Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Investigating the Performance of Neural Network Based Group Method of Data Handling to Pan's Daily Evaporation Estimation (Case Study: Garmsar City)

H. Karami[1*], H. Ghazvinian[1], M. Dehghanipour[1], M. Ferdosian[2]

1. Faculty of Civil Engineering, Semnan University, Semnan, Iran
2. Faculty of Civil Engineering, K. N. Toosi University of Technology, Tehran, Iran

Corresponding author: *hkarami@semnan.ac.ir*

https://doi.org/10.22115/SCCE.2021.274484.1282

ARTICLE INFO

Article history:
Received: 21 February 2021
Revised: 29 March 2021
Accepted: 19 April 2021

Keywords:
Pan evaporation,
GMDH-NN,
Hydrology,
Sensitivity analysis,
Garmsar.

ABSTRACT

Evaporation is a complex and nonlinear phenomenon due to the interactions of different climatic factors. Therefore, advanced models should be used to estimate evaporation. In the present study, the Neural Network-Based Group Method of Data Handling was used to estimate and simulate the evaporation rate from the pan in the synoptic station of Garmsar city located in Semnan province, Iran. For this purpose, the daily meteorological data of evaporation, minimum and maximum temperature, wind speed, relative humidity, air pressure, and sunny hours of the said station during the nine years (2009-2018) were used. The percent of data on training, test, number of the used layers, and the highest number of neurons were considered as 60%, 40%, 5%, and 30%, respectively. The studied method's accuracy was investigated using the statistical parameter of Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and correlation coefficient, and. Sensitivity analysis of the input parameters was performed using the GMDH-NN model. This study showed that R^2, RMSE, and MAE values in the test phase were obtained as 0.84, 2.65, and 1.91, respectively, in the most optimal state. From the third layer onwards, the amount of the best mean squared errors of the Validation data have converged to 0.062, and it is not affordable to use more layers for the modeling of the evaporation pan in the Garmsar station. The standard deviation and mean amounts of the errors are -0.1210 and 2.552 respectively. The amounts of the best mean squared errors of the validation data are presented. It shows that although the layers are increased, the amounts of the mean squared errors have not changed considerably. (Maximum 0.003). The sensitivity analysis results showed that the two input parameters of minimum temperature and relative humidity percent have a higher effect on evaporation pan modeling than other input parameters.

How to cite this article: Karami H, Ghazvinian H, Dehghanipour M, Ferdosian M. Investigating the performance of neural network based group method of data handling to pan's daily evaporation estimation (case study: Garmsar city). J Soft Comput Civ Eng 2021;5(2):01-18. https://doi.org/10.22115/scce.2021.274484.1282.

1. Introduction

Water is one of the most important factors in developing agriculture and industry in arid and semi-arid regions. Abnormal use of water resources and shortage of this vital matter through different methods have caused severe challenges and water stress [1]. The increasing population and the decline of the natural resources on the planet have forced humans to consider different ways to save these resources [2].

Water is one of the essential human needs, which has led to more requirements for its planning than those in the previous years due to time and space limitations and a very limited volume of fresh and exploitable [2].

One of the most important water resources meeting the agricultural, drinking, and industrial needs in the country, especially in arid and semi-arid regions, is the water stored in lakes behind dams and constructed storage pools[3,4]. In several hot and arid areas, a large volume of water stored behind dams, agricultural pools, and water storage tanks are wasted due to evaporation [5]. Evaporation plays an important role in managing water resources in the region, climate change, and agriculture [6]. Considering global climate change, researchers have conducted several studies on evaporation worldwide and its assessment for identification in the hydrological cycle [7,8]. Evaporation is one of the main phenomena of hydrology [9–11] ,Estimating evaporation plays an essential role in estimating the water balance of basins, designing and managing irrigation systems, and water resources management. One method for estimating evaporation is to use evaporation pans, which are used directly in most parts of the world to measure evaporation from the water's free surface [12].

The parameters that affect the evaporation rate are relative humidity, temperature, wind speed, sunny hours, etc. One of the methods for predicting evaporation is the use of soft computing methods. One of the advantages of this method is saving time, reducing trial and error [13,14]. Intelligent methods in modeling pan evaporation have been studied and approved by several researchers [15–20]. One of the models used in estimating evaporation pan is the developed neural network model of the Neural Network-Based Group Method of Data Handling (GMDH-NN), between input variables (such as temperature, sunny hours, etc.) and output variable (evaporation value). Then, some studies on applying intelligent methods in estimating the rate of pan evaporation are referred to.

J. M. Bruton et al. [21] used the artificial neural network to estimate the daily evaporation from the evaporation pan in different parts of the world, including Rome, between 1992 and 1996. In this method, precipitation, temperature, relative humidity, solar irradiance, and wind speed were used as input data. This study showed that the artificial neural network method has lower error than Priestley-Taylor linear multiple regression method. The coefficient of determination and root of the artificial neural network model's mean square error was 0.71 and 1.1 mm per day, respectively. The R-squared (R2) coefficient and Mean Absolute Error (MAE) of the artificial neural network were estimated to be 0.71 and 1.1 mm per day.

Keskin and Terzi [22] examined the data of a meteorological station near a lake in the west of Turkey to determine the daily evaporation of pan using the artificial neural network model. They compared the results of the above designed network model with the results of the Penman method. The artificial neural network model results showed a higher correlation with pan evaporation rate measured with the Penman method. The best structure of the artificial neural network model with 4 input data, including temperature, solar irradiance, air pressure, water surface temperature, wind speed, and relative humidity, have a low correlation with evaporation intensity in the study area. Qasem et al. [18] predicted the evaporation rate of Tabriz in Iran and Antalya in Turkey with three SVR models, ANN, and a combination of them with WSVR and WANN wavelet conversion. For both stations, the ANN model has had more reasonable results than other presented models had.

Ashrafzadeh et al. [23] compared the prediction of evapotranspiration in the north of Iran with the SARIMA time series model of the intelligent model of support vector machine and Based Group Method of Data Handling. The results showed that all three models have good efficiency in estimating evapotranspiration. Patle et al. [19] compared MLR and ANN models in estimating monthly evaporation of pan in two northern regions of India. The results of this study showed that the ANN model had better performance than the MLR model. Alsumaiei [24] modeled daily evaporation rate with artificial neural networks in Kuwait. The studied station was Kuwait International Airport (KIA). The Meteorological input data of the network include mean temperature, wind speed, and relative humidity, which were presented as 4 scenarios.The results of this study showed that the combined scenario of mean temperature and wind speed as input had better performance than other scenarios in estimating daily evaporation. Al-Mukhtar [25] predicted the evaporation rate from the pan in Basra, Mosul, and Baghdad in Iraq. Input parameters for artificial intelligence models were minimum and maximum temperature, relative humidity, and wind speed. Quantile Regression Forests Model had better performance than others. Ashrafzadeh et al. [26] by using the models MLP, SVM, and SOMNN, the pan evaporation in Bandar Anzali and Astara, two cities in northern Iran were estimated. Based on the high humidity of the studied regions, the input parameters of the three models included minimum, maximum, and average temperature, minimum, maximum, and average relative humidity, rain, wind speed, and sunshine duration. The SOMNN model worked more properly. Singh et al. [27] in another research, the two ANN and MLR models were used for the estimation of evaporation. The inputs of the neural network were rain, relative humidity delayed for one day, minimum and maximum temperature. The efficiency and correlation coefficient of the model ANN during calibration and validation were higher than MLR, but the amount of RMSE in the MLR model was higher.

