



Original Paper

A New Approach for Regional Groundwater Level Simulation: Clustering, Simulation, and Optimization

Zahra Kayhomayoon,^{1,5} Sami Ghordoyee Milan,² Naser Arya Azar,³ and Hamid Kardan Moghaddam⁴

Received 15 October 2020; accepted 20 June 2021
Published online: 13 July 2021

In this study, a new 3-stage approach that consists of clustering, simulation, and optimization stages is proposed for the simulation of groundwater level (GWL) in an arid region of eastern Iran. In the first stage, *K*-means clustering was used to divide the study aquifer into five different clusters based on precipitation, water recharge, water discharge, transmissivity, earth level, and water table. In the second stage, to simulate GWL in each cluster, several input variables, such as water level at the previous month, aquifer discharge, aquifer recharge, evaporation, temperature, and precipitation, were used in the form of various input patterns that were fed to an artificial neural network (ANN). Finally, in the third stage, two advanced optimization methods, i.e., particle swarm optimization (PSO) and whale optimization algorithm (WOA), were utilized to optimize the ANN results. Various patterns were identified as suitable clusters based on the studied models. A pattern including water level at the previous month, aquifer discharge, aquifer recharge, and precipitation was identified as the best model for four clusters, except for cluster 3. The validation with root mean squared error (RMSE), mean absolute percentage error (MAPE), and Nash Sutcliffe index (NSE) revealed RMSE = 0.01, NSE = 0.97, and MAPE = 0.13 for the first cluster, RMSE = 0.011, NSE = 0.99, and MAPE = 0.22 for the second cluster, RMSE = 0.003, NSE = 0.99, and MAPE = 0.30 for the fourth cluster, and RMSE = 0.001, NSE = 0.98, and MAPE = 0.05 for the fifth cluster. For the third cluster, a pattern including water level at the previous month, aquifer discharge, and aquifer recharge was identified as the best model resulting in RMSE = 0.006, NSE = 0.99, and MAPE = 0.05. Finally, according to the results, the ANN-PSO model was applied to three clusters, while the ANN-WOA model was applied to the remaining clusters. In general, this study showed that optimization algorithms can improve the simulation accuracy of ANN, and the efficient use of each method depends on the clustering type. The application of the approach proposed here can be extended to other aquifers that have a relatively large area and limited data availability.

KEY WORDS: Groundwater level, WOA, PSO, Aquifer hydrograph, Spatial clustering.

¹Department of Geology, Payame Noor University, Tehran, Iran.

²Department of Irrigation and Drainage Engineering, Aburaihan Campus, University of Tehran, Tehran, Iran.

³Department of Water Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran.

⁴Department of Water Resources Research, Water Research Institute, Tehran, Iran.

⁵To whom correspondence should be addressed; e-mail: Zkayhomayoon@pnu.ac.ir

INTRODUCTION

One of the approaches to groundwater resources management is the use of modeling and simulation tools to determine the status of water balance. To achieve this important goal, the use of new methods that can reduce the simulation error

and uncertainty of model variables is essential (Milan et al., 2018). Although physical and mathematical models are basic tools for demonstrating hydrogeological variables and understanding the processes that take place in a system, they suffer from practical and temporal limitations. Moreover, they require accurate information on proper inputs, which are often not available in many regions around the world. Therefore, the application of intelligent models is suggested in the case of sparse and incomplete data (Kardan Moghaddam et al., 2019; Nguyen et al., 2020a; Nhu et al., 2020a; Pham et al., 2019). Artificial neural networks (ANNs) are among the most widely used artificial intelligence methods, which have been proven to work well in various simulation studies (Nhu et al., 2020b; Xu et al., 2020). Using the least possible information from a system, this method can develop a regression model for predicting the output with satisfactory performance. Numerous studies have used ANNs with different structures for groundwater level (GWL) and potential simulation and reported the satisfactory performance of this method (Lallahem et al., 2005; Mirarabi et al., 2019; Nguyen et al., 2020b; Taormina et al., 2012). However, some researchers (e.g., Banadkooki et al., 2020; Jaafari et al., 2019a, b, c; Khedri et al., 2020; Kombo et al., 2020; Maroufpoor et al., 2020) believe that primary regression machines such as ANNs should be optimized by optimization methods to achieve the highest accuracy of results. The network structure, parameters, and type of network training algorithm directly affect the quality of prediction results. Although ANNs can use error back-propagation training algorithms for error convergence, they suffer from a low convergence rate, and sometimes, they are trapped in local minima (Asefpour Vakilian, 2020; Sarlaki et al., 2021), calling for the application of advanced optimization algorithms to achieve the best performance.

The application of advanced optimization algorithms such as particle swarm optimization (PSO) has been reported for groundwater management to minimize the cost of pumping (Gaur et al., 2013; Milan et al., 2021). Whale optimization algorithm (WOA) is another optimization algorithm that provides good results for optimization problems (Abd El Aziz et al., 2017; Ling et al., 2017; Mirjalili & Lewis, 2016). WOA has been successfully used to optimize the parameters of ANN, adaptive neuro-fuzzy inference system (ANFIS), and support vector regression (SVR) (Sai & Huajing, 2017; Aljarah

et al., 2018; Heydari et al., 2019; Chen et al., 2019; Mohammadi & Mehdizadeh, 2020; Vaheddoost et al., 2020). Seifi and Soroush (2020) showed that a hybrid ANN–WOA model outperformed ANN and ANN-genetic algorithm (GA) models for simulating water evaporation. Similar results were reported by Samadianfard et al. (2020) for predicting wind speed. Overall, the literature shows that the use of evolutionary optimization models, along with simulation, can provide timely predictions with acceptable performance in many real-world problems.

In this study, the GWL of an aquifer was simulated using an ANN and the PSO and WOA optimization algorithms. This study is the first for proposing and comparing hybrid ANN–WOA and ANN–PSO models for improving the simulation accuracy of GWL with expectation of significantly increasing the computational accuracy and reliability over a single ANN model. In favor of a more accurate simulation process, we also propose a clustering technique that identifies those regions of the study aquifer that represent similar characteristics. Combining the clustering, simulation, and optimization concepts into a single methodological framework distinguish our study from other analogous researches reported in the literature.

MATERIALS AND METHODS

Study Area and Dataset

With an area of 428.9 km², the Birjand aquifer is located in an arid region with a cold climate in eastern Iran. The aquifer is of alluvial type with average thickness of 75 m. The general direction of groundwater flow is from north to south and then to the west of the aquifer. The average saturation thickness of the aquifer is estimated at 25 m. The location of the aquifer in the study area is such that the aquifer is fed from the northern, southern, and eastern parts through surface flows and groundwater, while the output of groundwater flows is located at the western part of the aquifer.

All water demand in this area is supplied from groundwater, and there is a shortage of more than 10 MCM (million cubic meters) of groundwater resources annually (Kardan Moghaddam et al., 2019). Continuation of the current trend of average annual withdrawal in the region of about 50 cm will result in serious environmental problems and a shortage in water supply for drinking and agriculture. An annual

volume of about 90 MCM of water is extracted from this aquifer by about 100 discharge wells. Eighteen observation wells across the aquifer are responsible for monitoring GWL. Figure 1 shows the location of the aquifer and observation wells.

To simulate GWL, independent observational data are needed to estimate the amount of GWL at the end of the month. Therefore, according to previous research (Coulibaly et al., 2001; Jalalkamali et al., 2011; Guzman et al., 2018; Khaki et al., 2015; Ebrahimi & Rajaei, 2017; Rajaei et al., 2019; Karadan Moghaddam et al., 2019), six variables were selected: GWL at the previous month (GWL_{n-1}), precipitation (P), aquifer recharge (R), aquifer discharge (D), temperature (T), and evaporation (E).

Studies show that in parts of aquifers where GWL is close to the surface, two parameters (i.e., temperature and evaporation) are effective on GWL and its simulation (Karadan Moghaddam et al., 2019; Moghaddam et al., 2021). Therefore, these two parameters were considered for simulation of GWL. The time-series data of the climate of the region (precipitation, temperature, and evaporation) were extracted based on the statistics of the synoptic station of the region. GWL data per observation well were obtained from the Regional Water Company. Moreover, the amount of aquifer discharge in a Thiessen polygon network of each observation well was defined based on the inventory of resources and consumptions and the definition of time series during the simulation period. According to the sequence of three periods (2003, 2011, and 2017) and the amount of aquifer over-exploitation, the time series of discharge from each well was determined, and the sum per Thiessen polygon was determined as the discharge per observation well (Karadan Moghaddam et al., 2019; Moghaddam et al., 2021). The amount of aquifer recharge in each Thiessen polygon per observation well was defined as a time series based on the return water coefficient of consumption and infiltration due to precipitation and runoff in the area according to the regional balance reports (Ministry of Power, 2017).

Artificial Neural Network

ANN models have been studied for many years in the hope of achieving performance similar to human performance in speed and cognition (Hopfield, 1988). After selecting the model inputs to an ANN, several parameters such as numbers of hidden

and output layers and number of primary neurons in each middle layer should be determined. Next, the network evaluation criterion is selected to calculate the network's prediction error. The network determines the weights and biases by different algorithms according to training data. This step is repeated until the difference between the observed values and values predicted by the network is minimized (Haykin, 1999). The different architectures of the multilayer perceptron (MLP) network are determined by the number of neurons in the (hidden) layers, the number of hidden layers, and the type of transfer function in the hidden and output layers. A suitable architecture performs simulation with reasonable accuracy. The selected architecture and the appropriate number of neurons in the hidden layers, transfer functions, and the selected algorithm for network training are given in the results section. To perform the modeling, after data normalization in the [0,1] interval, the data were divided into two subsets, namely training data (75% of total data) and test data (rest of data).

Hybrid Models

Figure 2 depicts the flowchart of the present study to combine the ANN model with optimization algorithms for the simulation of GWL. First, by having the effective parameters, the aquifer's clustering was performed using the *K*-means method. Different patterns were developed by combining the variables that affect GWL as the output variable. These patterns were first implemented by the ANN model and then by the modified ANN with PSO and WOA algorithms to obtain a reliable model. Finally, a suitable pattern and a model were proposed for each cluster. In optimization using evolutionary algorithms, optimization variables are weights and biases of the network (Chen et al., 2018; Toghyani et al., 2016). The modeling is such that *N* position vectors are considered for X_i , where the vectors are generated randomly. ANN is executed considering the values of these vectors as its parameters, and minimizing the error obtained from each execution is considered the objective function of the model. This process is repeated until final convergence is achieved, where the weights and biases are optimized so that training error is minimized. Then, the ANN's optimal weights and biases are used, and the results are evaluated. If the results are desirable, the

