has been applied widely in scientific fields and it exploited little water resources and hydrology thus, it is distinguished by its good performance that is noticed in most the applications of prediction and rainfall-runoff modelling [25]. The training for regression or classification of SVM may involve the decreasing the error function [25]:

3.2 Artificial Neural Network (ANN)

(ANNs) mechanisms seek the activation and simulation of the human brain's behaves for developing a form or a pattern of the Artificial Intelligence System. For the past few years, this model has been used in many applications that include the prediction attached with the problems surrounding water resources. Furthermore, and in many cases, the AI models overcame the process-based models [20, 26].

A neural network includes three layers or more. These layers are input one and output one. The third one consists of several hidden layers respectively and as shown in Figure 2. The neuron of one layer must be connected or attached to the neurons in the next layer, but the units of the same layer are left without connections. The kind of problem determines the number of neurons in each layer. Figure 2 shows the basic component of the node of ANNs.

The interrelationships that are found in the process of evaporation and the factors that affect its formation are, in most cases, not clear and non-linear [3]. In addition, the data of dry areas are characterised by being low and sometimes inaccurate in the process of the formation of evaporation. This obstacle motivated the researchers to find other ways to deal with or tackle this kind of problem including artificial intelligence techniques [4].

The specialists prefer the adoption of the ANN models because of their power to deal with or study the linear and nonlinear systems [27]. Furthermore, several artificial intelligence techniques which are reflected in artificial neural networks (ANNs) and support vector machine (SVM), with their efficiency for determining the relationships within hydrological processes have proven as a new tool or means of new modelling that replace the traditional models used within the same area [11]. In ANNs, the availability of the data seeks to develop empirical relationships such as those of (cause-effect or input-output). They can communicate the physical action. They are utilized to estimate the output using fresh input data [4, 28].

In the present study, the MLP and RBF which are regarded as intelligent techniques and they represent the ANN technique were used.

3.3 Multilayer Perceptron (MLP)

The current understanding of the biological nervous system forms the base on which Multilayer Perceptrons (MLP) are built. In spite of the fact that much of the biological information and detail is neglected. MLP is a parallel system that composes of many processing elements. These elements are connected by links of variable weights. The backpropagation network is the most popular of the many MLP paradigms [29]. Layers of parallel processing elements form the network. These elements are called neurons and they are connected with the layer that precedes them by interconnection strengths, or weights, (W). Figure 2 clarifies a three-layer neural network with its layers i, j, and k, with the interconnection weights Wij and Wjk between layers of neurons. The estimation of the initial weight values is corrected during the process of training that seeks to compare predicted outputs with known outputs and backpropagate any errors (from right to left in Figure 2) to determine the appropriate weight adjustments whose appearance is necessary to minimize the errors.

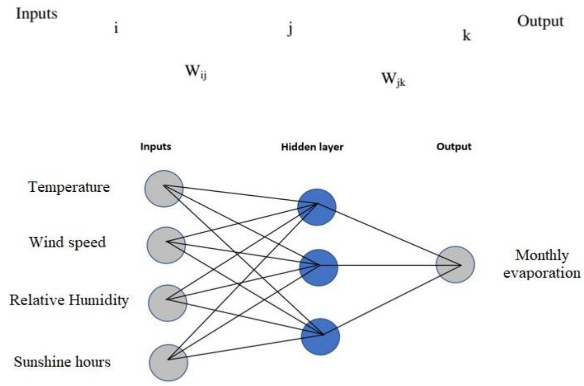

Figure 2. Architecture for Multilayer Perceptron (MLP) method

The multi-layer perceptron (MLP) neural network's design includes three layers: the input is reflected by the first layer, the second implies the hidden and the last one is devoted to the output. In MLP networks, the weights should be processed with input data before their entrance into the hidden layer (see Figure 2). In the MLP network, the algorithm is the component that is responsible for taking the inputs that are distributed through the layers of hidden features when the process of the modification of weights is in progress so that the input combined with weights is going to enter into the hidden layers for getting one output.

3.4 Radial Base Function (RBF)

Photoenhanced et al. [30] introduced RBF into the neural network literature. It consists of two layers. The output nodes of these two layers appear in a linear combination of the basis functions. The basis functions found in the hidden layer are responsible for the production of a nonzero response to the input stimulus. When the input locates in a constrained area of the input space, this process takes place.This paradigm can also be defined as a localized receptive field network [31]. Figure 2 clarifies the relation that attach inputs and outputs. A transformation that happened with inputs is essential to fight the curse of dimensionality happened in empirical modelling. The type of input transformation belongs to RBF can be regarded the projection of local nonlinear that uses a basis function of radial fixed shapes. Nonlinearly squashing is followed by the multi-dimensional inputs without considering the output space, a role as regressor is played by the radial basis functions. The implementation of the output layer causes a linear regressor in which the only changeable parameters are the weights of this regressor. They are adjustable. Therefore, these parameters can be determined by the use the linear least square method giving an important advantage for convergence. The basic concept and algorithm of the RBNN is described in [32].