For accurate prediction with neural networks, a lot of data is needed, so in this study, the gmdh neural network was used to predict evaporation from the pan. considering that the GMDH-NN intelligent model has not been extensively studied in estimating the daily evaporation rate of pan, this study tried to investigate the efficiency of the GMDH-NN model to estimate pan evaporation in Garmsar located in Semnan province, Iran. In this study, meteorological data of Garmsar synoptic station for 10 years (2009- 2018) were used for modeling. Then, the sensitivity analysis of input parameters was performed.

2. Methods

2.1. Study area

Garmsar city is located in the west of Semnan province in Iran. The distance between this city and the capital of Iran (Tehran) is 114 km. The minimum longitude of Garmsar is 51 degrees and 51 minutes, and the minimum northern latitude is 34 degrees and 18 minutes. The height of the meteorological station of Garmsar city center is 899.9 m from sea level.

Data were analyzed from 2009 to 2018. Data included daily minimum and maximum temperature, mean temperature, air pressure, relative humidity, sunny hours, and wind speed that were received from the main synoptic station of Garmsar city. The total position of the studied station is seen in Figure 1.

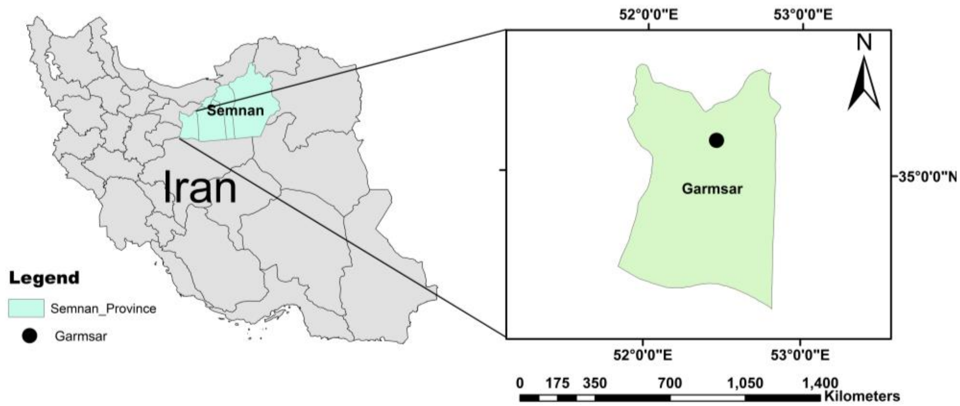


Fig. 1. Geographical position of Garmsar city

2.2. Parameters and statistical specifications of data

In this study, the GMDH-NN model's efficiency for predicting pan evaporation was evaluated based on the data on the minimum and maximum temperature, average temperature, relative humidity, wind speed, sunny hours, and air pressure, all of which are daily parameters.

Table 1 shows the studied parameters, abbreviation, and statistical specifications of this research. Figure 2 shows the histogram of the input and output data. For better performance, the input and output data have been normalized using Relation 1 and Table 2. Thus, all data were between 0.1 and 0.9 and then used to develop the Relation. This method was used as in the studies [28–30].

Table 1
Statistical Specifications of Data

Parameter	Unit	Symbol	Mean	Standard deviation	Minimum	Maximum
Minimum temperature	c̊	T_{min}	13.07	9.71	-12.6	35
Maximum temperature	c̊	T_{max}	26.31	11.08	-1.6	47
Relative humidity	%	RH_{mean}	37.14	19.32	4.5	97.625
Wind speed	m/s	WS	7.27	3.99	0	35
Sunshine hours	hr	n	8.79	3.27	0	13.8
Air pressure	hPa	Pa	914.84	6.26	896.88	936.98
Evaporation	mm	E	6.69	7.36	0	39.1

$$Parameter_{Scaled} = \left[(0.8 \left(\frac{Parameter - Parameter_{\min}}{Parameter_{\max} - Parameter_{\min}} \right) \right] + 0.1 \quad (1)$$