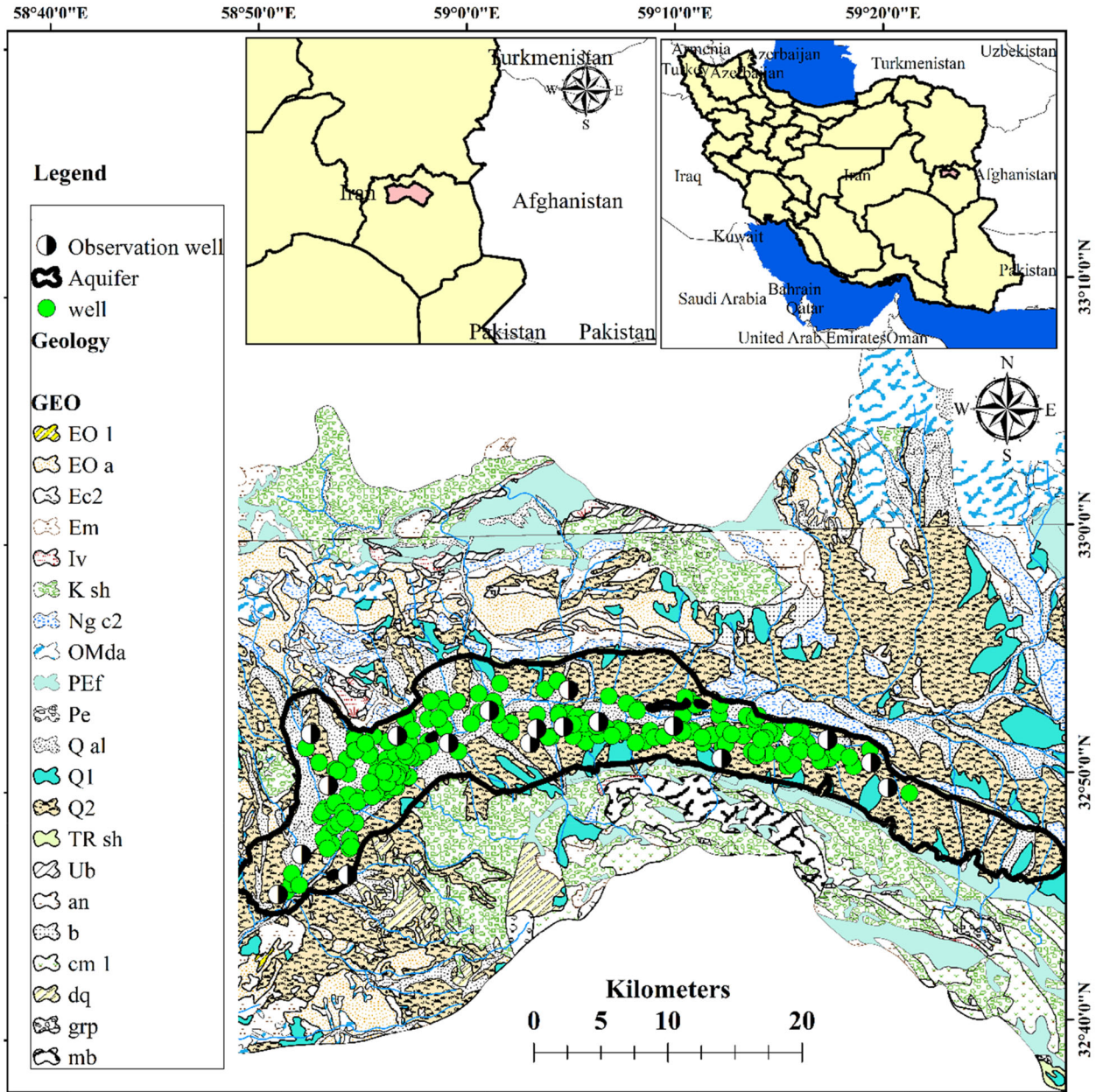


Figure 1. Location of aquifer and observation wells.

model training is completed, and the optimal network evaluates the test data.

Particle Swarm Optimization

Introduced by Kennedy and Eberhart (1995), PSO is a nature-inspired optimization algorithm.

Similar to other optimization algorithms, PSO starts with generating a random population. The components in this method are different sets of decision variables whose optimal values are provided by moving the variables to optimal points with a determined velocity (Arumugam et al., 2008). PSO includes a velocity vector and a position vector, which force the population to change their positions

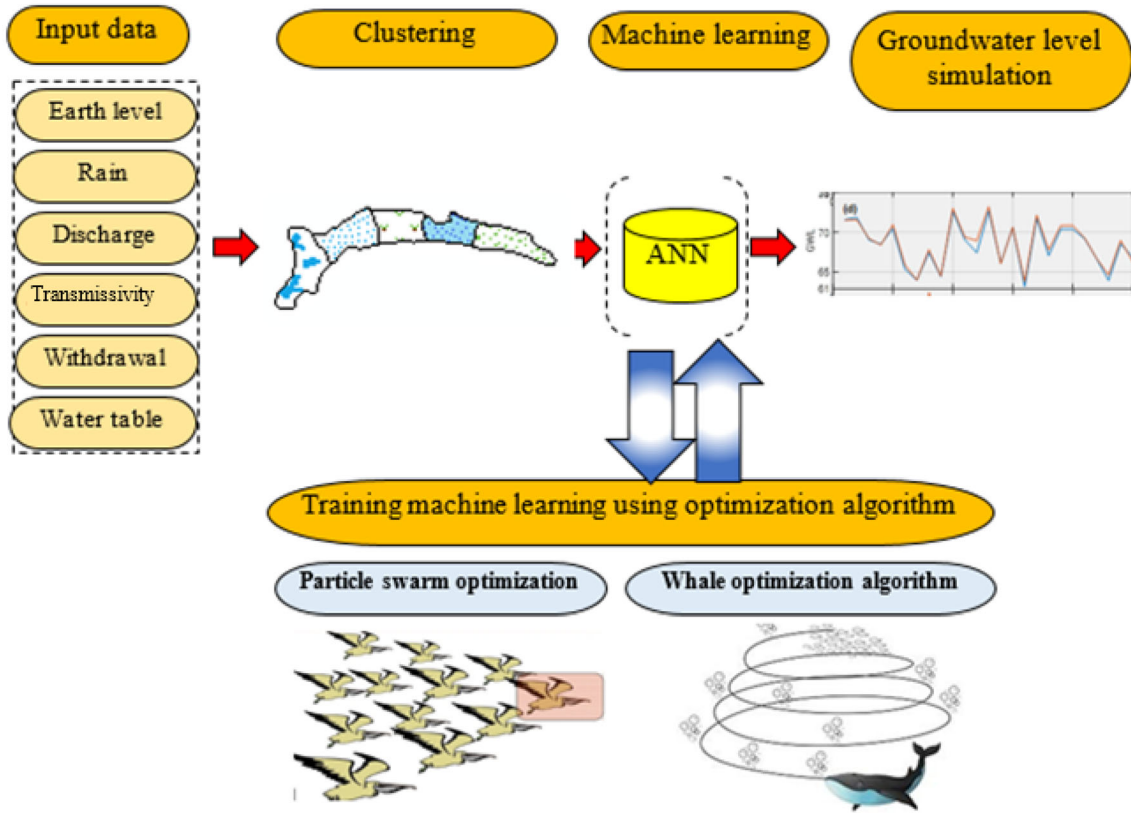


Figure 2. Flowchart of the proposed approach for simulation of GWL.

in the search space. The velocity consists of two vectors, i.e., p and p_g ; p is the best position that particle i has ever reached, while p_g is the best position that the neighborhood particle of i has ever reached. In the search for a d -dimensional space, the position of particle i is represented by a d -dimensional vector called $X_i = (X_{i1}, X_{i2}, \dots, X_{id})$. The velocity of each particle is represented by a d -dimensional velocity vector called $V_i = (V_{i1}, V_{i2}, \dots, V_{id})$. Finally, the variables move to the optimum points using Eqs. 1 and 2:

$$V_{id}^{n+1} = \omega \cdot v_{id}^n + c_1 r_1^n (p_{id}^n - x_{id}^n) + c_2 r_2^n (p_{pg}^n - x_{id}^n) \tag{1}$$

$$x_{id}^{n+1} = x_{id}^n + v_{id}^{n+1} \tag{2}$$

where ω is the shrinkage factor used for convergence rate determination, r_1 and r_2 are random numbers between 0 and 1, N is number of iterations, c_1 is the best solution obtained by a particle, and c_2 is the best solution identified by the whole population (Ken-

nedy & Eberhart, 1995). In this study, to develop the ANN-PSO algorithm, N random vectors with initial X_i position were created. ANN was then implemented with the particle positions, and the PSO objective function was used to minimize the prediction error. The particles were then moved to find better positions, and new parameters were obtained for the ANN (Anand & Suganthi, 2020; Spina, 2006). This process was repeated until the prediction error converged to a minimum.

Whale Optimization Algorithm

WOA is a nature-inspired algorithm proposed by Mirjalili and Lewis (2016). It uses the bubble-net hunting strategy of humpback whales. Each whale releases air bubbles under the sea, which create walls of rising air in the water. The krill and small fish herds inside the aerial wall, because of fear of being trapped, go to the center of the bubble circle when the whale hunts and eats a large number of

them. The whale can detect the position of the prey and thus surround the prey. However, because the search space's optimal position is unclear, it assumes that the best current answer is the adjacent prey. After determining this point, the search for other optimal points and position updates continues, which is indicated by Eqs. (3) and (4) (Mirjalili & Lewis, 2016):

$$D = \left| \vec{c} \cdot \vec{X}_t^* - \vec{X}_t \right| \quad (3)$$

$$\vec{X}_{t+1} = \vec{X}_t^* - \vec{A} \cdot \vec{D} \quad (4)$$

where t is the current iterator, C and A are the coefficient vectors, X^* is the best position vector so far, and X is the position vector. The vectors A and C are calculated, respectively, as:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \quad (5)$$

$$\vec{C} = 2 \cdot \vec{r} \quad (6)$$

where a is a vector in both exploration and exploitation phases and it is reduced from 2 to 0 per repetition. The vector r is a random vector in the range [0, 1] (Mirjalili & Lewis, 2016).

Aquifer Clustering

Clustering is a method that does not deal with the distribution of existing data and it often uses data similarity and dissimilarity criteria to sort the data (Bisht & Paul, 2013; Rokach & Maimon, 2005; Shah & Mahajan, 2012). Clustering is an unsupervised machine learning method that has many applications in engineering and science. Clustering aims to divide data into different groups based on the greater similarity within the groups and the greater dissimilarity between them.

The K -means algorithm is one of the most popular and the simplest clustering algorithms (Heil et al., 2019; Nayak et al., 2016; Zhang et al., 2013). It has been used to better manage and understand problems in water resources, water distribution systems, and water consumption management (Javadi et al., 2021; Mohammadrezapour et al., 2020). To identify similar regions in the study aquifer based on the selected criteria, namely, earth level, precipitation, water recharge, water discharge, transmissivity, and water table, the K -means clustering method was used. The purpose of K -means clustering is to min-

imize the objective function J (Dehariya et al., 2010), thus:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|X_{ij} - c_j\|^2 \quad (7)$$

where $\|X_{ij} - c_j\|^2$ is the Euclidean distance between X_{ij} and c_j ; the former is the data point, latter is the center of the cluster. The clustering procedure used here included four steps. In the first step, K initial clusters were selected randomly, and the centers of the clusters were determined individually. In the second step, each data sample was assigned to a cluster whose center had the shortest distance to data. After assigning all the data to clusters, a new point was considered the center of each cluster obtained by averaging the points belonging to each cluster. At the final step, steps 2 and 3 were repeated until no more change in the center of the clusters was observed, and the objective function was minimized.

Initial validation included deleting the unreal and out-of-range data, and selecting the number of input variables. Then, normalization was performed to normalize the data to the [0,1] range as:

$$X^* = \frac{X_i - \text{Min}(X)}{\text{Max}(X) - \text{Min}(X)} \quad (8)$$

where X^* and X_i represent normalized and original values of variable X , and Min and Max represent the lowest and highest values of variable X .

Statistical Evaluation

To evaluate the performance of the models and input patterns, four evaluation criteria, i.e., root mean squared error (RMSE) (Eq. 9), mean absolute percentage error (MAPE) (Eq. 10), Nash Sutcliffe index (NSE) (Eq. 11), and coefficient of determination (R^2) (Eq. 12) were used.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (S_o - S_p)^2}{n}} \quad (9)$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{S_o - S_p}{S_o} \right| \quad (10)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (S_p - S_o)^2}{\sum_{i=1}^n (S_o - \bar{S}_o)^2} \quad (11)$$

$$R^2 = \frac{\left[\sum_{i=1}^n (S_p - \bar{S}_p)(S_o - \bar{S}_o) \right]^2}{\sum_{i=1}^n (S_p - \bar{S}_p)^2 \sum_{i=1}^n (S_o - \bar{S}_o)^2} \quad (12)$$

where S_p and S_o are the i th simulated and observational data, respectively, \bar{S}_p and \bar{S}_o are the means of the simulated and observational data, respectively, and n is the number of samples.

RESULTS

Aquifer Clustering

Eighteen observation wells measure GWL on a monthly basis according to, among others, the locations of recharge-discharge sources, inflow and outflow, and land position. The hydrological condition of the aquifer was provided based on the analysis of these observation wells. To determine suitable clusters and their observation wells, six factors, including precipitation (P), water recharge (R), water discharge (D), transmissivity (T), water table (WT), and earth level (EL), were considered. Then, clustering was performed using the K -means method. According to this study's objective, the clustering method selected five observation wells based on the design criteria of the quantitative groundwater network. To evaluate and verify the number of clusters and observation wells selected at the center of each cluster, changes in water level per cluster and the entire aquifer in clustering conditions and without clustering conditions were compared, which showed appropriate clustering. Therefore, these five wells show the quantitative behavior of the aquifer.

According to the locations of the aquifer's recharge and discharge sources, the locations of residential areas, especially the city, groundwater inlets and outlets, land use and hydrological characteristics of the aquifer, spatial clustering of the aquifer was performed based on the Thiessen polygon network in the region. Table 1 shows the average value of the parameters considered for aquifer clustering.

