On the Determination of Groundwater Level using Temporal and Spatial Parameters: Advanced Machine Learning Methods

Amirhossein Najafabadipour (mailto:najafabadipour@eng.uk.ac.ir)
Shahid Bahonar University Faculty of Engineering

Gholamreza Kamali
Shahid Bahonar University Faculty of Engineering

Hossein Nezamabadi-pour
Shahid Bahonar University Faculty of Engineering

Research Article

Keywords: CF, Committee machine intelligent system, GR, Groundwater Level, MLP, RBF

Posted Date: April 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-234142/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
On the Determination of Groundwater Level using Temporal and Spatial Parameters: Advanced Machine Learning Methods

Amirhossein Najafabadipour 1. Gholamreza Kamali 1. Hossein Nezamabadi-pour 2

Abstract

Prediction of groundwater level is a useful tool for managing groundwater resources in the mining area. Water resources management requires identifying potential periods for groundwater drainage to prevent groundwater from entering the mine pit and reduce high costs. For this purpose, four multilayer perceptron (MLP) neural network models and four cascade forward (CF) neural network models optimized with Bayesian Regularization (BR), Levenberg-Marquardt (LM), Resilient Backpropagation (RB), and Scaled Conjugate Gradient (SCG), as well as a radial basis function (RBF) neural network model and a generalized regression (GR) neural network model were developed to predict groundwater level using 1377 data point. This data set includes 12 spatial parameters divided into two categories of sediments and bedrock, and besides, 6 time series parameters have been used. Also, to determine the best models and combine them, 165 extra

Amirhossein Najafabadipour

najafapidipour@eng.uk.ac.ir

1 Department of Mining Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

2 Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran
validation data points have been used. After identifying the best models from the three candidate
models with lower average absolute relative error (AARE) value, the committee machine
intelligence system (CMIS) model has been developed. The proposed CMIS model predicts
groundwater level data with high accuracy with an AARE value of less than 0.11%. Also, the
proposed model was compared with ten other models through graphical and statistical error
analysis. The results show that the developed CMIS model performs better than other existing
models in terms of precision and validity range. The relevancy factor indicates that the electrical
resistivity of sediments had the highest effect on the groundwater level. Eventually, the quality of
the data used was investigated both statistically and graphically, and the results show satisfactory
reliability of the data used.

Keywords: CF, Committee machine intelligent system, GR, Groundwater Level, MLP, RBF

1. Introduction

Groundwater is a natural resource that has costly adverse effects on mining operations (Brawner
1986). Due to the increased depth of mining, excavation may be done below the water table, which
leads to the movement of water towards mining works. Too much water entering the mining
environment may delay the project or impede production, in addition to causing environmental and
safety problems (Singh and Atkins 1985). Increased equipment failure, lack of access to part of the
mining area, increased use of explosives, loading problems, unsafe working conditions, and a
damaging effect on pit slope stability are among the undesirable impacts of groundwater entering
the mining environment. Therefore, to overcome these problems, it is necessary to develop an
efficient dewatering system that prediction of groundwater level can contribute significantly to this
design. Various classical and modern methods can be used to model and predict groundwater fluctuations.

For classic physical-based modeling, the study area’s exact characteristics and boundary conditions must be accurately identified, and this model is based on some assumptions. Therefore, a physical-based model is not proper for all hydrogeological situation. The numerical models are extensively used to simulate the quantity and quality of groundwater (Xu et al. 2011). Numerical modeling of groundwater by this model (i.e., MODFLOW) requires some input parameters; hence preparing proper values for these parameters is a time-consuming and costly activity. To control limitations, soft computing techniques are a valid alternative to providing more precision with less computational time to predict groundwater levels (Golestani Kermani et al. 2019).