A nonlinear function can be achieved by this equation ***h(x, t),*** (x) refers to the variable input and (***t***) is used as an indication to its center. Because of its dependence on the radial distance, it is also called a radial basis function r = ‖ x - t ‖. Given ***N*** different points {xi ϵ R^n | I = 1, 2,…,N} and ***N*** real numbers {yi ϵ R | I= 1, 2,…,N }, a function ***f*** can be found from $\boldsymbol{R^n}$ to ***R to*** satisfy the interpolation conditions: $f(x_i) = y_i$, i = 1, 2,…, N. The RBF approach consists of a linear space of dimension ***N***, depending on the data points *{xi}*.

The basis of the space in question is chosen to be the set of functions

$$\{h(\|x - xi\|) \mid i = 1,2,\ldots,N\}$$

Three layers were exploited in RBF method. The first layer represents the input layer. The second layer reflects a hidden layer including hidden units (neurons). They differ from one network to another. The output layer is manifested by the third layer that estimates the value of a dependent variable (see Figure 3). The training and testing phases make up the two phases of these networks. There is a difference in the function of RBF from that of the MLP. It reverses the MLP in the process of specifying the input variables and then it is divided into hidden layers resorting to group form (G). After that, the smart process is finished, and only then may the outputs (one output) be retrieved.

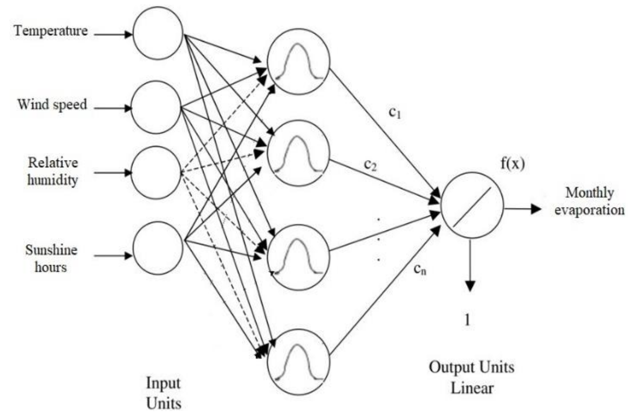

Figure 3. Architecture for RBF model

4. RESULTS and DISCUSSIONS

4.1 Prediction performance indicators

Sahoo et al. [28] reported that when comparing the performance efficiency of many models with each other, performance indicators that are used will determine which model best performance, because each performance indicator may give different results between models. Therefore, the performance indicator for each model should be evaluated according to all performance indicators [33]. To examine the suitable AI and ANNs models, many performance indicators should be used which are calculated by the monitored data and predicted data. The most common performance indicators used in the model evaluation are the coefficient of determination (R^2), root mean square error (RMSE), normalized absolute error (NAE), "Nash Sutcliffe Efficiency" abbreviated as (NSE) and "mean average percentage error" with this abbreviation (MAPE) [34]. NAE and RMSE should have zero for a successful pattern. In contrast, NSE and R^2 should be closer to one [35].

The following equations are used to calculate the performance indicators:

$$R^2 = \left(\frac{\sum_{i-1}^{N}(P_i - \bar{P})\,(O_i - \bar{O})}{N.S_{pred}.S_{obs}}\right)^2 \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N-1}\sum_{i-1}^{N}(P_i - O_i)^2} \quad (3)$$

$$NAE = \frac{\sum_{i=1}^{n} | Pi - Oi |}{\sum_{i=1}^{n} Oi} \quad (4)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{n} \left| \frac{Oi - Pi}{Oi} \right| * 100 \quad (5)$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{N}(Oi - P_i)^2}{\sum_{i=1}^{N}(Oi - \bar{O})^2}\right] \quad (6)$$

4.2 Model development

The application of ANNs (MLP and RBF) and SVM techniques for modelling evaporation with their results are presented in this section. The choice of the appropriate predictors can be regarded as one of the significant steps in designing a robust predictive model [2]. In the present study, the selected groups of evaporation and the best influencing and co-predicting variables were adopted in this study based on [35] in order to apply artificial intelligence techniques to these groups. Table 1 clarifies that Group one that includes E as the dependent variable and T, WS, RH and SS as independent variables and Group two consists of E as the dependent variable and T, WS and SS as independent ones.

Table 1. Best selected groups for AI evaporation prediction

Model no	Input parameter	R^2	Adjusted R^2	SE	Equation
Group one	T, WS, RH, SS	0.946	0.946	37.30	= - 168.05 + 8.47*T + 26.13*WS -0.535*RH + 20.43*SS
Group two	T, WS, SS	0.945	0.945	37.48	= - 220.87 + 9.014*T + 25.612*WS + 22.66*SS

4.3 ANN methods

4.3.1 Multilayer Perceptron (MLP)

Table 2 illustrates the results in each group MLP method that were given by the best architectural network. It clarifies that the network building in Group one is (4 - 7- 1), Group two (3 - 3 - 1). The performance indicators with the abbreviation (PI) were estimated for evaporation MLP models in each group (see Table 3). The results of MLP Models indicate that the highest R^2 in the training phase in Group one is 0.962 and the highest R^2 in the test phase in Group one is 0.989, the lowest RMSE concerning the training phase in Group one is 0.048 and the lowest RMSE in the test phase is 0.033 Groups one.