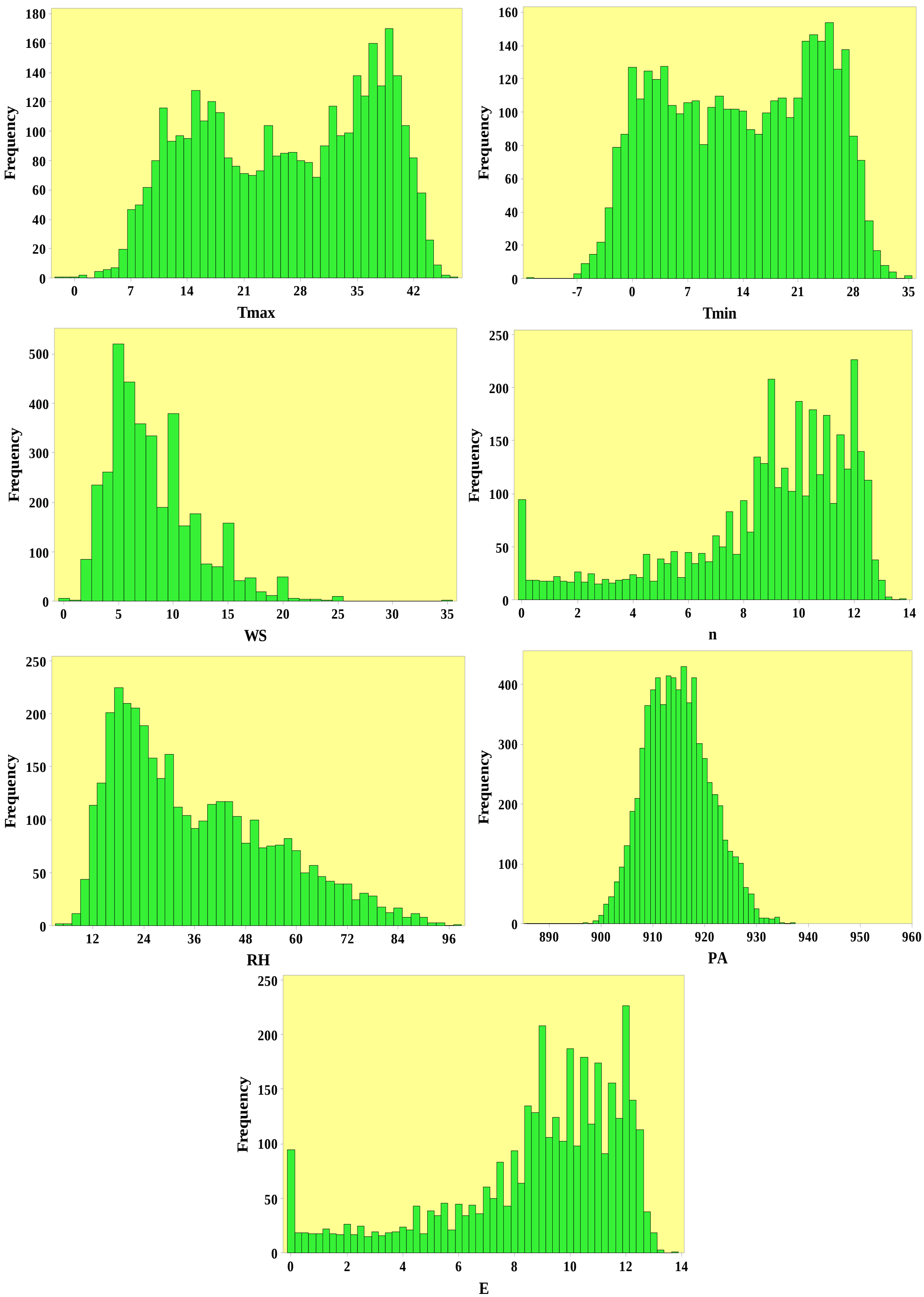


Fig. 2. Histogram of the input and output data

Table 2
Normalization of the considered data.

Symbol	Parameter	Normalized value
T_{min}	Minimum temperature	$T_{\min_{normal}} = 0.8\frac{T_{\min} - (-6.87)}{35.41} + 0.1$
T_{max}	Maximum temperature	$T_{\max_{normal}} = 0.8\frac{T_{\max} - 2.14}{38.26} + 0.1$
RH_{mean}	Relative humidity	$RH_{mean_{normal}} = 0.8\frac{RH_{mean} - 13.8}{63.07} + 0.1$
WS	Wind speed	$WS_{normal} = 0.8\frac{WS - 1.71}{8.55} + 0.1$
n	Sunshine hours	$n_{normal} = 0.8\frac{n - 3.73}{8.81} + 0.1$
PA	Air pressure	$PA_{normal} = 0.8\frac{PA - 878.11}{16.52} + 0.1$
E	Evaporation	$E_{normal} = 0.8\frac{P - 0}{87.2} + 0.1$

2.3. Class a evaporation pan

Class A pan is used in synoptic stations of Iran to estimate the evaporation rate. The evaporation data were collected using this pan at the synoptic station of Garmsar. Class A pan is one of the most known types of standard evaporation pans used to directly measure the evaporation rate. Internal diameter, depth, and water depth are 120, 25, and 20 cm, respectively. The pan has been painted with a galvanized sheet. The pan is placed on the wooden bases with a height of 15 cm to be protected against heat exchanges with the ground with air rotation below it. Figure 3 shows different parts of the evaporation pan Class A.

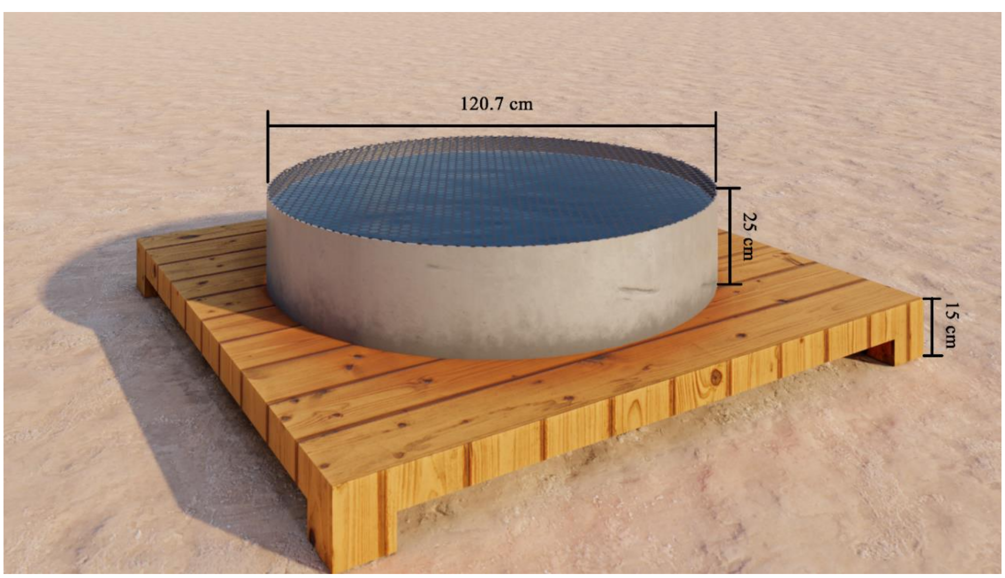


Fig. 3. Class A standard evaporation pan.

2.4. Neural network-based group method of data handling (GMDH-NN) model

The GMDH algorithm was first introduced by a Ukrainian scientist named Ivakhnenko [31]. GMDH neural network is a self-organizing and unilateral network obtained from several layers, each composed of several neurons. All neurons have a similar structure, so that they have two inputs and one output, and each neuron with six weights and one bias establishes the processing operation among the input and output data based on Relation 2.

$$y_{ik}^{*} = N(x_i, x_j) = b^k + w_1^k x_{i\alpha} + w_2^k x_{i\beta} + w_3^k x_{i\alpha}^2 + w_4^k x_{i\beta}^2 + w_5^k x_{i\alpha} x_{i\beta} \tag{2}$$

In Relation 2, (i=1,2,3,…, N) where N is the number of observations and (K=1,2,3,…, C_m^2) and $\beta \in \{1,2,3,...,m\}$ where m is the number of the previous layer neurons.

The weights are calculated based on the Minimum Mean Square Error method and then substituted inside each neuron as specified and constant values.

The obvious characteristic of such a network is that the neurons of the previous phase or the previous layer produce new neurons C_m^2 obtained from Relation 3.