The Nasrabad observation well N2 is located in the western part of the aquifer (where the groundwater discharges), and changes in GWL are affected by groundwater flows and return water of agricul-

tural lands upstream. The Sivjan observation well N6 is located in the central part of the aquifer and in agricultural lands. The Shamsabad observation well N12 is located at the central part of the Birjand aquifer, the upstream of the region's agricultural lands. The Hajiabad observation well N13 is located downstream of Birjand city; it is affected by the return water of drinking and industry sectors, and in recent years, the construction of a treatment plant has also affected the trend of changes in GWL. The Bojd observation well N17 is located in the eastern part of the aquifer; it is affected by inlet groundwater flows. Spatial clustering was performed on the surface of the Birjand aquifer (Fig. 3).

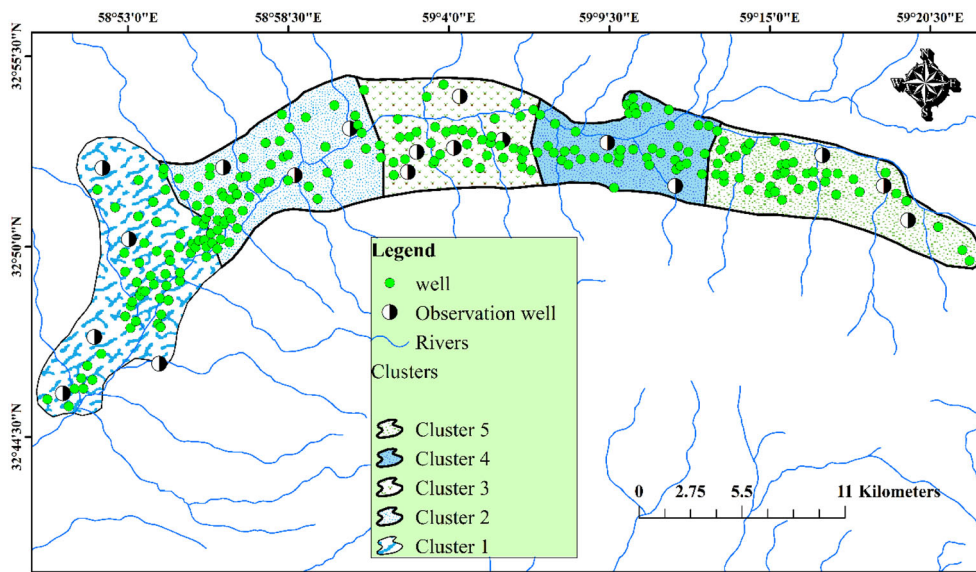
Table 2 shows the patterns developed based on different combinations of input variables. These patterns were developed based on literature review to determine the best and most cost-effective combination and to identify essential input variables among the various factors. Table 2 shows the patterns and their input variables, including GWL_{n-1} , aquifer discharge (D), aquifer recharge (R), evaporation (E), temperature (T), and precipitation (P). All patterns were implemented for each model investigated in this study, and the best combination and the most appropriate model were selected per cluster.

ANN Model

The architecture of the developed ANN was comprised of two hidden layers and an outer layer. From a range of 10 to 20 neurons, 12 proper neurons were selected for the hidden layer and one neuron was selected for the outer layer considering the number of output parameters. The sigmoid function was used for the middle transfer layer because it yielded better results compared to the hyperbolic function. The identical function was chosen as the transferring function of the outer layer. Finally, the ANN structure was trained by the Levenberg–Marquardt (LM) back-propagation algorithm. This architecture was selected from the various architectures developed for efficient prediction of GWL. The results of error evaluation criteria per cluster and per pattern are shown in Table 3. In each cluster, a different pattern was selected as the most appropriate pattern. The selected patterns had the least prediction error for the test data compared to other patterns. Besides, their error values were similar for the training and test datasets. Examina-

Table 1. Average values of aquifer parameters per cluster

| Cluster | Observation well | P (mm/year) | R (mm/year) | D (MCM/year) | T (m ² /day) | WT (3) | EL |
|---------|------------------|----------------|----------------|-----------------|----------------------------|-----------|--------|
| 1 | N2 | 150 | 0.46 | 28.2 | 800.0 | 8.80 | 1291.7 |
| 2 | N6 | 200 | 5.50 | 17.7 | 1800 | 38.8 | 1342.0 |
| 3 | N12 | 180 | 5.90 | 11.1 | 2000 | 61.6 | 1360.1 |
| 4 | N13 | 170 | 5.30 | 15.7 | 1500 | 59.4 | 1366.0 |
| 5 | N17 | 210 | 0.43 | 11.5 | 500.0 | 60.5 | 1539.8 |

**Figure 3.** Location of observation wells in clusters of the aquifer.**Table 2.** Patterns for prediction of GWL and their input variables

| Pattern | Input variables | Output variable |
|---------|--|-----------------|
| P1 | GWL _{n-1} and D | GWL |
| P2 | GWL _{n-1} and D and P | |
| P3 | GWL _{n-1} and D and R | |
| P4 | GWL _{n-1} and D and E and T | |
| P5 | GWL _{n-1} and D and R and P | |
| P6 | GWL _{n-1} and D and E and E | |
| P7 | GWL _{n-1} and D and R and T and P | |
| P8 | GWL _{n-1} and D and R and E and P | |
| P9 | GWL _{n-1} and D and R and E and T | |
| P10 | GWL _{n-1} and D and R and E and T and P | |

tion of the GWL prediction patterns shows that, in cluster 1, P5 was the selected pattern. This pattern consists of GWL at the previous month, aquifer recharge, aquifer discharge, and precipitation.

In cluster 2, P4 was selected according to the values of the error criteria. This pattern consists of four variables, namely GWL at the previous month, aquifer discharge, temperature, and evaporation. In

Table 3. Error evaluation criteria for the studied clusters using the ANN model. Values in bold represent the selected patterns, which have the overall least prediction error for the test data compared to other patterns

| Cluster | | | Pattern | | | | | | | | | |
|-----------|------|-------|---------|-------|-------|--------------|--------------|-------|-------|-------|-------|-------|
| | | | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| Cluster 1 | RMSE | Train | 0.085 | 0.09 | 0.058 | 0.03 | 0.017 | 0.021 | 0.025 | 0.028 | 0.03 | 0.03 |
| | | Test | 0.083 | 0.088 | 0.053 | 0.032 | 0.018 | 0.038 | 0.045 | 0.032 | 0.034 | 0.032 |
| | NSE | Train | 0.72 | 0.63 | 0.4 | 0.24 | 0.2 | 0.23 | 0.23 | 0.23 | 0.24 | 0.25 |
| | | Test | 0.79 | 0.82 | 0.3 | 0.25 | 0.21 | 0.26 | 0.28 | 0.31 | 0.27 | 0.26 |
| | MAPE | Train | 0.56 | 0.76 | 0.78 | 0.85 | 0.9 | 0.89 | 0.89 | 0.89 | 0.92 | 0.92 |
| | | Test | 0.52 | 0.72 | 0.81 | 0.87 | 0.87 | 0.83 | 0.85 | 0.9 | 0.89 | 0.87 |
| Cluster 2 | RMSE | Train | 0.022 | 0.022 | 0.023 | 0.021 | 0.023 | 0.023 | 0.022 | 0.022 | 0.022 | 0.022 |
| | | Test | 0.026 | 0.026 | 0.026 | 0.096 | 0.022 | 0.024 | 0.027 | 0.025 | 0.026 | 0.025 |
| | NSE | Train | 0.28 | 0.28 | 0.3 | 0.22 | 0.29 | 0.3 | 0.27 | 0.28 | 0.28 | 0.29 |
| | | Test | 0.34 | 0.31 | 0.3 | 0.24 | 0.27 | 0.25 | 0.33 | 0.3 | 0.34 | 0.31 |
| | MAPE | Train | 0.95 | 0.95 | 0.94 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 |
| | | Test | 0.93 | 0.94 | 0.95 | 0.93 | 0.94 | 0.95 | 0.93 | 0.93 | 0.94 | 0.94 |
| Cluster 3 | RMSE | Train | 0.017 | 0.018 | 0.018 | 0.018 | 0.018 | 0.018 | 0.017 | 0.017 | 0.018 | 0.018 |
| | | Test | 0.015 | 0.015 | 0.016 | 0.015 | 0.008 | 0.015 | 0.016 | 0.015 | 0.014 | 0.014 |
| | NSE | Train | 0.180 | 0.200 | 0.200 | 0.18 | 0.25 | 0.2 | 0.16 | 0.18 | 0.2 | 0.2 |
| | | Test | 0.260 | 0.220 | 0.450 | 0.24 | 0.30 | 0.19 | 0.32 | 0.22 | 0.3 | 0.24 |
| | MAPE | Train | 0.960 | 0.960 | 0.960 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 |
| | | Test | 0.960 | 0.960 | 0.940 | 0.96 | 0.96 | 0.96 | 0.95 | 0.96 | 0.96 | 0.96 |
| Cluster 4 | RMSE | Train | 0.009 | 0.011 | 0.011 | 0.011 | 0.01 | 0.011 | 0.01 | 0.009 | 0.01 | 0.009 |
| | | Test | 0.022 | 0.018 | 0.023 | 0.018 | 0.019 | 0.018 | 0.02 | 0.023 | 0.019 | 0.022 |
| | NSE | Train | 0.67 | 0.5 | 0.5 | 0.59 | 0.58 | 0.59 | 0.58 | 0.56 | 0.58 | 0.58 |
| | | Test | 0.85 | 0.69 | 0.65 | 0.61 | 0.51 | 0.58 | 0.51 | 0.56 | 0.59 | 0.56 |
| | MAPE | Train | 0.93 | 0.91 | 0.91 | 0.92 | 0.92 | 0.9 | 0.92 | 0.94 | 0.93 | 0.93 |
| | | Test | 0.9 | 0.94 | 0.64 | 0.93 | 0.94 | 0.96 | 0.93 | 0.88 | 0.91 | 0.85 |
| Cluster 5 | RMSE | Train | 0.011 | 0.008 | 0.005 | 0.006 | 0.005 | 0.004 | 0.006 | 0.007 | 0.005 | 0.005 |
| | | Test | 0.013 | 0.014 | 0.006 | 0.007 | 0.006 | 0.007 | 0.008 | 0.007 | 0.007 | 0.006 |
| | NSE | Train | 0.32 | 0.12 | 0.08 | 0.09 | 0.06 | 0.05 | 0.08 | 0.08 | 0.06 | 0.05 |
| | | Test | 0.36 | 0.22 | 0.09 | 0.09 | 0.08 | 0.09 | 0.2 | 0.07 | 0.09 | 0.07 |
| | MAPE | Train | 0.85 | 0.93 | 0.9 | 0.81 | 0.92 | 0.95 | 0.91 | 0.89 | 0.92 | 0.94 |
| | | Test | 0.74 | 0.84 | 0.88 | 0.8 | 0.88 | 0.85 | 0.85 | 0.85 | 0.77 | 0.87 |

clusters 3 to 5, similar to cluster 1, P5 was the selected pattern. In cluster 3, P9 also had results similar to those of P5. However, because the P5 pattern had fewer input variables, it was more suitable than the P9 pattern. An essential point in the ANN model is the importance of the three parameters, namely GWL at the previous month, precipitation, and aquifer discharge, which are required in all the selected patterns for simulation. In addition, the spatial evaluation indicates that observation wells in the aquifer's central parts required more variables for efficient prediction.

Time-series plots for the observed and simulated values of the test data for the selected patterns are depicted in Fig. 4. In general, although the ANN was able to detect correctly the trend of changes in GWL, it has performed poorly in some steps. Figure 5 shows a comparison of observed and simulated

data. The goodness of fit is defined relative to the regression line; the closer the observed and simulated values are to each other, the more they lie on the regression line, and the more the accuracy of the model's performance. Figure 5 shows the density and dispersion of test data per selected pattern. According to the graphs, it is clear that all models have a relatively good density compared to the regression line. However, better results, or in other words, higher R^2 values, can be obtained in some clusters. R^2 values of the observational and simulation values in selected patterns vary from 0.84 to 0.99. The lowest R^2 value belonged to the first cluster, which was equal to 0.84. The hybrid ANN-evolutionary optimization methods can help to improve the performance of the ANN model in obtaining more reliable results (see next section).