The selection of three models (MLP, RBF and SVM) which are different in their nature for predicting monthly evaporation was done. The hydrometeorological data were distributed into two phases. The first consists of 80% training while the latter consists of 20% testing. The results data were obtained from (MLP, RBF and SVM) depending on the principle of a trial and error by continuing to train until getting the best R^2 value in the test phase and a small error that indicates the best prediction performance was attained. AI modelling MLP, RBF and SVM) were performed to design an evaporation predictive equation by the use of climatologic variables, such as T, WS, SS and RH. Table 3 shows the results obtained for two groups.

- **Group one**

The Table 3 shows that the best R^2 values calculated in train and test phase were 0.96 and 0.967 respectively. Figures 4 and 5 illustrates the relationship between observed and expected evaporation values using MLP method in train and test phase. The relationship suggesting that there is a strong correlation between observed and predicted values.

Table 2. The architecture of evaporation - MLP method models

Model	Network architecture	Inputs	Hidden units	Output
Group one	MLP 4-7-1	4	7	1
Group two	MLP 3-3-1	3	3	1

Table 3. Evaporation performance indicators of MLP models groups

Model	Phase	r	R^2	RMSE	NAE	MAE	NSE	MAPE
Group one	Training	0.980	0.96	0.048	0.064	0.0263	0.960	7.03
	Test	0.983	0.967	0.033	0.069	0.0266	0.981	7.70
Group two	Training	0.976	0.953	0.056	0.069	0.0284	0.953	7.39
	Test	0.979	0.96	0.04	0.082	0.03	0.978	8.45

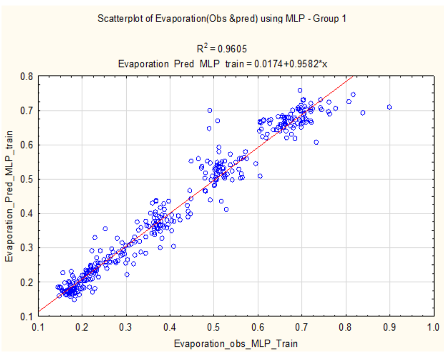

Figure 4. Observed versus predicted evaporation for MLP model group one for train phase

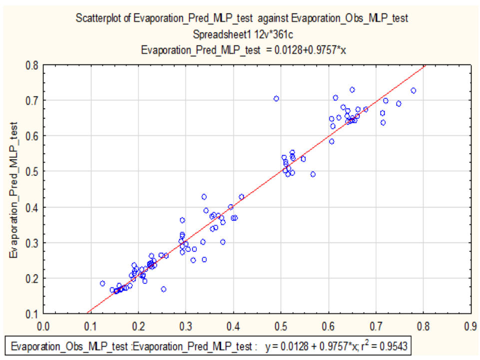

Figure 5. Observed versus predicted evaporation for MLP model group one for test phase

- **Group two**

The Table 3 shows that the best R^2 values calculated in train and test phase were 0.953 and 0.960 respectively. Figures 6 and 7 illustrate the relationship between observed and expected evaporation values using MLP method in train and test phase. The relationship suggesting that there is a strong correlation between observed and predicted values.

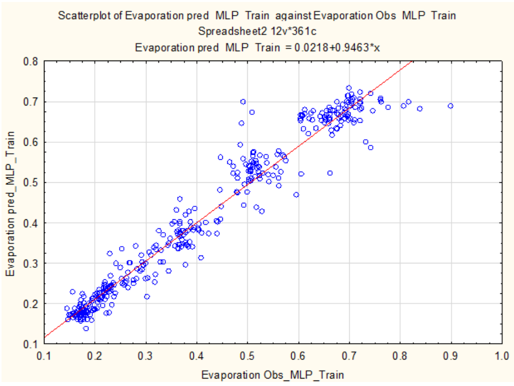

Figure 6. Observed versus predicted evaporation for MLP model group two for train phase

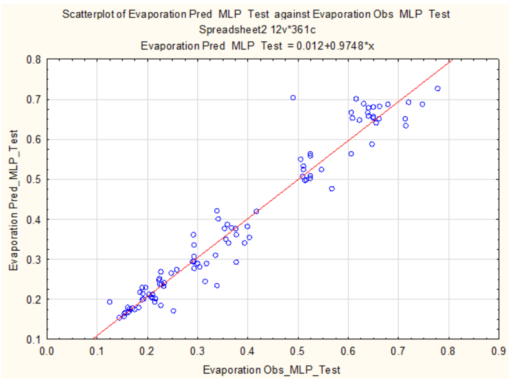

Figure 7. Observed versus predicted evaporation for MLP model group two for test phase

4.3.2 Radial Basis Functions (RBF) method

- **Group one**

Table 4 illustrates the results of each group in the RBF method that was given by the best architectural network. It shows that the network architecture in Group one is (4 - 3- 1), Group two (3 - 3 - 1). The PI was estimated for evaporation RBF models in each group as shown in Table 5. The results of RBF Models refers to that the regression (R^2) in the training phase for Group one is 0.949 and the highest regression (R^2) in the test phase of group one is 0.958, the RMSE in the training phase in group one is 0.05 and the RMSE in the test phase is 0.037.

Figures 8 and 9 illustrate the relationship between observed and expected evaporation values using RBF method in train and test phase. The relationship suggesting that there is a strong correlation between observed and predicted values.