Some of the produced neurons are necessarily removed to prevent network divergence, and neurons that remain to expand the network may also be removed due to lack of direct or indirect communication with the last layer and creating a network convergence form, called passive neurons . The criterion for excluding and selecting a set of neurons in a layer is the Mean Square Error (MSE) between the real output and output of each neuron. This criterion for jth neuron output, i.e. (y_{ij}^{*}), is shown as Relation 4.

$$C_m^2 = \frac{m(m-1)}{2} \tag{3}$$

$$mse_j = \frac{\sum_{i=1}^{N} (y_i - y_i^{*})^2}{N} \tag{4}$$

In the above Relation, $j \in \{1,2,3,...,C_m^2\}$ where m is the number of selected neurons in the previous layer. The mapping established between the input and output variables with such neural networks as the Volterra nonlinear function is shown as Relation 5.

$$\hat{y} = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} a_{ijk} x_i x_j x_k \tag{5}$$

The structure considered for neurons will be in a brief form of two quadratic variables of Relation 5:

$$\hat{y} = f(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i^2 + a_4 x_j^2 + a_5 x_i x_j \tag{6}$$

Unknown coefficients α_i in Equation 6 are obtained with regression technique such that the difference between real output y and calculated values $\hat{y}$ for each pair of the input variables x_i x_j is minimized. Sets of polynomials are obtained using Equation 6, of which unknown coefficients are obtained with the least-squares (LS) method. For each function G_i (each neuron produced), the coefficients of equations of each neuron are obtained to minimize its error in order to optimally match inputs with all pairs of input-output sets (Relation 7) [32].

$$E = \frac{\sum_{i=1}^{m} (y_i - G_i)^2}{m} \rightarrow \min \tag{7}$$

In the GMDH algorithm basic method, all binary compounds (neurons) have made of the n input variable, and unknown coefficients of all neurons are obtained using the Least Squares Method. Therefore, neurons are made in the second layer according to Relation 8, displayed as Set 9.

$$\binom{n}{2} = \frac{n(n-1)}{2} \tag{8}$$

$$\left\{ (y_i, x_{ip}, x_{iq} \left| \begin{matrix} (i = 1, 2, \ldots, m) \\ p, q \in (1, 2, \ldots, m) \end{matrix} \right. \right\} \tag{9}$$

We use the quadratic form of the function expressed in Equation 6 for each M of the triple row. These equations can be expressed as the matrix (10):

$$Aa = Y \tag{10}$$

Where A is the unknown coefficient vector of the quadratic Equation shown in Equation (6), i.e.,

$$a = \{a_0, a_1, \ldots, a_5\} \tag{11}$$

And

$$Y = \{y_1, y_2, \ldots, y_m\} \tag{12}$$

It can be easily shown from values of the input vectors and function form:

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & x_{1p}^{\ 2} & x_{1q}^{\ 2} \\ 1 & x_{2q} & x_{2q} & x_{2p}x_{2q} & x_{2p}^{\ 2} & x_{2q}^{\ 2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_{Mq} & x_{Mq} & x_{Mp}x_{Mq} & x_{Mp}^{\ 2} & x_{Mq}^{\ 2} \end{bmatrix} \tag{13}$$

The Least Squares method of the multiple regression analysis yields the solution of equations as Equation (14):

$$a = (A^T A)^{-1} A^T Y \tag{14}$$

This Equation creates the coefficients vector of Equation (5) for the whole M triad set [32].
In this study, GMGH neural network was used to model the evaporation pan of Garmsar Station for 10 years (6 years for modeling the evaporation pan of Garmsar station and 4 years for the model validation). The percent of the Train data, the percent of the Validation data, the number of layers used, and the highest number of the used neurons are presented in Table 3. Figure 4 shows a schematic view of a gmdh -nn algorithm.

Table 3
Specifications of the GMDH Neural Network trained for Modeling Evaporation Pan of Garmsar Station

P_Train	**P_Validation**	**Max number of nrouns**	**Max number of layars**
60%	40%	30	5

2.5. Modeling accuracy assessment criteria

Statistical indices of the R-squared coefficient (R^2)[33], Root Mean Square Error (RMSE) [34,35], and Mean Absolute Error (MAE)[36] were calculated to evaluate the accuracy of the intelligent models[37]. The values of these indicators are calculated from the following relations:

$$R^2 = \left[\frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} \right]^2 \tag{15}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - x_i)^2}{N}} \tag{16}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{n} \left| (y_i - x_i) \right| \tag{17}$$

In relations (15) to (17), x_i is the evaporation rate measured per day, y_i is the predicted evaporation rate of the same day, $\bar{x}$ is the average values of the measured evaporation, and $\bar{y}$ is the corresponding mean for the predicted values.

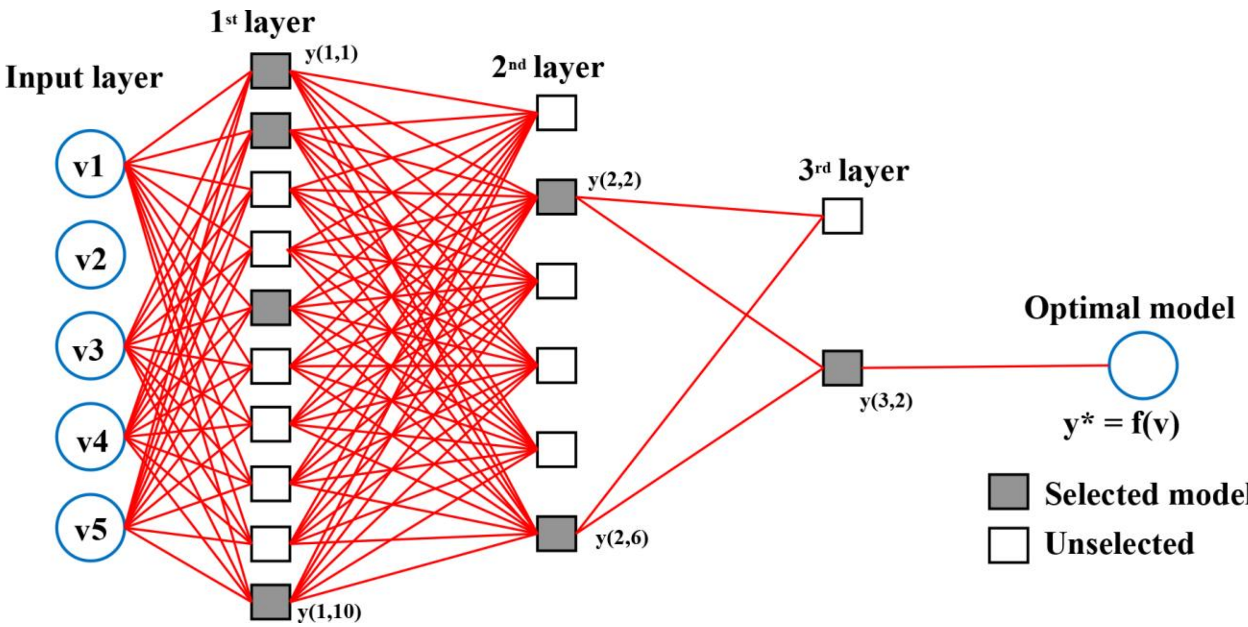


Fig. 4. An illustration of modeling process for the GMDH-NN.

3. Results

3.1. Results of GMDH-NN modeling

In this study, the pan evaporation was calculated every day through the nonlinear GMDH model in the Garmsar station. R^2, RMSE, and MAE evaluation criteria were used to evaluate the performance of this model. Table 4 shows the results of GMDH neural network layers in modeling the evaporation pan of the Garmsar station. This table presents the number of neurons used in each layer of the trained GMDH neural network and the Mean Squared Error of the Validation data for the best neuron of each layer in simulating the evaporation pan of the Garmsar station. The results of Table 4 show that the Mean Squared Error of the Validation data has converged to 0.062 from the third layer onwards, and it is not cost-effective to use more layers in modeling the evaporation pan of Garmsar station. According to Table 5, the GMDH-NN model was evaluated for two training and testing phases. The R-squared coefficient of the model in the two training and testing phases is 0.86 and 0.84, respectively, indicating that the model has a good performance.