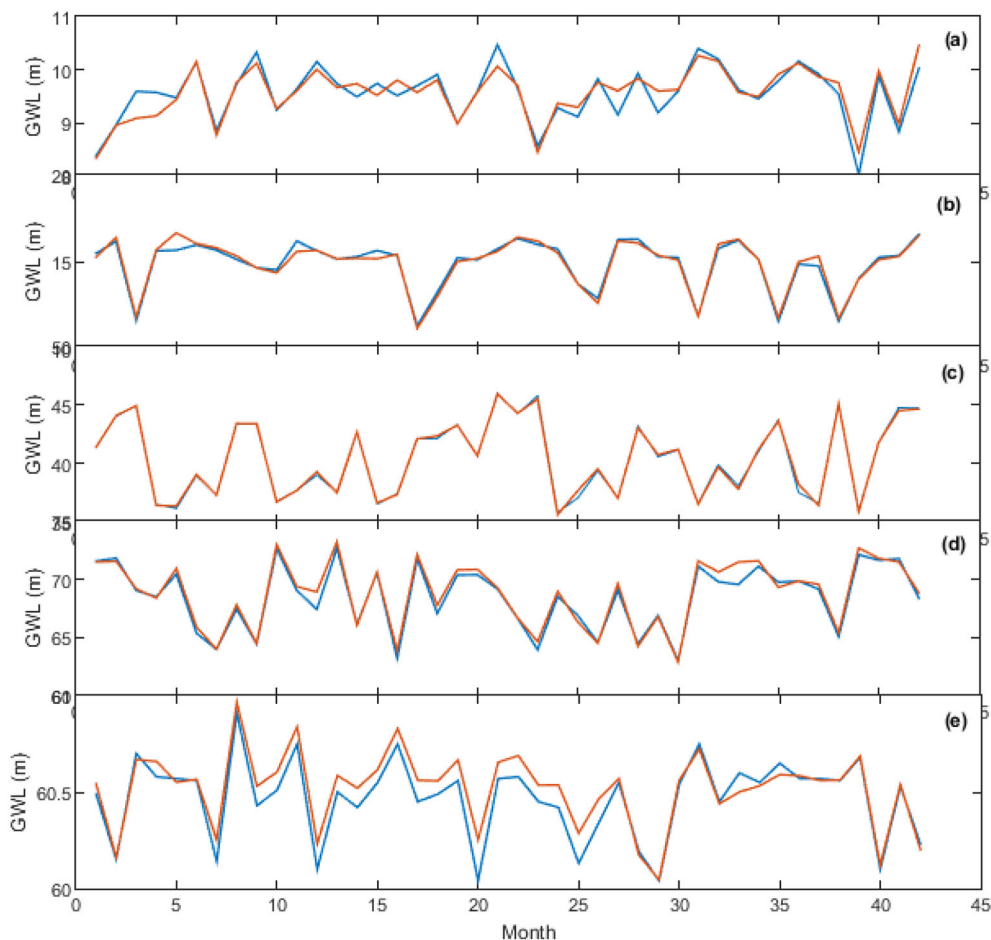


Figure 4. Time series of observed and simulated test data per cluster by the selected pattern using ANN model, (a) cluster 1, (b) cluster 2, (c) cluster 3, (d) cluster 4, (e) cluster 5. Blue lines = observed values. Red lines = simulated values.

Hybrid Machine Learning Models: Evolutionary Algorithms

The patterns were implemented using the ANN-PSO and ANN-WOA models. The initial population and the maximum iteration number were considered equal to 30 and 1500, respectively. By increasing or decreasing the population, the optimization accuracy did not improve. Furthermore, after 1500 repetitions, no change in the optimization results was observed.

Table 4 shows the results of the error evaluation criteria of the ANN-PSO and ANN-WOA models. Because the approached to reach the optimal points of both algorithms are different, a suitable pattern was selected per cluster and per model. Therefore, a maximum of two patterns was selected per cluster. Patterns P5 and P7 were selected for cluster 1. P5

was selected for the ANN-PSO algorithm, while P7 was selected for the ANN-WOA. P5 included GWL at the previous month, aquifer discharge, aquifer recharge, and precipitation; its RMSE, MAPE, and NSE were 0.01 m, 0.13 m, and 0.97, respectively. The pattern closest to P5 was P6, which had similar results to this pattern, but it was not considered because of its high prediction error on training data. P7 included the parameters of P6 plus the temperature. In this pattern, the RMSE, MAPE, and NSE were 0.01 m, 0.95 m, and 0.12, respectively. P3, which included three parameters of GWL at the previous month, aquifer discharge, and aquifer recharge, was the selected model for both hybrid models. The best ANN-PSO model had RMSE, MAPE, and NSE values of 0.006 m, 0.99, and 0.12 m for the test data, respectively. In cluster 4, the P5 and P8 patterns resulted in the highest performance

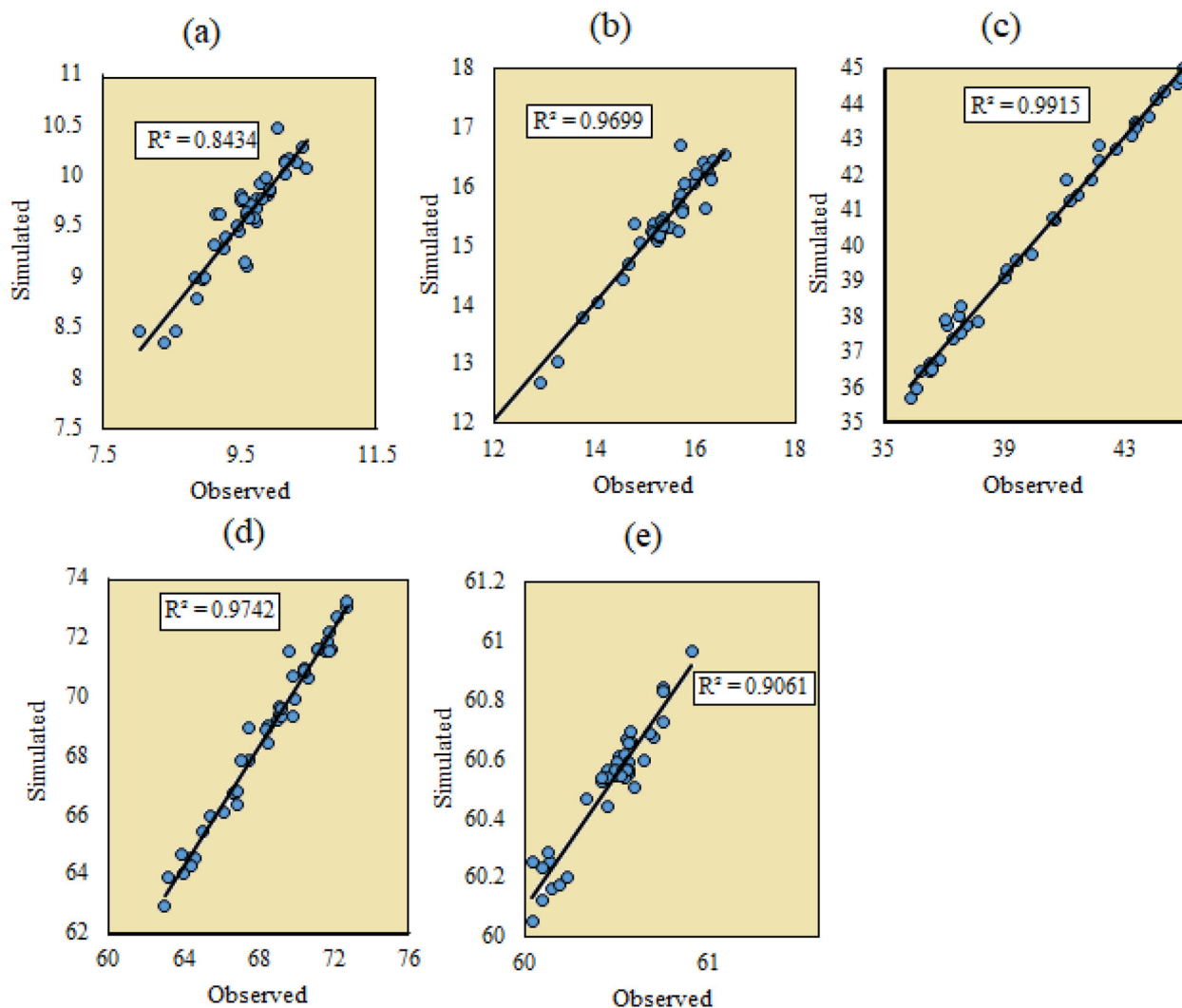


Figure 5. Scatter points of ANN results for the selected pattern per cluster (test data): (a) cluster 1; (b) cluster 2; (c) cluster 3; (d) cluster 4; and (e) cluster 5.

for the ANN-PSO and ANN-WOA models, respectively. P8 included the GWL at the previous month, aquifer discharge, aquifer recharge, evaporation, and precipitation. In pattern P5, RMSE, MAPE, and NSE criteria were equal to 0.003 m, 0.99, and 0.21 m, respectively. For the P8 pattern, these values were 0.004 m, 0.98, and 0.4 m, respectively.

It is observed that the PSO algorithm performed better than the WOA algorithm. It resulted in better performance with fewer inputs and was more accurate than the P8 pattern with more inputs. Of course, in both models, other patterns also had good evaluation results, and this shows that both

algorithms have a high ability to train the ANN model. P9 and P5 were the selected patterns of the ANN-PSO and ANN-WOA models in cluster 5. Error evaluation criteria for P9 were 0.001 m, 0.96, and 0.04 for the RMSE, MAPE, and NSE, respectively. These values for the P5 pattern were 0.001 m, 0.98, and 0.05 m, respectively. In this cluster, in contrast with the fourth cluster, the WOA algorithm performed better than the PSO algorithm since the WOA has provided better results using lower input variables. In this model, except for the first two patterns, which included two input variables, other patterns had appropriate results close to the selected pattern. In this cluster, the combination of all input

Table 4. Error evaluation criteria for the ANN-PSO and ANN-WOA models. The best model performance is shown in bold