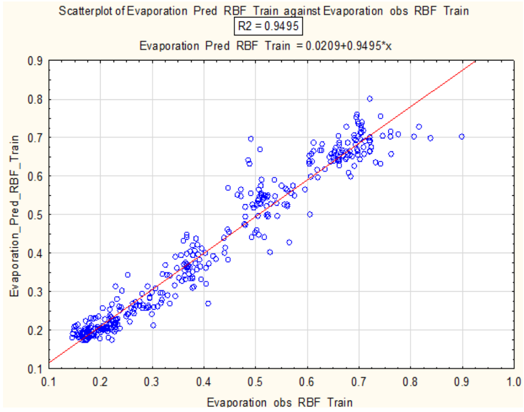

Figure 8. Observed versus predicted evaporation for RBF model group one for train phase

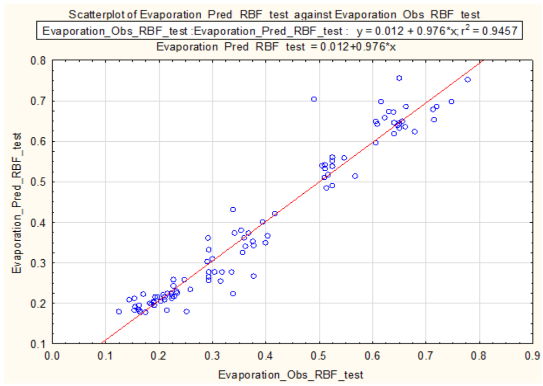

Figure 9. Observed versus predicted evaporation for RBF model group one for test phase

- **Group two**

The results of RBF Models refer to that the regression (R^2) in the training phase for Group two was 0.952 and the (R^2) in the test phase was 0.957, the RMSE in the train and test phase were 0.057 and 0.04 respectively.

Table 4. The architecture of evaporation - RBF method models

Model	Network architecture	Inputs	Hidden units	Output
Group one	RBF 4-3-1	4	3	1
Group two	RBF 3-3-1	3	3	1

Table 5. The evaporation performance evaluation indicators of RBF models groups

Model	Phase	r	R^2	RMSE	NAE	MAE	NSE	MAPE
Group one	Training	0.974	0.949	0.05	0.076	0.0314	0.949	8.47
	Test	0.978	0.958	0.037	0.073	0.0312	0.978	9.61
Group two	Training	0.975	0.952	0.057	0.091	0.0286	0.952	7.36
	Test	0.978	0.957	0.04	0.083	0.0314	0.978	8.45

Table 6. The evaporation performance evaluation indicators of SVM models groups

Model	Phase	r	R^2	RMSE	NAE	MAE	NSE	MAPE
Group one	Training	0.979	0.959	0.051	0.09	0.028	0.959	8.046
	Test	0.981	0.962	0.035	0.06	0.0294	0.976	8.47
Group two	Training	0.977	0.955	0.059	0.106	0.0305	0.954	8.74
	Test	0.979	0.958	0.042	0.069	0.0309	0.974	9.08

Figures 10 and 11 illustrate the relationship between observed and expected evaporation values using RBF method in train and test phase. The relationship suggesting that there is a strong correlation between observed and predicted values.

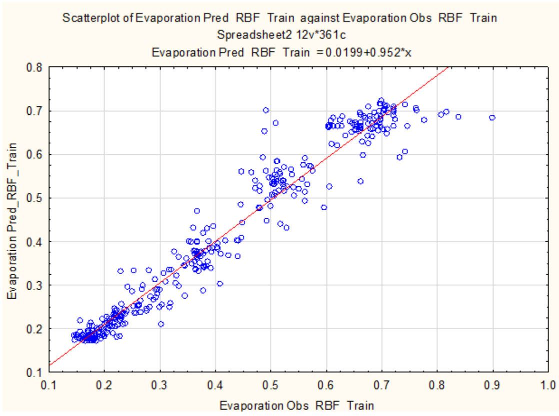

Figure 10. Observed versus predicted evaporation for RBF model group two for train phase

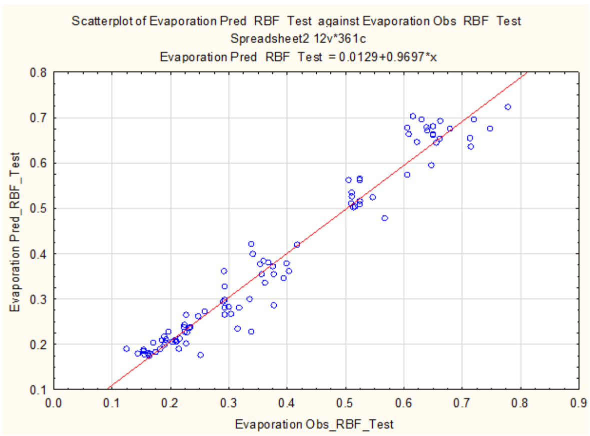

Figure 11. Observed versus predicted evaporation for RBF model group one for test phase

4.4 Support vector machine for modelling

PI was carried out to find out the best-fit evaporation models in 1 group using the SVM method for testing and training phases as shown in Table 6.

- **Group one**

The results of SVM Models refers to that the regression (R^2) in the training phase for group one was 0.959 and the (R^2) in the test phase was 0.962, the RMSE in the train and test phase were 0.051 and 0.035 respectively.

Figures 12 and 13 illustrate the relationship between observed and expected evaporation values using SVM method in train and test phase. The relationship suggesting that there is a strong correlation between observed and predicted values.