Figure 5 shows the time series of the measured and simulated values with the GMDH model. The horizontal axis shows the time series (in terms of the month), and the vertical axis shows evaporation values (mm). The more simulated values match the measured values, the more accuracy and the less error the model has.

Table 4
Number of layers, number of neurons used, and the best Mean Square Error in GMDH neural network trained to model evaporation pan of Garmsar station.

Layar number	Number of nrouns used	Best validation RMSE
1	10	0.065
2	25	0.063
3	30	0.062
4	30	0.062
5	1	0.062

Table 5
The value of error measurement parameters and the accuracy of the proposed model based on GMDH-NN.

	R^2	MAE	RMSE
Training	0.86	1.78	2.49
Testing	0.84	1.91	2.65

Figure 6 shows the data predicted by the GMDH-NN model based on the measurement data in two training and testing phases. The horizontal axis shows the measured evaporation data (mm), and the vertical axis shows the simulated evaporation data (mm). The less the dispersion of data around the best fitting line, the more correlation and the fewer errors are achieved. As can be seen, the correlation between the measured and simulated fitting data in two training and test phases with correlation coefficients of 0.8081 and 0.8598 is relatively high. In addition, Figure 7 shows that the error values obtained from the developed model based on the group method of data handling are small and can estimate the daily evaporation values of the pan. In Figure 8, the histogram of the error obtained from the modeling is drawn. The mean and standard deviation values of errors are -0.1210 and 2.552, respectively.

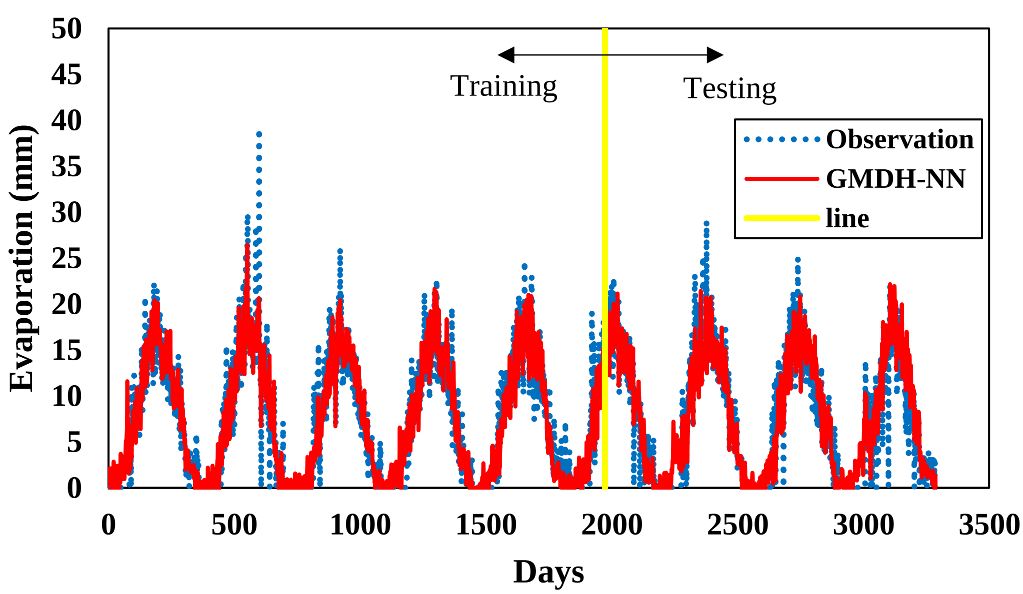
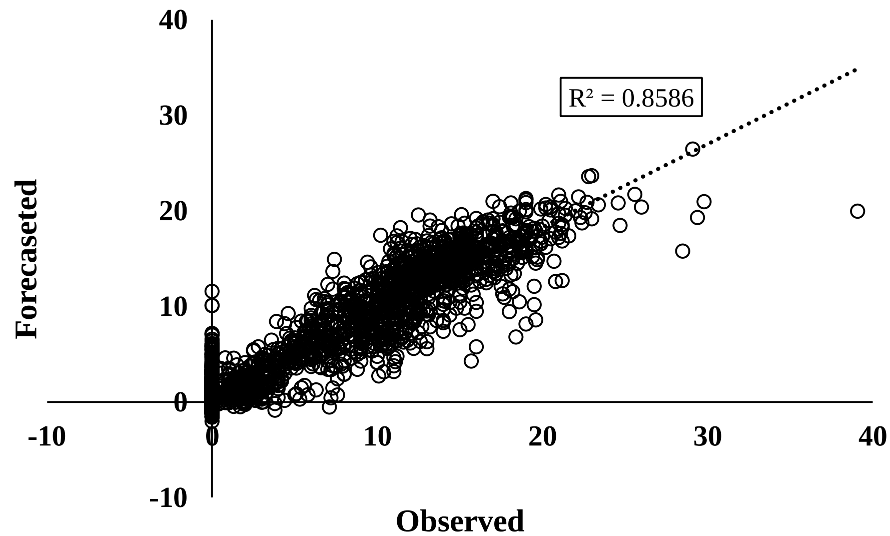
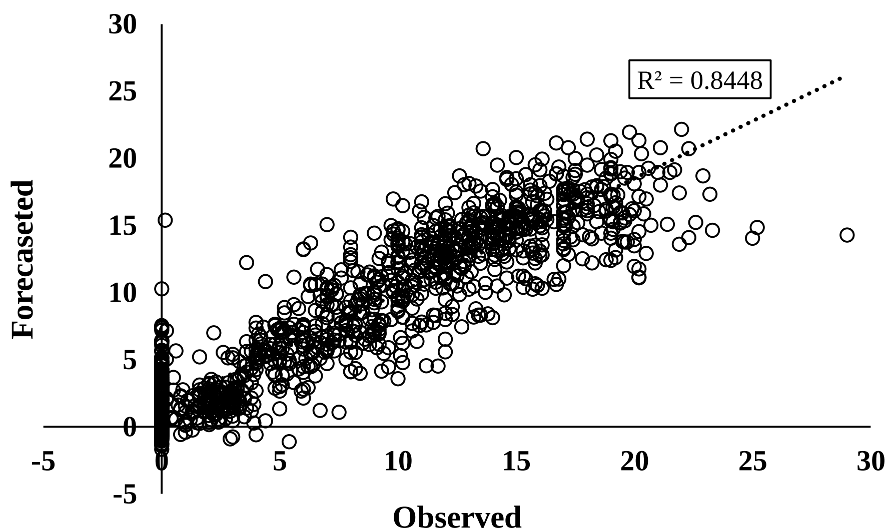