| Model | Cluster | | | Pattern | | | | | | | | | |
|-----------|-----------|-------|-------|---------|-------|--------------|-------------|--------------|-------|-------------|--------------|--------------|-------|
| | | | | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| ANN-PSO | Cluster 1 | RMSE | Train | 0.041 | 0.008 | 0.007 | 0.009 | 0.01 | 0.051 | 0.015 | 0.01 | 0.009 | 0.008 |
| | | | Test | 0.05 | 0.013 | 0.013 | 0.013 | 0.01 | 0.009 | 0.012 | 0.011 | 0.011 | 0.011 |
| | | NSE | Train | 0.44 | 0.98 | 0.98 | 0.92 | 0.98 | 0.92 | 0.88 | 0.92 | 0.97 | 0.97 |
| | | | Test | 0.44 | 0.9 | 0.86 | 0.92 | 0.97 | 0.97 | 0.91 | 0.9 | 0.87 | 0.87 |
| | | MAPE | Train | 0.54 | 0.11 | 0.1 | 0.13 | 0.12 | 0.14 | 0.2 | 0.14 | 0.13 | 0.13 |
| | | | Test | 0.59 | 0.18 | 0.18 | 0.18 | 0.13 | 0.12 | 0.15 | 0.14 | 0.16 | 0.16 |
| | Cluster 2 | RMSE | Train | 0.018 | 0.02 | 0.017 | 0.018 | 0.011 | 0.014 | 0.012 | 0.011 | 0.011 | 0.013 |
| | | | Test | 0.02 | 0.02 | 0.02 | 0.02 | 0.012 | 0.016 | 0.018 | 0.018 | 0.011 | 0.016 |
| | | NSE | Train | 0.96 | 0.97 | 0.96 | 0.98 | 0.99 | 0.98 | 0.98 | 0.99 | 0.98 | 0.98 |
| | | | Test | 0.95 | 0.96 | 0.94 | 0.97 | 0.98 | 0.98 | 0.98 | 0.98 | 0.99 | 0.98 |
| | | MAPE | Train | 0.28 | 0.25 | 0.26 | 0.21 | 0.2 | 0.3 | 0.21 | 0.21 | 0.23 | 0.26 |
| | | | Test | 0.28 | 0.27 | 0.25 | 0.2 | 0.25 | 0.33 | 0.35 | 0.32 | 0.21 | 0.34 |
| | Cluster 3 | RMSE | Train | 0.015 | 0.018 | 0.006 | 0.008 | 0.008 | 0.009 | 0.007 | 0.009 | 0.009 | 0.01 |
| | | | Test | 0.015 | 0.015 | 0.006 | 0.009 | 0.008 | 0.01 | 0.008 | 0.009 | 0.011 | 0.011 |
| | | NSE | Train | 0.97 | 0.96 | 0.99 | 0.98 | 0.98 | 0.98 | 0.97 | 0.96 | 0.97 | 0.96 |
| | | | Test | 0.96 | 0.96 | 0.99 | 0.98 | 0.98 | 0.97 | 0.98 | 0.97 | 0.96 | 0.96 |
| | | MAPE | Train | 0.16 | 0.2 | 0.10 | 0.12 | 0.15 | 0.13 | 0.12 | 0.14 | 0.17 | 0.18 |
| | | | Test | 0.23 | 0.22 | 0.12 | 0.14 | 0.12 | 0.15 | 0.13 | 0.17 | 0.17 | 0.16 |
| | Cluster 4 | RMSE | Train | 0.01 | 0.005 | 0.003 | 0.004 | 0.002 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 |
| | | | Test | 0.012 | 0.005 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 | 0.003 | 0.004 | 0.004 |
| NSE | | Train | 0.85 | 0.96 | 0.99 | 0.97 | 0.99 | 0.99 | 0.99 | 0.98 | 0.99 | 0.98 | |
| | | Test | 0.82 | 0.97 | 0.99 | 0.98 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.96 | |
| MAPE | | Train | 0.65 | 0.32 | 0.3 | 0.33 | 0.21 | 0.25 | 0.24 | 0.26 | 0.28 | 0.27 | |
| | | Test | 0.78 | 0.36 | 0.29 | 0.31 | 0.30 | 0.33 | 0.33 | 0.32 | 0.36 | 0.35 | |
| Cluster 5 | RMSE | Train | 0.008 | 0.008 | 0.006 | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.002 | |
| | | Test | 0.01 | 0.006 | 0.005 | 0.003 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.002 | |
| | NSE | Train | 0.85 | 0.89 | 0.94 | 0.9 | 0.97 | 0.94 | 0.95 | 0.92 | 0.95 | 0.9 | |
| | | Test | 0.86 | 0.91 | 0.92 | 0.93 | 0.86 | 0.85 | 0.88 | 0.93 | 0.96 | 0.91 | |
| | MAPE | Train | 0.16 | 0.12 | 0.06 | 0.07 | 0.04 | 0.05 | 0.04 | 0.06 | 0.05 | 0.06 | |
| | | Test | 0.22 | 0.12 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.06 | 0.04 | 0.08 | |
| ANN-WOA | Cluster 1 | RMSE | Train | 0.03 | 0.01 | 0.01 | 0.018 | 0.011 | 0.011 | 0.01 | 0.013 | 0.01 | 0.011 |
| | | | Test | 0.035 | 0.014 | 0.014 | 0.019 | 0.01 | 0.011 | 0.01 | 0.01 | 0.013 | 0.015 |
| | | NSE | Train | 0.72 | 0.98 | 0.94 | 0.8 | 0.89 | 0.96 | 0.94 | 0.89 | 0.97 | 0.93 |
| | | | Test | 0.63 | 0.89 | 0.92 | 0.73 | 0.96 | 0.95 | 0.95 | 0.93 | 0.9 | 0.9 |
| | | MAPE | Train | 0.45 | 0.13 | 0.14 | 0.24 | 0.15 | 0.15 | 0.14 | 0.13 | 0.13 | 0.16 |
| | | | Test | 0.53 | 0.19 | 0.19 | 0.22 | 0.15 | 0.14 | 0.12 | 0.18 | 0.18 | 0.2 |
| | Cluster 2 | RMSE | Train | 0.019 | 0.02 | 0.019 | 0.018 | 0.012 | 0.015 | 0.013 | 0.011 | 0.013 | 0.012 |
| | | | Test | 0.021 | 0.022 | 0.022 | 0.019 | 0.011 | 0.014 | 0.017 | 0.018 | 0.015 | 0.017 |
| | | NSE | Train | 0.96 | 0.97 | 0.98 | 0.98 | 0.99 | 0.99 | 0.89 | 0.99 | 0.99 | 0.98 |
| | | | Test | 0.96 | 0.95 | 0.96 | 0.97 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.98 |
| | | MAPE | Train | 0.27 | 0.23 | 0.24 | 0.22 | 0.23 | 0.3 | 0.17 | 0.21 | 0.22 | 0.24 |
| | | | Test | 0.3 | 0.26 | 0.25 | 0.23 | 0.22 | 0.27 | 0.35 | 0.3 | 0.28 | 0.33 |
| | Cluster 3 | RMSE | Train | 0.012 | 0.011 | 0.011 | 0.01 | 0.008 | 0.008 | 0.008 | 0.012 | 0.01 | 0.009 |
| | | | Test | 0.011 | 0.013 | 0.01 | 0.01 | 0.009 | 0.01 | 0.01 | 0.011 | 0.013 | 0.012 |
| | | NSE | Train | 0.97 | 0.97 | 0.96 | 0.98 | 0.99 | 0.98 | 0.97 | 0.98 | 0.98 | 0.99 |
| | | | Test | 0.97 | 0.96 | 0.95 | 0.97 | 0.99 | 0.98 | 0.98 | 0.97 | 0.96 | 0.97 |
| | | MAPE | Train | 0.18 | 0.17 | 0.18 | 0.18 | 0.15 | 0.2 | 0.16 | 0.18 | 0.2 | 0.2 |
| | | | Test | 0.21 | 0.2 | 0.22 | 0.2 | 0.14 | 0.19 | 0.32 | 0.22 | 0.3 | 0.24 |
| | Cluster 4 | RMSE | Train | 0.008 | 0.009 | 0.008 | 0.008 | 0.005 | 0.008 | 0.006 | 0.005 | 0.005 | 0.008 |
| | | | Test | 0.012 | 0.008 | 0.007 | 0.01 | 0.004 | 0.01 | 0.006 | 0.004 | 0.005 | 0.006 |
| NSE | | Train | 0.92 | 0.95 | 0.95 | 0.96 | 0.98 | 0.94 | 0.97 | 0.98 | 0.97 | 0.93 | |
| | | Test | 0.9 | 0.97 | 0.92 | 0.93 | 0.97 | 0.92 | 0.97 | 0.98 | 0.96 | 0.95 | |
| MAPE | | Train | 0.98 | 0.47 | 0.46 | 0.51 | 0.45 | 0.7 | 0.51 | 0.43 | 0.48 | 0.77 | |
| | | Test | 0.88 | 0.53 | 0.52 | 0.58 | 0.48 | 0.77 | 0.5 | 0.4 | 0.52 | 0.68 | |

Table 4. continued

| Model | Cluster | | Pattern | | | | | | | | | | |
|-----------|---------|-------|---------|-------|-------|-------|--------------|-------|-------|-------|-------|-------|-------|
| | | | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | |
| Cluster 5 | RMSE | Train | 0.009 | 0.008 | 0.005 | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.002 |
| | | Test | 0.01 | 0.006 | 0.003 | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.003 |
| | NSE | Train | 0.88 | 0.89 | 0.99 | 0.92 | 0.98 | 0.95 | 0.94 | 0.92 | 0.93 | 0.93 | 0.91 |
| | | Test | 0.85 | 0.91 | 0.99 | 0.93 | 0.98 | 0.87 | 0.9 | 0.93 | 0.96 | 0.96 | 0.89 |
| | MAPE | Train | 0.12 | 0.09 | 0.04 | 0.07 | 0.05 | 0.05 | 0.05 | 0.06 | 0.07 | 0.07 | 0.08 |
| | | Test | 0.25 | 0.07 | 0.07 | 0.08 | 0.05 | 0.07 | 0.06 | 0.06 | 0.06 | 0.06 | 0.09 |

variables did not improve the results compared to the four input variables.

It can be said that for such aquifers, no more than four input variables are required for the prediction of the GWL, and selecting the proper algorithm results in more efficient performance. In addition, the use of two input variables cannot accurately detect changes in the GWL. The relationship between the input variables and the changes in GWL is more complicated than that can be detected by two input variables. Patterns with three inputs, if the correct variables are selected, can result in promising predictions. For example, in cluster 3, the pattern with three inputs of the GWL at the previous month, aquifer discharge, and aquifer recharge can predict the GWL of the current month. It can be said that in each cluster, a suitable model and a suitable pattern should be proposed to predict the GWL. Finally, if appropriate algorithms are used, it may not be necessary to use different inputs to predict the GWL, which can reduce the cost of data collection, which has economic advantages.

The time series of the observational and simulation test data for the selected patterns and both hybrid models are depicted in Figure 6. It is observed that there is an acceptable correlation between observed and simulated values. Since the training and test data are randomly selected, the test data values are different for each cluster and the two hybrid models. Therefore, as can be seen in the diagrams, it can be concluded that there is acceptable accuracy in predicting the GWL using ANN-PSO and ANN-WOA models.

Figure 7 shows the scatter point of the observed and simulated values. A regression line is fitted to the data. The closer the observed and simulated values to each other, the more they lie on the $y = x$ line, which shows the accuracy of the model's performance. This figure shows the density and disper-

sion of test data for each of the selected patterns. If there are scattered points relative to the regression line, it indicates that the model cannot correctly predict the model output. It is clear that all models had a good density relative to the regression line, and this result, along with other results, shows the models' appropriate accuracy. R^2 values of the observational and simulation values in selected patterns vary from 0.90 to 0.99. The lowest R^2 value belonged to the first cluster and the ANN-PSO model, and the highest value belonged to cluster 4 and the SNN-PSO model. R^2 values and the fitted regression line do not solely indicate the performance of the model. However, in addition to other error evaluation criteria, they reveal the efficiency of a prediction model. However, all the selected patterns have acceptable R^2 and data density relative to the regression line.

The selected patterns in each cluster and for each model had acceptable accuracy. By using the selected pattern in each cluster and with the desired number of inputs, it is possible to predict the GWL. However, using the appropriate learning model is valuable to obtain an efficient approach. Using this approach, the additional costs of data collection are reduced, and unnecessary models are not implemented for some clusters. In the following, the appropriate model and pattern for each cluster are determined.

Evaluation of Selected Model

In this section, the appropriate pattern and model are selected for each cluster (Fig. 8). In the previous sections, three suitable patterns were selected for each cluster, and three different models were used to implement each selected pattern. Since the ANN had lower accuracy than the hybrid mod-

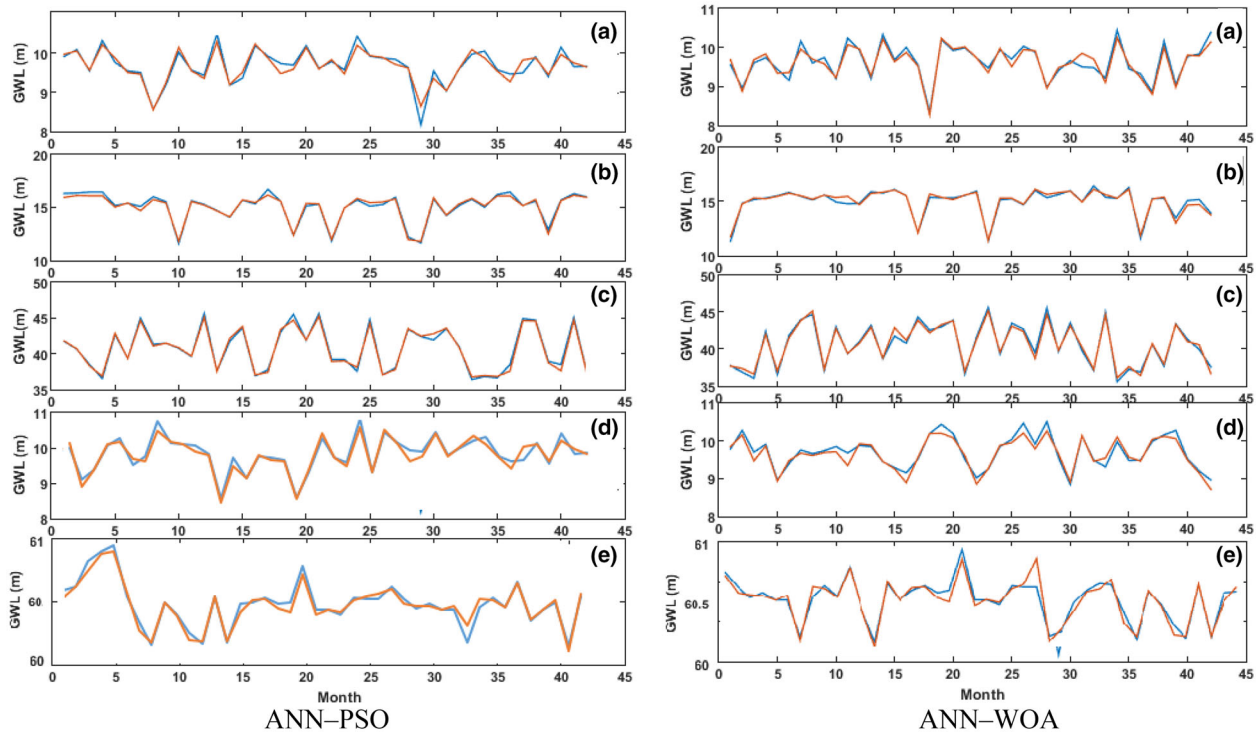


Figure 6. Time series of the observed and simulated test data per cluster by the selected pattern using the ANN-PSO and ANN-WOA models, (a) cluster 1, (b) cluster 2, (c) cluster 3, (d) cluster 4, (e) cluster 5. Blue lines: observed values and red lines: simulated values.

els, it will not be used hereafter as a proposed model. Furthermore, among the various selected patterns, patterns are selected with the appropriate accuracy and the lowest input variables. This choice helps managers reduce the uncertainty and cost of data collection and analysis because each input variable increases the costs. Therefore, the proposed approach in this study can reduce costs and uncertainty.