To predict groundwater levels and the complexity of subsurface conditions, machine learning methods based on nonlinear dependence can be used without deep knowledge of basic physical parameters (Parisouj et al. 2020). In recent years, artificial intelligence (AI) methods have been widely used to predict water system variables due to their high ability to learn complex mathematical relationships between output and prediction variables. One of the most common machine learning algorithms used to predict groundwater level is the artificial neural network (ANN) (Santos Finck and Correa Pedrollo 2021), multilayer perceptron (MLP), Cascade forward (CF), Radial basis function (RBF), General regression (GR) and, committee machine intelligence system (CMIS) are the most widely used artificial neural network methods based on their different architectures. Kouziokas et al. (2018) used MLP to predict groundwater levels in Montgomery, Pennsylvania. Hayder et al. (2020) modeled river flow using optimized CF and MLP in the Kelantan River in Malaysia. Hong et al. (2020) predicted trihalomethanes levels in tap water using RBF and gray relational analysis (GRA). Pal and Chakrabarty (2020) have evaluated GR neural network
models in simulating the groundwater contaminant transport. Malekpour and Mohammad Rezapour Tabari (2020) have implemented the supervised intelligence committee machine method to predict reservoir water level variation for the design and operation of dams. In addition to different methods for modeling and predicting groundwater levels, different spatial and temporal data can affect groundwater levels (Rajaee et al. 2019). Most studies have used spatial or temporal parameters to predict groundwater levels using machine learning, but both spatial and temporal parameters affect groundwater levels.

The main purpose of this research is to develop and compare widely used and accurate machine learning methods to predict groundwater levels using both spatial and temporal parameters. For this purpose, 1542 data points including 12 spatial parameters and 6 time series parameters were used. Out of 1377 data have been used to create different networks, and to evaluate the performance of these developed models, 165 extra validation data have been used. Four multilayer perceptron neural network (MLP) models and four cascade forward neural network (CF) models were then developed using four different optimized Bayesian Regularization (BR), Levenberg-Marquardt (LM), Resilient Backpropagation (RB), and Scaled Conjugate Gradient (SCG). Also, the radial basis function neural network (RBF) and general regression neural network (GR) methods have been used to model the groundwater level. Although CF and GR models have been used for water resource management, these models have not yet been developed for the prediction of groundwater level. After developing models, a committee machine intelligent system (CMIS) is combined of three candidate models with the least error. The validity of the proposed CMIS is evaluated through statistical and graphical error analysis. The innovation of this research, in addition to identifying the relevancy factor of the data relative to the groundwater level, is also identifying outliers.
The rest of this research is organized as follows: Section 2 provides information on the study area and the data used; Sections 3 and 4 describe the developed models and optimization techniques, respectively; Section 5 outlines the results and discussion of this research, and section 6 presents the conclusions of this research.

2. Experimental Data

Gol Gohar iron ore deposit, one of the most popular pivot points of the mining industry in the Middle East, with six separate anomalies and a reservoir of about 1200 million tons, is located in an area with a length of approximately 10 km and nearly 4 km in width. In anomaly No.3 (Gohar Zamin Iron Ore Mine), groundwater enters the pit, and also water permeates through the alluvium of the pit’s stairs. One of the probable factors of going groundwater inflow into Gohar Zamin Iron Ore Mine is the Kheyrabad plain with alluvial sediments situated in the northeast of the mine at a distance of 15 km (Fig. 1.). Around the Gohar Zamin Iron Ore Mine, water pumping wells are located around anomaly No. 1, which is considered as a discharge area.

To estimate the spatial and temporal groundwater level as the target and output of the neural network, two sets of spatial and temporal data have been used as the input of the neural network. The input spatial data set includes five piezometer data around the Gohar Zamin Iron Ore Mine. Besides, the input temporal data consists of the time series of parameters affecting the groundwater level in the period of March 21, 2019, to July 2, 2020. Because both sediments and bedrock affect the groundwater level, the spatial data input to the neural network is divided into two categories of bedrock and sediment parameters. In Table 1, input spatial data to the neural network including latitude and longitude of piezometers, hydraulic conductivity of sediments and bedrock, effective porosity of sediments and bedrock, the electrical resistivity of sediments and bedrock, the
piezometers surface level, bedrock height, Depth of sediments and the presence or absence of faults are shown. The input temporal data to the neural network include day, month, year, drainage volume, evaporation, and precipitation, which in Table 2, statistical explanation of these parameters are shown.

3. Models

3.1 Multilayer perceptron neural network (MLP)

Artificial neural networks (ANN) as a useful computational intelligence built on analogy with human information processing systems widely used in distributed processing systems (Hebb 1949). Although the structures of the artificial neural network (ANN) are very varied, multilayer perceptron (MLP) is still one of the most dominant as well as the most extensive structures of the artificial neural network. The MLP shown by Cybenko's theorem (1989) is a universal function approximation used to create mathematical models using regression analysis. With training on observation data, the network can learn specific features hidden in the collected data samples and even generalize what it has learned. MLP networks have a multilayered structure, and the first layer is the input data to the model, the last layer is the output data of the model, and the layers between the input and output are called hidden layers. The number of neurons in the input layer is the same as the input variables, the number of outputs is usually the same as the output parameter, and the hidden layers are responsible for the internal appearance of the relationship between the model inputs and the desired output. The value of each neuron in the hidden layer or the output layer is the sum of each neuron in the previous layer, multiplied in a particular weight for that neuron, and summed with bias, and passed from an activation function. The following is a summary of some of the common activation functions used for hidden and output layers.
Linear=Purelin: \( f(x) = x \)  