Fig. 5. Time series of observational and predicted values using GMDH-NN model



a



b

Fig. 6. Daily evaporation values simulated with GMDH-NN model based on the measured values a) training and b) test

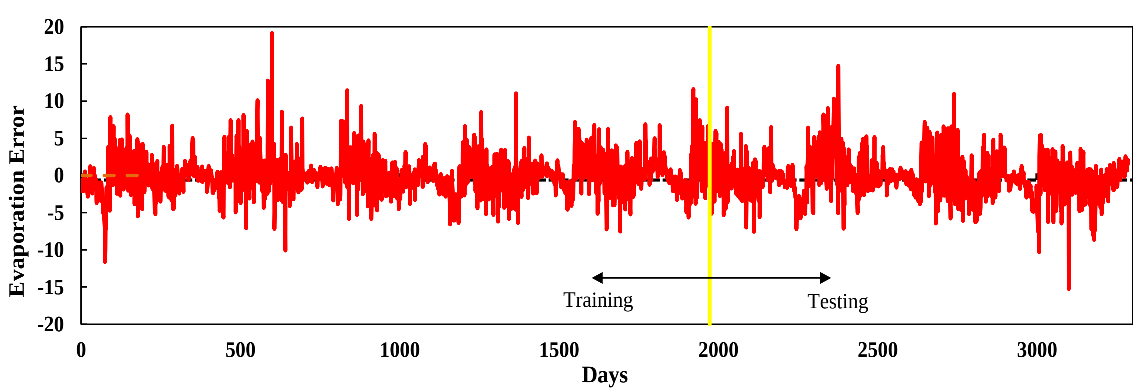


Fig. 7. Errors obtained from the developed relation in the training and test set.

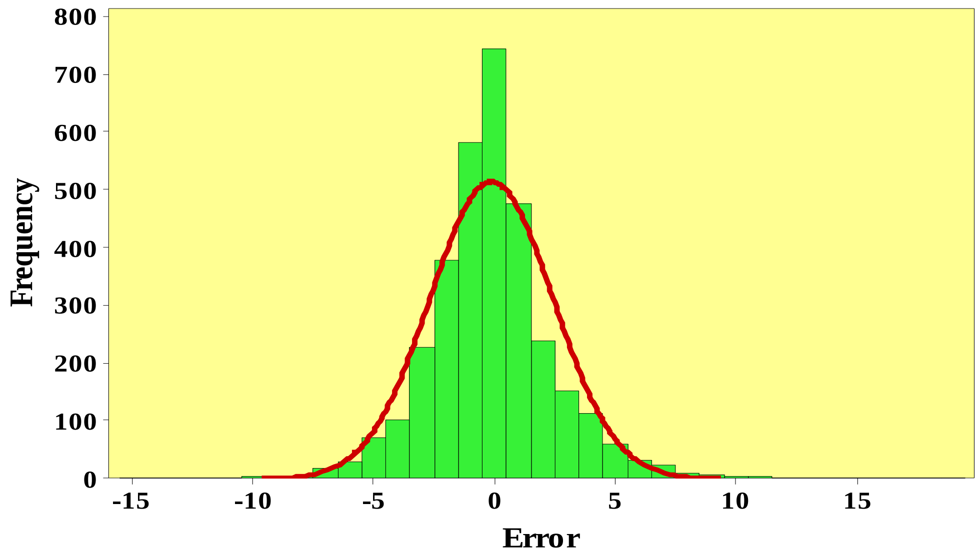


Fig. 8. Simulation error histogram with GMDH-NN

3.2. Sensitivity analysis results

GMDH neural network training results showed that after sorting neurons based on the mean squares value and removing additional neurons based on standard error, 10 neurons were used in the first layer (Table 4). The order of inputs used in the first layer neurons is shown in Table 6. Given the results of Table 6, input data 1 and 6, which indicate the minimum temperature and the mean relative humidity percent, have the lowest error in simulating the value of pan evaporation in Garmsar station, followed by two inputs of the minimum and maximum temperatures with the lowest error in simulating the evaporation pan of Garmsar station. The results of Table 6, in which the values of the best Mean Squared Error of the Validation data have been presented, show that the best Mean Squared Error values have not changed significantly with increasing the layers (maximum 0.003). Therefore, two input parameters of minimum temperature and relative humidity percent significantly impact the evaporation pan's modeling. Therefore, three input parameters of minimum temperature, maximum temperature, and relative humidity percent were selected as sensitive parameters for simulation of evaporation pan of Semnan station.

Table 6

Mean Squared Error values in the arranged neurons of the first layer of GMDH neural network.

Nroun number	Input data number	Input data name	RMSE of validation data
1	[1,6]	T_{min}, RH_{mean}	0.0651
2	[1,2]	T_{min}, T_{max}	0.0662
3	[1,3]	T_{min}, n	0.0672
4	[2,4]	T_{max}, WS	0.0674
5	[1,5]	T_{min}, PA	0.0676
6	[2,6]	T_{max}, RH_{mean}	0.0680
7	[2,5]	T_{max}, PA	0.0681
8	[2,3]	T_{max}, n	0.0682
9	[1,4]	T_{min}, WS	0.0682
10	[5,6]	PA, RH_{mean}	0.0856

4. Discussion

This study aimed to investigate the efficiency of the GMDH-NN model in simulating the value of pan evaporation for Garmsar station located in Semnan province, Iran.The study interval was from 2009 to 2018. The value of the R-squared coefficient during the test period was approximately 0.84. The mean squared error values of the first neuron in each layer of the GMDH neural network showed that the RMSE value converged to 0.062 mm after the three layers and it is not economical to use more layers in modeling the evaporation pan of this station. Also, after sorting the mean squared error in the first layer neurons, two input parameters of minimum temperature and mean relative humidity had the lowest RMSE values of 0.0651 and were selected as two sensitive parameters in simulating the evaporation pan of Garmsar station.

Comparing the present study results with the study by Asharfzadeh et al. (2020) [23] shows that GMDH neural network model has the necessary efficiency in estimating pan evaporation. Karbasi (2016) [38] studied the GMDH model to estimate the evaporation of synoptic stations in Ahwaz. The results of this study show that the mentioned model can be effective in estimating effective evaporation and can model nonlinear behaviors. In the sensitivity analysis, the results of the research by Traore et al. (2010) [39] and Nourani and Sayyah Fard (2012) [40] investigated the evaporation using the neural network method. In this research, the most effective parameter was the temperature, which is in line with the results of this study.