For cluster 1, the selected patterns were P5 for ANN-PSO and P7 for ANN-WOA. Because both patterns' accuracies were very close and because P5 required a fewer number of input variables, P5 and ANN-PSO can be introduced as the best pattern and model, respectively, for cluster 1. Among the P9 and P5 patterns for the second cluster, P5 is the selected pattern due to its good accuracy and fewer input parameters. This pattern is performed by the ANN-WOA model. In the third cluster, P3 was selected for both models, and since the ANN-PSO results were more accurate than the ANN-WOA, it was selected as the appropriate model for the cluster. In the fourth cluster, P5 and P8 were the superior patterns. P5 is preferred due to its better

performance and a lower number of input parameters. This template is implemented with the ANN-PSO model. For the fifth cluster, P5 and ANN-WOA were selected as the most suitable pattern and model, respectively. Therefore, patterns and the models can perform differently for each cluster in the prediction of the GWL. According to the results, all selected patterns have a maximum of four input variables.

P5 was the most suitable pattern for all clusters except the third cluster, which shows that it is not possible to define a single pattern and a single model for all the clusters. The results also show that it is impossible to determine an algorithm superior to other algorithms for the entire aquifer. It can be said that by having the GWL at the previous month, aquifer discharge, aquifer recharge, and precipitation, it is possible to predict the GWL in each cluster with appropriate accuracy, and there is no need to have temperature and evaporation information. In addition, both hybrid models are suitable to improve the prediction performance of the ANN in the study.

After selecting the appropriate patterns and models for each cluster, GWL values were predicted

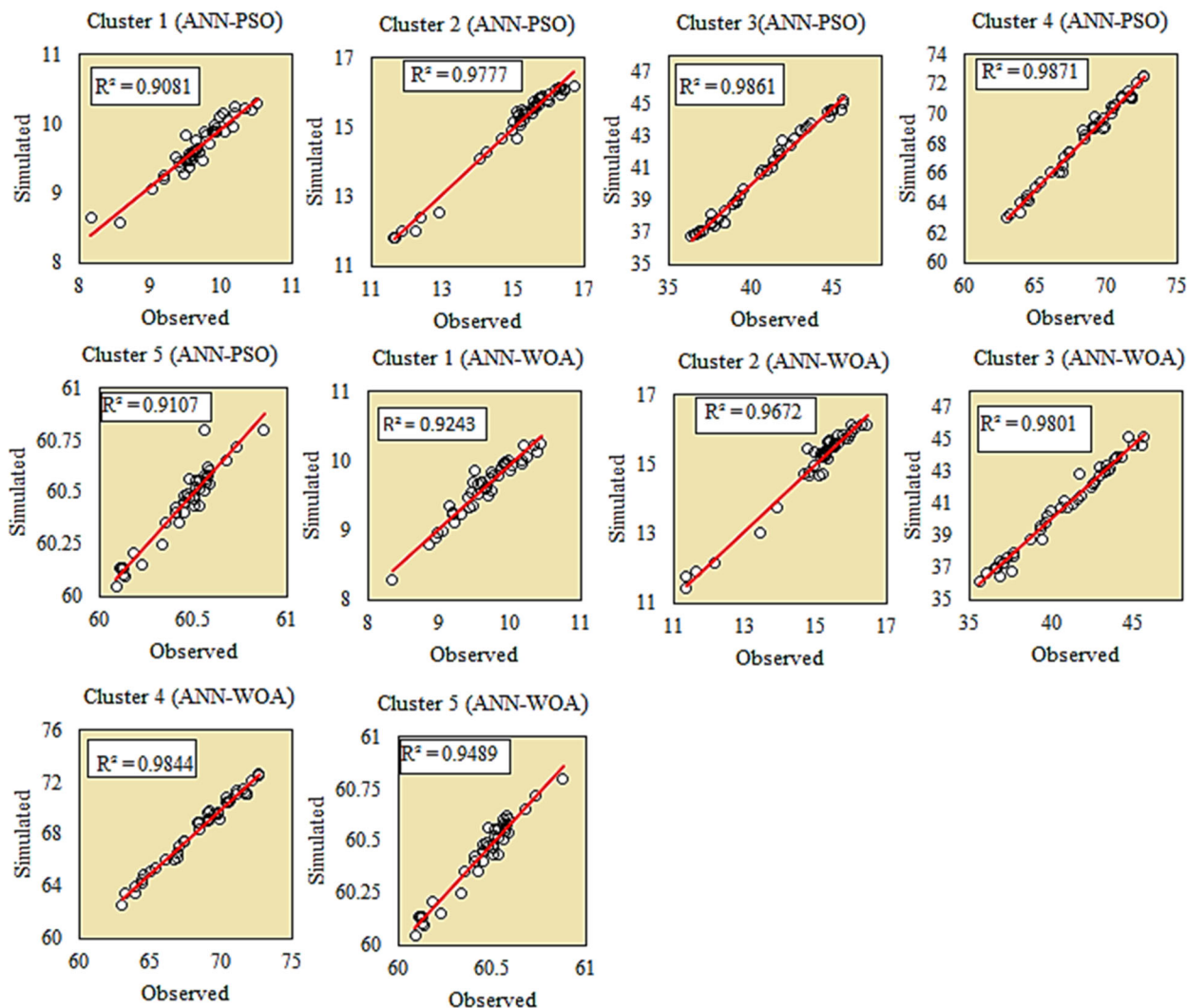


Figure 7. Scatter points of the ANN-PSO and ANN-WOA results for the selected pattern per cluster (test data): (a) cluster 1; (b) cluster 2; (c) cluster 3; (d) cluster 4; and (e) cluster 5.

for each cluster. Figure 9 shows a graph of observational and simulation data for the entire period. A good density is observed between the data and the regression line. Undesirable scattered points are not observed in the graphs, and the observed and simulated values are close to each other. R^2 of the diagrams varies from 0.92 to 0.99. The highest R^2 (0.99) was obtained for the fourth cluster. In this cluster, the observed and simulated data are close to each other, almost on the regression line.

Taylor diagram was also used to evaluate better the results obtained from selected patterns and models (Fig. 10). In this diagram, the horizontal and

vertical axes represent the standard deviation, and the arc shows the correlation coefficient. Arcs inside the diagram are used to represent the RMSD. In this diagram, the closer the models' predicted results to the observed values, the higher the correlation coefficient. For each cluster, Taylor diagram is plotted separately. In all diagrams, a remarkable correlation coefficient was observed between the observed and simulated data. The closest values of the simulated values to the observed values are observed in the third, fourth, and fifth clusters. The correlation coefficient of all clusters is more than 0.97. RMSD values for all the clusters are less than

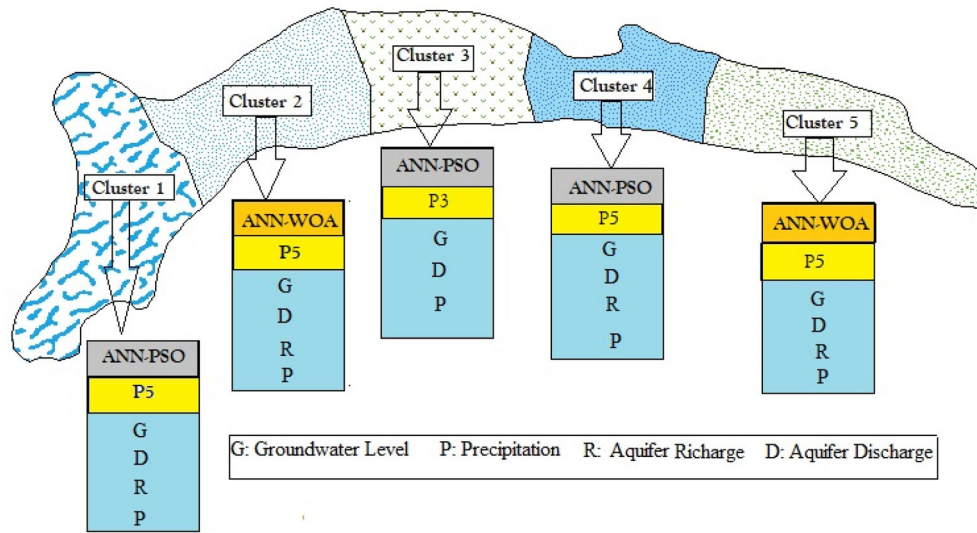


Figure 8. Proposed GWL simulation model per cluster in the study aquifer.

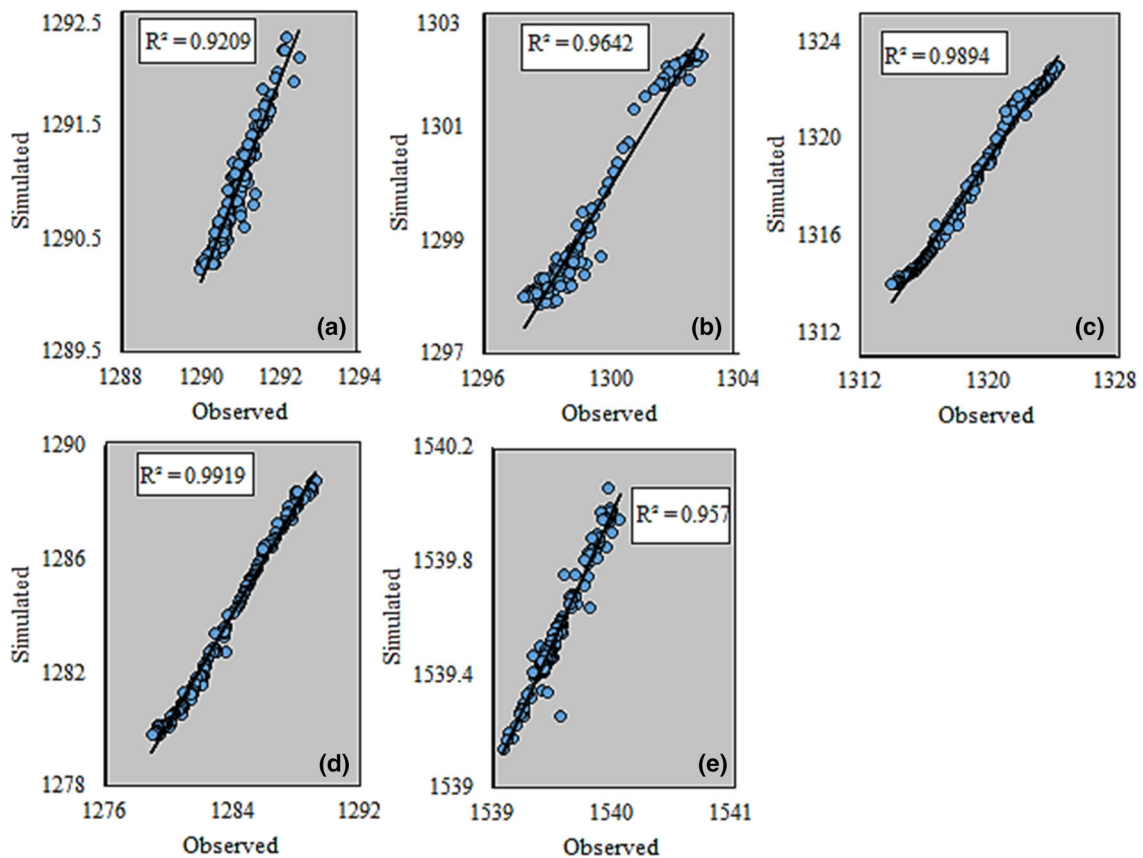


Figure 9. Scatter points of observed and simulated data by the proposed model for (a) cluster 1, (b) cluster 2, (c) cluster 3, (d) cluster 4, (e) cluster 5.