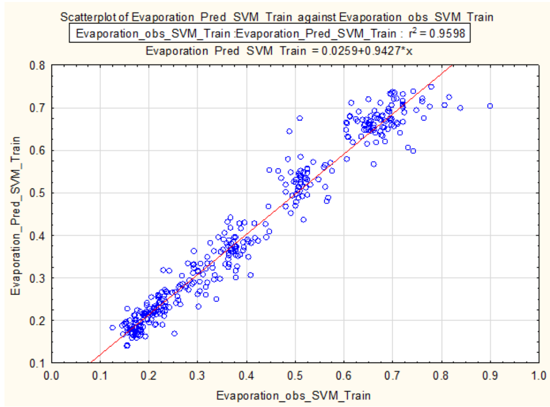

Figure 12. Observed versus predicted evaporation for SVM model group one for train phase

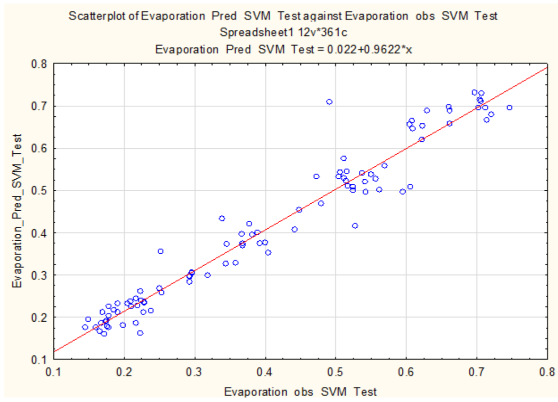

Figure 13. Observed versus predicted evaporation for SVM model group one for test phase

- **Group two**

The results of SVM Models refers to that the regression (R^2) in the training phase for group two was 0.955 and the (R^2) in the test phase was 0.958, the RMSE in the train and test phase were 0.059 and 0.042 respectively.

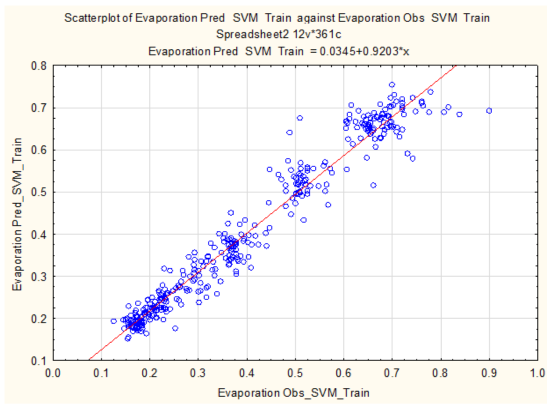

Figure 14. Observed versus predicted evaporation for SVM model group two for train phase

Table 7. The optimal model for each AI technique

Model	Phase	r	R^2	RMSE	NAE	MAE	NSE	MAPE
Group one (MLP)	Training	0.980	0.96	0.048	0.077	0.0263	0.960	7.03
	Test	0.983	0.967	0.033	0.068	0.0266	0.981	7.7
Group one (SVM)	Training	0.979	0.959	0.051	0.09	0.028	0.959	8.04
	Test	0.981	0.962	0.035	0.06	0.0294	0.976	8.47
Group one (RBF)	Training	0.974	0.949	0.05	0.076	0.0314	0.949	8.47
	Test	0.978	0.958	0.037	0.073	0.0312	0.978	9.61

Figures 14 and 15 illustrates the relationship between observed and expected evaporation values using SVM method in train and test phase. The relationship suggesting that there is a strong correlation between observed and predicted values.

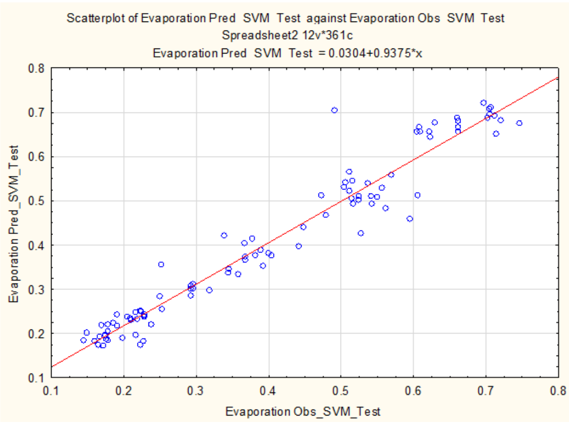

Figure 15. Observed versus predicted evaporation for SVM model group two for test phase