The functions of the intelligent methods in estimating evaporation in arid climates are similar to the results observed in the previous studies, which have investigated the application of intelligent methods for modeling evaporation rates under different climates. Piri et al. (2009) [41] were among the first to use ANNs to model pan evaporation rates in arid and semi-arid climates. They reported satisfactory performance for ANNs used in a research site located in the southeast of Iran. Their study reported R2 = 0.93 for an ANN model with an optimal combination of 4 meteorological inputs. In the present study, the best obtained R^2 value was 0.84 during the test period, as shown in Table 3. This shows that the models based on intelligent methods are effective in arid climates. In addition, the use of artificial intelligence models in arid regions has the same prediction error reported in the present study considering the high rate of pan evaporation. However, in arid climates, the frequency of such high evaporation rates is higher than the pan, leading to lower model performance.

The present study results are also comparable with the results of other artificial intelligence methods used in similar climates. Moghaddamnia et al. (2009) [14] used the ANFIS method on a research site located in the Southeast. The R^2 value was reported 0.91 for the best performance of the ANFIS model during the validation period. However, a similar prediction orientation was observed for the ANFIS model. Therefore, further research should improve artificial intelligence techniques to allow more reliable predictions for the high rate of pan evaporation. A correction-deviation method may indicate a suitable approach in this field. In addition, further studies may consider other meteorological variables. Abusada (1988) [42] compared class A pan evaporation data collected from Kuwait Airport Station from 1962 to 1977 with theoretical calculations of evaporation estimation using the Penman method at the same station for the same period.

The comparison showed that the Penman method estimated an annual evaporation rate of 2630 mm, while the annual evaporation rate of the pan measured for the same period was 3540 mm. Therefore, the error resulting from the Penman method is 910 mm per year. Although the Penman method is one of the best practical methods for estimating evaporation in an arid climate [43], it was found that this method had poor performance compared to the artificial neural network models. The evaporation pan wall prevents additional sunlight and increases heat exchange with the surrounding atmosphere [44].Therefore, physical models cannot be used to estimate the evaporation of the pan directly.

5. Conclusions

In this study, estimation of pan evaporation values in Garmsar city located in Semnan province, Iran, was investigated using the developed neural network model of group method of data handling. In this research, R^2, RMSE, and MBE evaluation criteria were used to evaluate the model results. The results of this study show that the GMDH-NN model is suitable for modeling "evaporation pan processes". Also, the mean squared error values in the first neurons of the trained GMDH neural network layers showed that the RMSE value of the first neuron in the third layer converges to 0.062 mm and it is not necessary to use more layers. The mean squared error values in the first layer neurons showed that the two input parameters of minimum temperature and relative humidity have the lowest RMSE values (0.0651 mm) in simulating the amount of evaporation pan of Garmsar station and were selected as two sensitive parameters. Arid and semi-arid climates have unique climate regimes that are characterized by "scarce water resources, "bare vegetation," and "high evaporation rate ".Considering the report of the Food and Agriculture Organization of the United Nations (FAO), excessively arid climates are defined as the regions where annual precipitation does not exceed 3% of the annual evaporation. Comparing the performance of the ANN model with other practical models is essential by dealing with the subject of contribution of pan wall to heat exchange. Then, it can be noted that considering that the time behavior of daily evaporation is non-stationary, it is better to investigate and compare other intelligent methods for modeling.

Conflicts of interest

The authors declare no conflict of interest.

References

[1] Helfer F, Lemckert C, Zhang H. Impacts of climate change on temperature and evaporation from a large reservoir in Australia. J Hydrol 2012;475:365–78. doi:10.1016/j.jhydrol.2012.10.008.

[2] Ghazvinian H, Karami H, Farzin S, Mousavi SF. Effect of MDF-Cover for Water Reservoir Evaporation Reduction, Experimental, and Soft Computing Approaches. J Soft Comput Civ Eng 2020;4:98–110. doi:10.22115/scce.2020.213617.1156.

[3] Sima S, Ahmadalipour A, Tajrishy M. Mapping surface temperature in a hyper-saline lake and investigating the effect of temperature distribution on the lake evaporation. Remote Sens Environ 2013;136:374–85. doi:10.1016/j.rse.2013.05.014.

[4] Ghazvinian H, Farzin S, Karami H, Mousavi SF. Investigating the Effect of using Polystyrene sheets on Evaporation Reduction from Water-storage Reservoirs in Arid and Semiarid Regions (Case study: Semnan city). J Water Sustain Dev 2020;7:45–52. doi:10.22067/jwsd.v7i2.81748.

[5] Torres EA, Calera A. Bare soil evaporation under high evaporation demand: a proposed modification to the FAO-56 model. Hydrol Sci J 2010;55:303–15. doi:10.1080/02626661003683249.

[6] Wang L, Niu Z, Kisi O, Li C, Yu D. Pan evaporation modeling using four different heuristic approaches. Comput Electron Agric 2017;140:203–13. doi:10.1016/j.compag.2017.05.036.

[7] Ghazvinian H, Karami H, Farzin S, Mousavi SF. Experimental Study of Evaporation Reduction Using Polystyrene Coating, Wood and Wax and its Estimation by Intelligent Algorithms. Irrig Water Eng 2020;11:147–65. doi:10.22125/iwe.2020.120727.

[8] Miralles DG, Jiménez C, Jung M, Michel D, Ershadi A, McCabe MF, et al. The WACMOS-ET project – Part 2: Evaluation of global terrestrial evaporation data sets. Hydrol Earth Syst Sci 2016;20:823–42. doi:10.5194/hess-20-823-2016.

[9] Teng J, Yasufuku N, Liu Q, Liu S. Experimental evaluation and parameterization of evaporation from soil surface. Nat Hazards 2014;73:1405–18. doi:10.1007/s11069-014-1138-z.

[10] Kumar N, Arakeri JH. Experimental and numerical investigation of evaporation from line sources of water in low porosity surfaces. J Hydrol 2019;569:795–808. doi:10.1016/j.jhydrol.2019.01.001.

[11] Zheng J, Chen L, Wang J, Zhou Y, Wang J. Thermodynamic modelling and optimization of self-evaporation vapor cooled shield for liquid hydrogen storage tank. Energy Convers Manag 2019;184:74–82. doi:10.1016/J.ENCONMAN.2018.12.053.

[12] Irmak S, Haman DZ, Jones JW. Evaluation of Class A Pan Coefficients for Estimating Reference Evapotranspiration in Humid Location. J Irrig Drain Eng 2002;128:153–9. doi:10.1061/(ASCE)0733-9437(2002)128:3(153).

[13] Kumar M, Raghuwanshi NS, Singh R, Wallender WW, Pruitt WO. Estimating Evapotranspiration using Artificial Neural Network. J Irrig Drain Eng 2002;128:224–33. doi:10.1061/(ASCE)0733-9437(2002)128:4(224).

[14] Moghaddamnia A, Ghafari Gousheh M, Piri J, Amin S, Han D. Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. Adv Water Resour 2009;32:88–97. doi:10.1016/j.advwatres.2008.10.005.

[15] Wang L, Kisi O, Hu B, Bilal M, Zounemat-Kermani M, Li H. Evaporation modelling using different machine learning techniques. Int J Climatol 2017;37:1076–92. doi:10.1002/joc.5064.