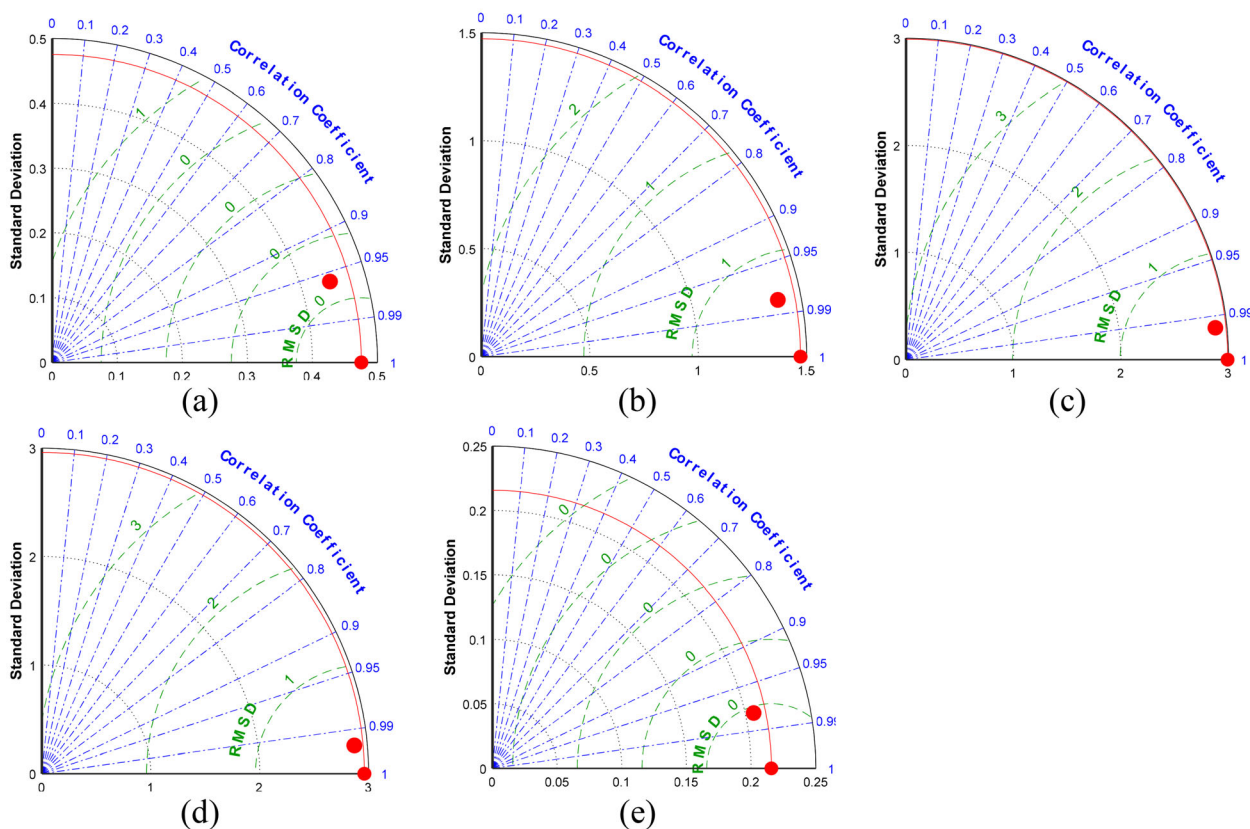


Figure 10. Taylor diagrams for (a) cluster 1, (b) cluster 2, (c) cluster 3, (d) cluster 4, (e) cluster 5.

0.50 m, which indicates that all clusters have high accuracy.

Finally, the time series of the observed and simulated values for each cluster using the selected pattern and model is depicted in Figure 11. The diagrams confirm the acceptable performance of the selected models and patterns. In the time series, the simulated and observed values are very close to each other. In some months, when GWL changes are significant, the models correctly detect the changes and exert perfect accuracy in the simulation. For example, in the fifth cluster, in the steps of 65 to 75, sudden changes in GWL have been correctly predicted by the model. Of course, other examples of sudden upward and downward changes can be observed in Figure 11. The figure shows that the GWL has a downward trend in two clusters, and the Groundwater drawdown is relatively steep, so a decline of ca. 10 m has occurred in both clusters during the study period. In the rest of the clusters, the groundwater drawdown slightly, and the decline values in these clusters vary from ca. 1 to 5 m.

Therefore, various approaches should be taken to properly plan and apply management strategies and scenarios for different aquifer regions.

DISCUSSION

The different performances of the simulation-optimization models due to their different responses toward input patterns show that these models are strongly data-driven, and the essential factor that improves the results is the alignment of changes in input variables with the output variable. However, considering the use of the clustering approach to reduce the number of inputs based on spatial characteristics, the minimum amount of data at the aquifer level was used in this study to simulate GWL. Thus, increasing the number of input variables not only increases the models' performance but also reduces their performance because of redundant data. The results also showed that simultaneous consideration of precipitation and

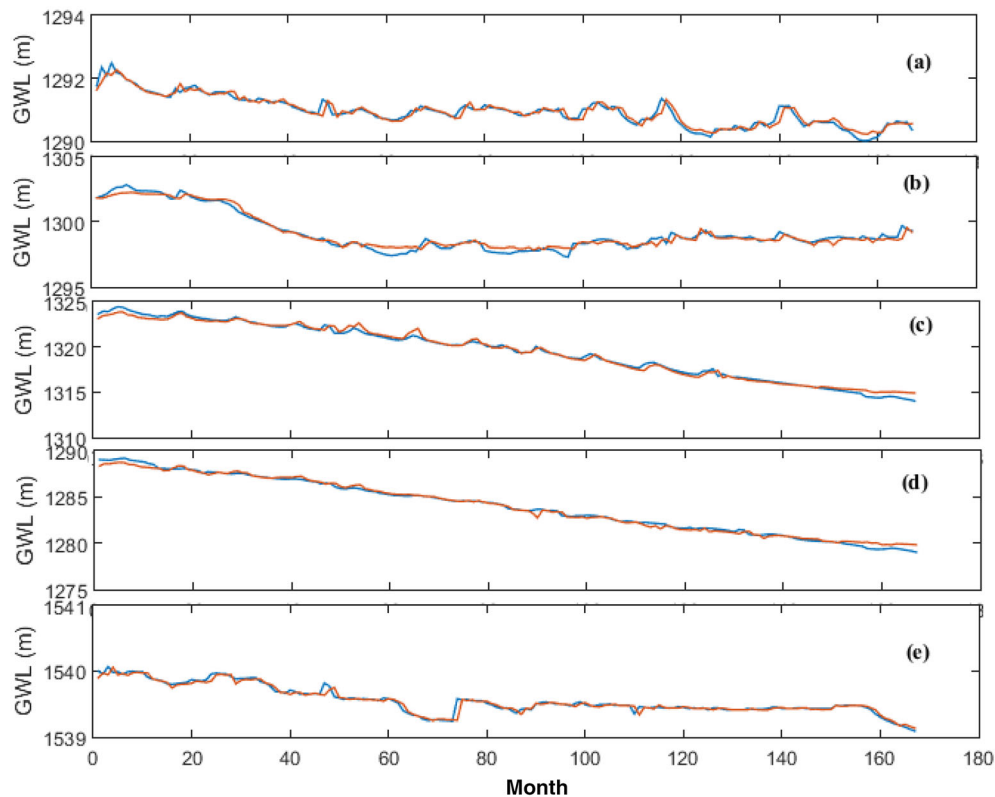


Figure 11. Observed and predicted values of GWL per cluster: (a) cluster 1; (b) cluster 2; (c) cluster 3; (d) cluster 4; and (e) cluster 5. Blue lines: observed values. Red lines: simulated values.

surface recharge as input variables improved the GWL simulation results. This indicates that aquifer recharge, caused by infiltration of precipitation, surface flows, and return water from consumption do not correlate with precipitation at the aquifer surface. In this study, unlike many studies that consider precipitation and recharge as a single input, these two variables were considered independently as inputs to the simulation model, which improved the model performance.

According to the selected patterns that were input to the simulation model, the results indicated that temperature and evaporation did not affect the model performance, even though evaporation from groundwater is remarkable at the study aquifer. Considering the importance of input and output variables of water balance and their effects on GWL, patterns including both recharge and discharge resulted in acceptable performances. Patterns P3 and P5, which exerted high performances, included the three variables of precipitation, aquifer recharge, and aquifer discharge. In addition, considering sim-

ilar regions in terms of aquifer characteristics and hydrogeological changes improved the simulation results during clustering. This improvement, which was the result of aquifer clustering along with utilizing management patterns, was effective in GWL simulation.

The application of optimization algorithms in improving the performance of ANN in simulating GWL of the aquifer was another finding of this study. However, the performance of these advanced models should be evaluated carefully before utilization because each optimization algorithm can have its own strengths and limitations in developing hybrid models. In general, the evaluation of various algorithms to optimize the structure of neuron-based learning methods is suggested to develop an accurate hybrid model for each portion of the aquifer with its own characteristics. In this regard, the results of the suggested hybrid models have shown that both algorithms are well-capable of GWL simulation. Although the PSO algorithm is more dated than the WOA algorithm, it provided results that are

more favorable so that the hybrid ANN–PSO model was the chosen model with three clusters out of five available clusters. This demonstrates that it has high efficiency in solving optimization problems.

While many algorithms have been proposed to solve optimization problems, the main question remains as to which optimization algorithm can provide the most accurate results. Therefore, it is of great importance to use different algorithms to check for the possibility of accuracy improvement. Applying such algorithms in other fields of science has also shown that they can be considered an appropriate approach to improve many machine learning models such as ANN and ANFIS (e.g., Milan et al., 2021).

CONCLUSIONS

Analyzing the quantitative status of an aquifer, this research demonstrated a novel approach that combines clustering, simulation, and optimization concepts into a single methodological framework for GWL simulation. A combination of various input parameters, such as monthly groundwater level, temperature, evaporation, precipitation, aquifer discharge, and aquifer recharge, was used to simulate GWL changes, among which evaporation was the least effective variable. The simulation results revealed that combining precipitation and aquifer recharge in each area, which had been neglected in the previous studies, had a positive impact on the accuracy of results. Moreover, separating the aquifer area into homogeneous quarters using the K-means clustering approach allowed for selection of the most effective models per region that can lead to definition of suitable management scenarios regarding the condition of each cluster that can enable managers and authorities to make decisions that are more informed in response to different situations. The application of ANN accompanied by either PSO or WOA was successful in improving the efficiency of GWL simulation. The results acquired from the hybrid models demonstrated that each algorithm has its own special ability in solving optimization problems, which should be further investigated in future studies. Overall, simulation of GWL using our approach is a step toward sustainable use and management of water resources that can reliably ensure water supply for urban and rural areas.