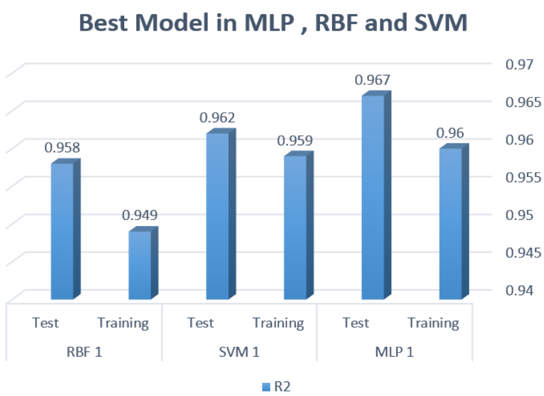

Figure 16. The best model in MLP, RBF and SVM

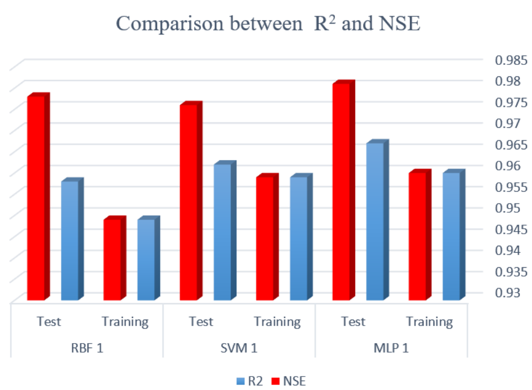

Figure 17. Comparison between R^2 and NSE for three models (MLP, RBF and SVM)

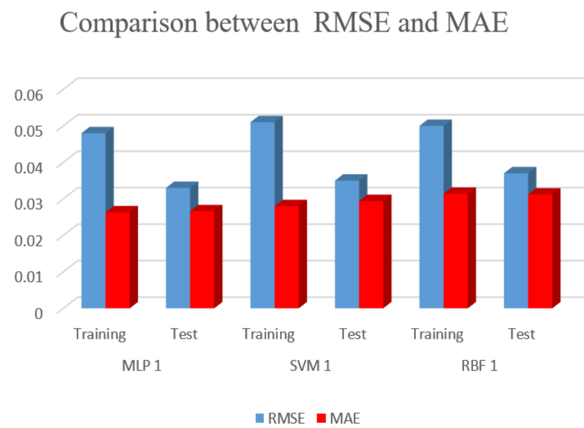

Figure 18. Comparison between R^2 and NSE for three models (MLP, RBF and SVM)

Table 7 and Figure 16 summarize the best group in each method (MLP, RBF and SVM) that used in current study. Figure 17 and Figure 18 shows a comparison between the best groups in terms of R^2 and RMSE, indicating that the MLP method is the best method to predict the evaporation parameter.

4.5 Sensitivity analysis of evaporation model

To determine which input factor of the model has the greatest influence on evaporation the modified Garson algorithm method can be used. The modified Garson algorithm was used to analyse the model's sensitivity and determine the importance of input parameters in a neural network. This method estimates the relative importance of each input variable within a model by dividing the neural network connection output weights as shown in Eq. (7) below [26].

$$Q_{ik} = \frac{\sum_{j=1}^{L} |w_{ij} v_{jk}| / \sum_{r=1}^{N} |w_{rj}|}{\sum_{i=1}^{N} \sum_{j=1}^{L} \left(|w_{ij} v_{jk}| / \sum_{r=1}^{N} |w_{rj}| \right)} \quad (7)$$

where,

- Q_{ik} refers to the percentage that belongs to the influence of the input variable on the output.
- w_{ij} refers to the connection weight that links the input neuron i with the hidden neuron j.
- v_{jk} refers to the connection weight that links the hidden neuron j with the output neuron k.
- $\sum_{r=1}^{N} |w_{rj}|$ refers to the connected sum of weights that links the N input neurons with the hidden neuron j.

Figures 19 show the importance of input factors on the monthly evaporation. the results of sensitivity analysis show that temperature is the most influence on evaporation followed by wind speed, sunshine, and humidity.

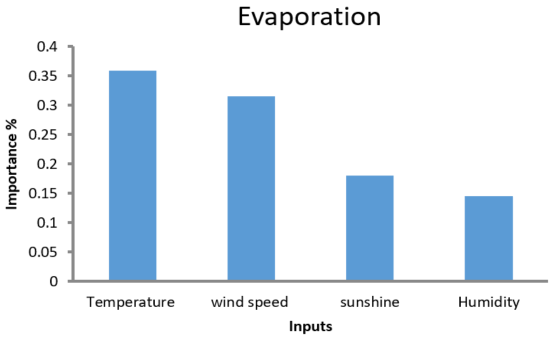

Figure 19. The importance of input factors on the monthly evaporation

5. CONCLUSION

Several AI models such as ANNs (MLP, RBF) and AI (SVM) techniques have been utilized by this study. They were used to select the best monthly evaporation prediction model in the arid region - Al- Ghadaf Valley in the western desert of Iraq. The study depended on the use of a statistical method to determine the sums of the variables most influential in predicting the monthly evaporation rate using multiple linear regression by using Enter method. The best two groups were selected as independent variables, comprising the first group (temperature, relative humidity, wind speed and solar brightness) and the second group (temperature, wind speed, and solar brightness). The variables of the first group gave slightly better results than the second group, and this indicates that the presence of the variable relative humidity does not have a significant impact on the accuracy of evaporation prediction. The performance of the proposed modelling was evaluated by the use of several statistical indicators. The results indicate that. MLP outperformed RBF and SVM methods in enhancing the predictive process and evaporation MLP Group one gets the best results compared to other models. The results of evaporation MLP Group 1 were; in train Phase (R^2= 0.96, RMSE 0.048, NAE = 0.077, MAE=0.026, NSE = 0.96) while in test phase R^2= 0.967, RMSE 0.033, NAE = 0.068, MAE = 0.026, NSE = 0.981). The lower prediction accuracy was obtained by RBF-NN compared to the other methods used. The SVM method gave an accurate prediction between MLP-NN and RBF -NN. On the other hand, for the model of evaporation the most influential input parameter is the temperature, followed by wind speed, sunshine, and humidity. The techniques that were used in this paper could be regarded as a basis for doing a more effective decision-making process on the part of the policymakers for getting the help to improve and maintain the water resources management, particularly in hydrology.