[16] Shimi M, Najjarchi M, Khalili K, Hezavei E, Mirhoseyni SM. Investigation of the accuracy of linear and nonlinear time series models in modeling and forecasting of pan evaporation in IRAN. Arab J Geosci 2020;13:59. doi:10.1007/s12517-019-5031-7.

[17] Sebbar A, Heddam S, Djemili L. Kernel extreme learning machines (KELM): a new approach for modeling monthly evaporation (EP) from dams reservoirs. Phys Geogr 2020:1–23. doi:10.1080/02723646.2020.1776087.

[18] Qasem SN, Samadianfard S, Kheshtgar S, Jarhan S, Kisi O, Shamshirband S, et al. Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural

networks in arid and humid climates. Eng Appl Comput Fluid Mech 2019;13:177–87. doi:10.1080/19942060.2018.1564702.

[19] Patle GT, Chettri M, Jhajharia D. Monthly pan evaporation modelling using multiple linear regression and artificial neural network techniques. Water Supply 2020;20:800–8. doi:10.2166/ws.2019.189.

[20] Adnan RM, Malik A, Kumar A, Parmar KS, Kisi O. Pan evaporation modeling by three different neuro-fuzzy intelligent systems using climatic inputs. Arab J Geosci 2019;12:606. doi:10.1007/s12517-019-4781-6.

[21] J. M. Bruton, R. W. McClendon, G. Hoogenboom. ESTIMATING DAILY PAN EVAPORATION WITH ARTIFICIAL NEURAL NETWORKS. Trans ASAE 2000;43:491–6. doi:10.13031/2013.2730.

[22] Keskin ME, Terzi Ö. Artificial Neural Network Models of Daily Pan Evaporation. J Hydrol Eng 2006;11:65–70. doi:10.1061/(ASCE)1084-0699(2006)11:1(65).

[23] Ashrafzadeh A, Kişi O, Aghelpour P, Biazar SM, Masouleh MA. Comparative study of time series models, support vector machines, and GMDH in forecasting long-term evapotranspiration rates in northern Iran. J Irrig Drain Eng 2020;146:4020010.

[24] Alsumaiei AA. Utility of Artificial Neural Networks in Modeling Pan Evaporation in Hyper-Arid Climates. Water 2020;12:1508. doi:10.3390/w12051508.

[25] Al-Mukhtar M. Modeling the monthly pan evaporation rates using artificial intelligence methods: a case study in Iraq. Environ Earth Sci 2021;80:39. doi:10.1007/s12665-020-09337-0.

[26] Ashrafzadeh A, Malik A, Jothiprakash V, Ghorbani MA, Biazar SM. Estimation of daily pan evaporation using neural networks and meta-heuristic approaches. ISH J Hydraul Eng 2020;26:421–9. doi:10.1080/09715010.2018.1498754.

[27] Singh A, Singh RM, Kumar ARS, Kumar A, Hanwat S, Tripathi VK. Evaluation of soft computing and regression-based techniques for the estimation of evaporation. J Water Clim Chang 2021;12:32–43. doi:10.2166/wcc.2019.101.

[28] Ghazvinian H, Bahrami H, Ghazvinian H, Heddam S. Simulation of Monthly Precipitation in Semnan City Using ANN Artificial Intelligence Model. J Soft Comput Civ Eng 2020;4:36–46. doi:10.22115/scce.2020.242813.1251.

[29] Naderpour H, Rezazadeh Eidgahee D, Fakharian P, Rafiean AH, Kalantari SM. A new proposed approach for moment capacity estimation of ferrocement members using Group Method of Data Handling. Eng Sci Technol an Int J 2020;23:382–91. doi:10.1016/j.jestch.2019.05.013.

[30] Naderpour H, Nagai K, Fakharian P, Haji M. Innovative models for prediction of compressive strength of FRP-confined circular reinforced concrete columns using soft computing methods. Compos Struct 2019;215:69–84. doi:10.1016/j.compstruct.2019.02.048.

[31] Ivakhnenko AG. Polynomial Theory of Complex Systems. IEEE Trans Syst Man Cybern 1971;SMC-1:364–78. doi:10.1109/TSMC.1971.4308320.

[32] Farlow SJ. Self-organizing methods in modeling: GMDH type algorithms. Vol 54 CrC Press 1984.

[33] Ehteram M, Karami H, Farzin S. Reservoir Optimization for Energy Production Using a New Evolutionary Algorithm Based on Multi-Criteria Decision-Making Models. Water Resour Manag 2018;32:2539–60. doi:10.1007/s11269-018-1945-1.

[34] Basser H, Karami H, Shamshirband S, Jahangirzadeh A, Akib S, Saboohi H. Predicting optimum parameters of a protective spur dike using soft computing methodologies – A comparative study. Comput Fluids 2014;97:168–76. doi:10.1016/j.compfluid.2014.04.013.

[35] Ghazvinian H, Mousavi S-F, Karami H, Farzin S, Ehteram M, Hossain MS, et al. Integrated support vector regression and an improved particle swarm optimization-based model for solar radiation prediction. PLoS One 2019;14:e0217634. doi:10.1371/journal.pone.0217634.

[36] Ehteram M, Karami H, Mousavi S-F, Farzin S, Kisi O. Evaluation of contemporary evolutionary algorithms for optimization in reservoir operation and water supply. J Water Supply Res Technol - Aqua 2018;67:54–67. doi:10.2166/aqua.2017.109.

[37] Dianatikhah M, Karami H, Hosseini K. Generation of Clean Hydropower Energy in Multi-Reservoir Systems Based on a New Evolutionary Algorithm. Water Resour Manag 2020;34:1247–64.

[38] Karbasi M. Forecasting of Daily Reference Evapotranspiration at Ahvaz synoptic station using wavelet-GMDH hybrid model. J Water Soil Conserv 2016;23:323–30. doi:10.22069/jwfst.2016.9610.2385.

[39] Traore S, Wang Y-M, Kerh T. Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. Agric Water Manag 2010;97:707–14. doi:10.1016/j.agwat.2010.01.002.

[40] Nourani V, Sayyah Fard M. Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes. Adv Eng Softw 2012;47:127–46. doi:10.1016/j.advengsoft.2011.12.014.

[41] Piri J, Amin S, Moghaddamnia A, Keshavarz A, Han D, Remesan R. Daily Pan Evaporation Modeling in a Hot and Dry Climate. J Hydrol Eng 2009;14:803–11. doi:10.1061/(ASCE)HE.1943-5584.0000056.

[42] Abusada SM. The essentials of groundwater resources of Kuwait. Kuwait Inst Sci Res Rep No KISR 1988;2665.

[43] Brutsaert W. Evaluation of some practical methods of estimating evapotranspiration in arid climates at low latitudes. Water Resour Res 1965;1:187–91. doi:10.1029/WR001i002p00187.

[44] Linacre ET. Estimating U.S. Class A Pan Evaporation from Few Climate Data. Water Int 1994;19:5–14. doi:10.1080/02508069408686189.