REFERENCES

- Abd El Aziz, M., Ewees, A. A., & Hassanien, A. E. (2017). Whale optimization algorithm and moth-flame optimization for multilevel thresholding image segmentation. *Expert Systems with Applications*, 83, 242–256.
- Aljarah, I., Faris, H., & Mirjalili, S. (2018). Optimizing connection weights in neural networks using the whale optimization algorithm. *Soft Computing*, 22(1), 1–15.
- Anand, A., & Suganthi, L. (2020). Forecasting of electricity demand by hybrid ANN-PSO models. In *Deep learning and neural networks: Concepts, methodologies, tools, and applications* (pp. 865–882). IGI Global.
- Arumugam, M. S., Rao, M. V. C., & Chandramohan, A. (2008). A new and improved version of particle swarm optimization algorithm with global–local best parameters. *Knowledge and Information Systems*, 16(3), 331–357.
- Asefpour Vakilian, K. (2020). Machine learning improves our knowledge about miRNA functions towards plant abiotic stresses. *Scientific Reports*, 10, 3041.
- Banadkooki, F. B., Ehteram, M., Ahmed, A. N., Teo, F. Y., Fai, C. M., Afan, H. A., Sapitang, M., & El-Shafie, A. (2020). Enhancement of groundwater-level prediction using an integrated machine learning model optimized by whale algorithm. *Natural Resources Research*, 29(5), 3233–3252.
- Bisht, S., & Paul, A. (2013). Document clustering: A review. *International Journal of Computer Applications*, 73(11), 26–33.
- Chen, W., Hong, H., Panahi, M., Shahabi, H., Wang, Y., Shirzadi, A., Pirasteh, S., Alesheikh, A. A., Khosravi, K., Panahi, S., & Rezaie, F. (2019). Spatial prediction of landslide susceptibility using GIS-based data mining techniques of ANFIS with whale optimization algorithm (WOA) and grey wolf optimizer (GWO). *Applied Sciences*, 9(18), 3755.
- Chen, X. L., Fu, J. P., Yao, J. L., & Gan, J. F. (2018). Prediction of shear strength for squat RC walls using a hybrid ANN–PSO model. *Engineering with Computers*, 34(2), 367–383.
- Coulibaly, P., Anciaux, F., Aravena, R., & Bobée, B. (2001). Artificial neural network modeling of water table depth fluctuations. *Water Resources Research*, 37(4), 885–896.
- Dehariya, V. K., Shrivastava, S. K., & Jain, R. C. (2010). Clustering of image data set using k-means and fuzzy k-means algorithms. In *2010 International conference on computational intelligence and communication networks* (pp. 386–391). IEEE.
- Ebrahimi, H., & Rajaei, T. (2017). Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. *Global and Planetary Change*, 148, 181–191.
- Gaur, S., Ch, S., Graillot, D., Chahar, B. R., & Kumar, D. N. (2013). Application of artificial neural networks and particle swarm optimization for the management of groundwater resources. *Water Resources Management*, 27(3), 927–941.
- Haykin, S. (1999). *Neural network and its application in IR, a comprehensive foundation*, Upper Saddle River. New Jersey: Prentice Hall, 13, 775–781.
- Heil, J., Häring, V., Marschner, B., & Stumpe, B. (2019). Advantages of fuzzy k-means over k-means clustering in the classification of diffuse reflectance soil spectra: A case study with West African soils. *Geoderma*, 337, 11–21.
- Heydari, A., Astiaso Garcia, D., Keynia, F., Bisegna, F., & De Santoli, L. (2019). Hybrid intelligent strategy for multifactor influenced electrical energy consumption forecasting. *Energy Sources, Part B: Economics, Planning, and Policy*, 14(10–12), 341–358.
- Hopfield, J. J. (1988). Artificial neural networks. *IEEE Circuits and Devices Magazine*, 4(5), 3–10.
- Jaafari, A., Panahi, M., Pham, B. T., Shahabi, H., Bui, D. T., Rezaie, F., & Lee, S. (2019a). Meta optimization of an

- adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility. *CATENA*, 175, 430–445.
- Jaafari, A., Termeh, S. V. R., & Bui, D. T. (2019b). Genetic and firefly metaheuristic algorithms for an optimized neuro-fuzzy prediction modeling of wildfire probability. *Journal of Environmental Management*, 243, 358–369.
- Jaafari, A., Zenner, E. K., Panahi, M., & Shahabi, H. (2019c). Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. *Agricultural and Forest Meteorology*, 266, 198–207.
- Jalalkamali, A., Sedghi, H., & Manshouri, M. (2011). Monthly groundwater level prediction using ANN and neuro-fuzzy models: A case study on Kerman plain, Iran. *Journal of Hydroinformatics*, 13(4), 867–876.
- Javadi, S., Saatsaz, M., Shahdany, M. H., Neshat, A., Milan, S., & Akbari, S. (2021). A new hybrid framework of site selection for groundwater recharge. *Geoscience Frontiers*, 12(4), 101144.
- Kardan Moghaddam, H., Kardan Moghaddam, H., Kivi, Z. R., Bahreinimotlagh, M., & Alizadeh, M. J. (2019). Developing comparative mathematic models, BN and ANN for forecasting of groundwater levels. *Groundwater for Sustainable Development*, 9, 100237.
- Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks* (Vol. 4, pp 1942–1948). IEEE.
- Khaki, M., Yusoff, I., & Islami, N. (2015). Simulation of groundwater level through artificial intelligence system. *Environmental Earth Sciences*, 73(12), 8357–8367.
- Khedri, A., Kalantari, N., & Vadiati, M. (2020). Comparison study of artificial intelligence method for short term groundwater level prediction in the northeast Gachsaran unconfined aquifer. *Water Supply*, 20(3), 909–921.
- Kombo, O. H., Kumaran, S., Sheikh, Y. H., Bovim, A., & Jayavel, K. (2020). Long-term groundwater level prediction model based on hybrid KNN-RF technique. *Hydrology*, 7(3), 59.
- Lallahem, S., Mania, J., Hani, A., & Najjar, Y. (2005). On the use of neural networks to evaluate groundwater levels in fractured media. *Journal of Hydrology*, 307(1–4), 92–111.
- Ling, Y., Zhou, Y., & Luo, Q. (2017). Lévy flight trajectory-based whale optimization algorithm for global optimization. *IEEE Access*, 5, 6168–6186.
- Maroufpoor, S., Bozorg-Haddad, O., & Maroufpoor, E. (2020). Reference evapotranspiration estimating based on optimal input combination and hybrid artificial intelligent model: Hybridization of artificial neural network with grey wolf optimizer algorithm. *Journal of Hydrology*, 588, 125060.
- Milan, S. G., Roozbahani, A., Azar, N. A., & Javadi, S. (2021). Development of adaptive neuro fuzzy inference system–evolutionary algorithms hybrid models (ANFIS-EA) for prediction of optimal groundwater exploitation. *Journal of Hydrology* 126258.
- Milan, S. G., Roozbahani, A., & Banihabib, M. E. (2018). Fuzzy optimization model and fuzzy inference system for conjunctive use of surface and groundwater resources. *Journal of Hydrology*, 566, 421–434.
- Ministry of power (2017). Water Resources Balances for Birjand aquifer area, Iran water resource.
- Mirarabi, A., Nassery, H. R., Nakhaei, M., Adamowski, J., Akbarzadeh, A. H., & Aljani, F. (2019). Evaluation of data-driven models (SVR and ANN) for groundwater-level prediction in confined and unconfined systems. *Environmental Earth Sciences*, 78(15), 489.
- Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in Engineering Software*, 95, 51–67.
- Moghaddam, H. K., Milan, S. G., Kayhomayoon, Z., & Azar, N. A. (2021). The prediction of aquifer groundwater level based on spatial clustering approach using machine learning. *Environmental Monitoring and Assessment*, 193(4), 1–20.
- Mohammadi, B., & Mehdizadeh, S. (2020). Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. *Agricultural Water Management*, 106145.
- Mohammadrezapour, O., Kisi, O., & Pourahmad, F. (2020). Fuzzy c-means and K-means clustering with genetic algorithm for identification of homogeneous regions of groundwater quality. *Neural Computing and Applications*, 32(8), 3763–3775.
- Nayak, J., Kanungo, D. P., Naik, B., & Behera, H. S. (2016). Evolutionary improved swarm-based hybrid K-means algorithm for cluster analysis. In *Proceedings of the second international conference on computer and communication technologies* (pp. 343–352). Springer.
- Nguyen, P. T., Ha, D. H., Avand, M., Jaafari, A., Nguyen, H. D., Al-Ansari, N., Phong, T. V., Sharma, R., Kumar, R., Le, H. V., & Ho, L. S. (2020a). Soft computing ensemble models based on logistic regression for groundwater potential mapping. *Applied Sciences*, 10(7), 2469.
- Nguyen, P. T., Ha, D. H., Jaafari, A., Nguyen, H. D., Van Phong, T., Al-Ansari, N., Prakash, I., Le, H. V., & Pham, B. T. (2020b). Groundwater potential mapping combining artificial neural network and real adaboost ensemble technique: The DakNong Province case-study. *Vietnam. International Journal of Environmental Research and Public Health*, 17(7), 2473.
- Nhu, V. H., Mohammadi, A., Shahabi, H., Shirzadi, A., Al-Ansari, N., Ahmad, B. B., Chen, W., Khodadadi, M., Ahmadi, M., Khosravi, K., & Jaafari, A. (2020a). Monitoring and assessment of water level fluctuations of the lake urmia and its environmental consequences using multitemporal landsat 7 etm+ images. *International Journal of Environmental Research and Public Health*, 17(12), 4210.
- Nhu, V. H., Shirzadi, A., Shahabi, H., Singh, S. K., Al-Ansari, N., Clague, J. J., Jaafari, A., Chen, W., Miraki, S., Dou, J., & Luu, C. (2020b). Shallow landslide susceptibility mapping: A comparison between logistic model tree, logistic regression, naïve bayes tree, artificial neural network, and support vector machine algorithms. *International Journal of Environmental Research and Public Health*, 17(8), 2749.
- Pham, B. T., Jaafari, A., Prakash, I., Singh, S. K., Quoc, N. K., & Bui, D. T. (2019). Hybrid computational intelligence models for groundwater potential mapping. *CATENA*, 182, 104101.
- Rajaei, T., Ebrahimi, H., & Nourani, V. (2019). A review of the artificial intelligence methods in groundwater level modeling. *Journal of Hydrology*, 572, 336–351.
- Rokach, L., & Maimon, O. (2005). Clustering methods. In *Data mining and knowledge discovery handbook* (pp. 321–352). Springer.
- Sai, L., & Huajing, F. (2017). A WOA-based algorithm for parameter optimization of support vector regression and its application to condition prognostics. In *2017 36th Chinese control conference (CCC)* (pp. 7345–7350). IEEE.
- Samadianfard, S., Hashemi, S., Kargar, K., Izadyar, M., Mostafaeipour, A., Mosavi, A., Nabipour, N., & Shamshirband, S. (2020). Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm. *Energy Reports*, 6, 1147–1159.
- Sarlaki, E., Sharif Paghaleh, A., Kianmehr, M. H., & Asefpour Vakilian, K. (2021). Valorization of lignite wastes into humic acids: Process optimization, energy efficiency and structural features analysis. *Renewable Energy*, 163, 105–122.
- Seifi, A., & Soroush, F. (2020). Pan evaporation estimation and derivation of explicit optimized equations by novel hybrid meta-heuristic ANN based methods in different climates of Iran. *Computers and Electronics in Agriculture*, 173, 105418.

- Shah, N., & Mahajan, S. (2012). Document clustering: A detailed review. *International Journal of Applied Information Systems*, 4(5), 30–38.
- Spina, R. (2006). Optimisation of injection moulded parts by using ANN-PSO approach. *Journal of Achievements in Materials and Manufacturing Engineering*, 15(1–2), 146–152.
- Taormina, R., Chau, K. W., & Sethi, R. (2012). Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon. *Engineering Applications of Artificial Intelligence*, 25(8), 1670–1676.
- Toghyani, S., Ahmadi, M. H., Kasaeian, A., & Mohammadi, A. H. (2016). Artificial neural network, ANN-PSO and ANN-ICA for modelling the Stirling engine. *International Journal of Ambient Energy*, 37(5), 456–468.
- Vaheddoost, B., Guan, Y., & Mohammadi, B. (2020). Application of hybrid ANN-whale optimization model in evaluation of the field capacity and the permanent wilting point of the soils. *Environmental Science and Pollution Research*, 27(12), 13131–13141.
- Xu, Z., Huang, X., Lin, L., Wang, Q., Liu, J., Yu, K., & Chen, C. (2020). BP neural networks and random forest models to detect damage by *Dendrolimus punctatus* Walker. *Journal of Forestry Research*, 31(1), 107–121.
- Zhang, C. L., Jing, Z. L., Pan, H., Jin, B., & Li, Z. X. (2013). Robust visual tracking using discriminative stable regions and K-means clustering. *Neurocomputing*, 111, 131–143.