The importance of the curent study comes from the fact that the study area is ungauged basin, and scientific research studies in it are rare, in addition to being an unsave area, which impedes conducting field visits to it. The future direction for the results of the current study is to apply the predictive evaporation equation that derived from this current study in another ungauged basins in Anbar Governorate.

ACKNOWLEDGMENT

Thanks to the Upper Euphrates Basin Developing Center, University of Anbar, Ramadi, Iraq for their support in achieving this work.

REFERENCES

[1] Yaseen, Z.M., Al-Juboori, A.M., Beyaztas, U., et al. (2020). Prediction of evaporation in arid and semi-arid regions: A comparative study using different machine learning models. Engineering applications of computational fluid mechanics, 14(1): 70-89. https://doi.org/10.1080/19942060.2019.1680576

[2] Sayl, K.N., Muhammad, N.S., Yaseen, Z.M., El-shafie, A. (2016). Estimation the physical variables of rainwater harvesting system using integrated GIS-based remote sensing approach. Water Resources Management, 30(9): 3299-3313. https://doi.org/10.1007/s11269-016-1350-6.

[3] Dorofki, M., Elshafie, A., Jaafar, O., Karim, O., Mastura, S. (2012). A GIS-ANN-Based Method for Calculating the Runoff. Energy, 1(1): 1-5.

[4] Alagha, J.S., Said, M.A.M., Mogheir, Y., Seyam, M. (2012). Modelling of chloride concentration in coastal aquifers using artificial neural networks-a case study: Khanyounis governorate gaza strip-palestine. Caspian Journal of Applied Sciences Research, 2: 158-165.

[5] Almawla, A.S. (2017). Predicting the daily evaporation in Ramadi city by using artificial neural network. Anbar Journal of Engineering Sciences, 7(2): 134-139. ISSN: 1997-9428.

[6] Ali Ghorbani, M., Kazempour, R., Chau, K.W., Shamshirband, S., Taherei Ghazvinei, P. (2018). Forecasting pan evaporation with an integrated artificial neural network quantum-behaved particle swarm optimization model: A case study in Talesh, Northern Iran. Engineering Applications of Computational Fluid Mechanics, 12(1): 724-737. https://doi.org/10.1080/19942060.2018.1517052

[7] Lu, X., Ju, Y., Wu, L., Fan, J., Zhang, F., Li, Z. (2018). Daily pan evaporation modeling from local and cross-station data using three tree-based machine learning models. Journal of Hydrology, 566: 668-684. https://doi.org/10.1016/j.jhydrol.2018.09.055

[8] Sammen, S.S. (2013). Forecasting of evaporation from Hemren reservoir by using artificial neural network. Diyala Journal of Engineering Sciences, 6(4): 38-33. ISSN: 1999-8716.

[9] Rajanayaka, C., Samarasinghe, S., Kulasiri, D. (2002). Solving the inverse problem in stochastic groundwater modelling with artificial neural networks. Brigham Young University BYU Scholars Archive, 1st International Congress on Environmental Modelling and Software - Lugano, Switzerland, pp. 154-159.

[10] Iliadis, L.S., Maris, F. (2007). An artificial neural network model for mountainous water-resources management: The case of Cyprus mountainous watersheds. Environmental Modelling & Software, 22(7): 1066-1072. https://doi.org/10.1016/j.envsoft.2006.05.026

[11] Chen, S.H., Jakeman, A.J., Norton, J.P. (2008). Artificial intelligence techniques: An introduction to their use for modelling environmental systems. Mathematics and computers in simulation, 78(2-3): 379-400. https://doi.org/10.1016/j.matcom.2008.01.028

[12] Khan, N., Shahid, S., Ismail, T. B., Wang, X. J. (2019).

Spatial distribution of unidirectional trends in temperature and temperature extremes in Pakistan. Theoretical and Applied Climatology, 136(3): 899-913. https://doi.org/10.1007/s00704-018-2520-7

[13] Moghaddamnia, A., Gousheh, M.G., Piri, J., Amin, S., Han, D. (2009). Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. Advances in Water Resources, 32(1): 88-97. https://doi.org/10.1016/j.advwatres.2008.10.005

[14] Keskin, M.E., Terzi, Ö., Taylan, D. (2004). Fuzzy logic model approaches to daily pan evaporation estimation in western Turkey/Estimation de l'évaporation journalière du bac dans l'Ouest de la Turquie par des modèles à base de logique floue. Hydrological Sciences Journal, 49(6): 1001-1010. https://doi.org/10.1623/hysj.49.6.1001.55718

[15] Pallavi, K., Rajeev, S. (2016). Predicting reservoir evaporation using artificial neural network. International Journal of Innovative Research in Science, Engineering and Technology, (An ISO 3297: 2007 Certified Organization, 5(4): 5851-5857. https://doi.org/10.15680/IJIRSET.2016.0504122

[16] Zounemat-Kermani, M., Kisi, O., Piri, J., Mahdavi-Meymand, A. (2019). Assessment of artificial intelligence-based models and metaheuristic algorithms in modeling evaporation. Journal of Hydrologic Engineering, 24(10): 04019033. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001835

[17] Allawi, M.F., Aidan, I.A., El-Shafie, A. (2021). Enhancing the performance of data-driven models for monthly reservoir evaporation prediction. Environmental Science and Pollution Research, 28(7): 8281-8295. https://doi.org/10.1007/s11356-020-11062-x

[18] Guan, Y., Mohammadi, B., Pham, Q.B., et al. (2020). A novel approach for predicting daily pan evaporation in the coastal regions of Iran using support vector regression coupled with krill herd algorithm model. Theoretical and Applied Climatology, 142(1): 349-367. https://doi.org/10.1007/s00704-020-03283-4

[19] Emadi, A., Zamanzad-Ghavidel, S., Fazeli, S., Zarei, S., Rashid-Niaghi, A. (2021). Multivariate modeling of pan evaporation in monthly temporal resolution using a hybrid evolutionary data-driven method (case study: Urmia Lake and Gavkhouni basins). Environmental Monitoring and Assessment, 193(6): 1-32. https://doi.org/10.1007/s10661-021-09060-8

[20] Banerjee, P., Singh, V.S., Chatttopadhyay, K., Chandra, P.C., Singh, B. (2011). Artificial neural network model as a potential alternative for groundwater salinity forecasting. Journal of Hydrology, 398(3-4): 212-220. https://doi.org/10.1016/j.jhydrol.2010.12.016

[21] Wang, H., Bengio, S. (2002). The mnist database of handwritten upper-case letters (No. REP_WORK). Idiap.

[22] Deka, P.C. (2014). Support vector machine applications in the field of hydrology: A review. Applied Soft Computing, 19: 372-386. https://doi.org/10.1016/j.asoc.2014.02.002

[23] Liu, J.P., Chang, M.Q., Ma, X.Y. (2009). Groundwater quality assessment based on support vector machine. In HAIHE River Basin Research and Planning Approach-Proceedings of 2009 International Symposium of HAIHE Basin Integrated Water and Environment Management, Beijing, China, pp. 173-178.

[24] Kanevski, M., Kanevski, M.F., Maignan, M. (2004). Analysis and Modelling of Spatial Environmental Data. EPFL Press.

[25] Mirzaee, P., Fazloula, R. (2016). Runoff prediction by support vector machine for Chalous River Basin of Iran. International Journal of Geography and Geology, 5(6): 113-118. https://doi.org/10.18488/journal.10/2016.5.6/10.6.113.118

[26] Almawla, A.S., Lateef, A.M., Kamel, A.H. (2022). Modelling the effects of hydraulic force on strain in hydraulic structures using ANN (Haditha dam in Iraq as a case study). Mathematical Modelling of Engineering Problems, 9(1): 150-158. https://doi.org/10.18280/mmep.090119

[27] Rajurkar, M.P., Kothyari, U.C., Chaube, U.C. (2004). Modeling of the daily rainfall-runoff relationship with artificial neural network. Journal of Hydrology, 285(1-4): 96-113. https://doi.org/10.1016/j.jhydrol.2003.08.011

[28] Sahoo, G.B., Ray, C., Mehnert, E., Keefer, D.A. (2006). Application of artificial neural networks to assess pesticide contamination in shallow groundwater. Science of the Total Environment, 367(1): 234-251. https://doi.org/10.1016/j.scitotenv.2005.12.011

[29] Haykin, S., Thomson, D.J. (1998). Signal detection in a nonstationary environment reformulated as an adaptive pattern classification problem. Proceedings of the IEEE, 86(11): 2325-2344. https://doi.org/10.1109/5.726792

[30] Photoenhanced, U., Cadmium, M.O., Haq, S., Sr, R., Malvern, E., Haq, S. (1988). Royal signals and radar establishment. Emerging Technologies for in Situ Processing, 139(4148): 221.

[31] Lee, G.C., Chang, S.H. (2003). Radial basis function networks applied to DNBR calculation in digital core protection systems. Annals of Nuclear Energy, 30(15): 1561-1572. https://doi.org/10.1016/S0306-4549(03)00099-9

[32] Maier, H.R., Jain, A., Dandy, G.C., Sudheer, K.P. (2010). Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. Environmental Modelling & Software, 25(8): 891-909. https://doi.org/10.1016/j.envsoft.2010.02.003

[33] Solaimani, K. (2009). Rainfall-runoff prediction based on artificial neural network (a case study: Jarahi watershed). American-Eurasian J. Agric. & Environ. Sci, 5(6), 856-865. ISSN 1818-6769.

[34] Majhi, B., Naidu, D. (2021). Pan evaporation modeling in different agroclimatic zones using functional link artificial neural network. Information processing in agriculture, 8(1): 134-147. https://doi.org/10.1016/j.inpa.2020.02.007

[35] Mohammed, A.S., Md Said, M.A., Kamel, A.H., Abdullah, R. (2022). Develop evaporation model using multiple linear regression in the Western Desert of Iraq – Horan valley. International Journal of Design & Nature and Ecodynamics, 17(1): 137-143. https://doi.org/10.18280/ijdne.170118