(1)

Logsig=Sigmoid: \( f(x) = \frac{1}{1+e^{-x}} \)  

(2)

Tansig=Tanh: \( f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \)  

(3)

Elliot 2 Symmetric Sigmoid: \( f(x) = \frac{x^2 \text{SIGN}(x)}{1+x^2} \)  

(4)

The output of an MLP model with two hidden layers whose activation functions for these two layers are Logsig and Tansig, and Purlin activation function for the output layer are as follows (Hemmati-Sarapardeh et al. 2018):

\[
\text{Output} = \text{Purelin}(w_3 \times (\text{Tan sig}(w_2 \times (\text{Log sig}(x) + b_1)) + b_2) + b_3)
\]

(5)

Where \( w_1 \) and \( b_1 \) are the weight matrixes and the bias vectors of the first hidden layers, \( w_2 \) and \( b_2 \) are the weight matrixes and the bias vectors of the second hidden layers, and \( w_3 \) and \( b_3 \) are the weight matrixes and the bias vectors of the output layers.

3.2 Cascade forward neural network (CF)

There is a direct relationship between input and output in the perceptron connection, whereas in the feedforward neural network connection, there is an indirect relationship between input and output, which is a hidden layer through a nonlinear activation function (Fahlman and Lebiere 1989). If the connection form is combined in a multilayer network and perceptron, the network can be formed with a direct connection and the indirect connection between the input layer and the output layer (Hayder et al. 2020). The network formed of this connection pattern is called cascade forward neural network (CF), and the equations of this model can be written as follows:

\[
y = \sum_{i=1}^{n} f^i w^i x_i + f^o (w^b + \sum_{j=1}^{k} w^o_j f^h (w^b_j + \sum_{i=1}^{n} w^h_{ji} x_i))
\]

(6)
Where $f^i$ is the activation function between the input layer and the output layer, $w_{ij}^i$ is the weight between the input layer and the output layer, $w^b$ is the weight from bias to output, $w_{ij}^b$ is the weight from bias to hidden layer, and $f^h$ is the activation function of each neuron in the hidden layer.

Fig. 2. shows the CF architecture where there is a direct relationship between input and output. Because the optimization algorithm used for training has influential impress in the efficiency of the MLP and CF model, in this research, four important Bayesian Regularization (BR), Levenberg-Marquardt (LM), Resilient Backpropagation (RB), and Scaled Conjugate Gradient (SCG) optimization algorithms have been used.

### 3.3 Radial basis function neural network (RBF)

Radial basis function neural network (RBF) is one of the powerful feedforward neural networks that use performance approximation theory in solving regression problems. Broomhead and Lowe (1988), based on adaptive function interpolation, introduced an approach to local functional approximation. In general, an RBF neural network has a three-layer feedforward structure in which the input layer and the output layer are connected through a hidden layer. The principal part of the RBF neural network is the hidden layer, which transmits data from the input space to the hidden space with higher dimensions (Hemmati-Sarapardeh et al. 2018). Each point in the hidden layer with a particular radius is located at a given space, that in each neuron, the distance between the input vector and its center is calculated. Euclidean distance is used to measure the distance between centers and inputs, which is calculated from the following equation:

$$r_j = \sqrt{\sum_{i=1}^{p} (x_i - c_{ij})^2}$$  (7)
For a model with ten input variables, \( P=10 \). To transfer the Euclidean distance from each neuron in the hidden layer to the output, a radial basis function has been used. The most common radial basis function is the Gaussian, which is obtained from the following interface:

\[
\text{Gaussian: } \phi(r) = \exp\left(\frac{-r^2}{2\sigma^2}\right)
\]  

(8)

Due to its smoother and flexible behavior, the Gaussian function has been utilized as the activation function in this research. When \( x = c_j \), \( \phi(r) \) is maximum, and as \( r \) increases, the \( \phi(r) \) decreases. When \( |r| \to \infty \), \( \phi(r) \to 0 \). \( \sigma \) is the spreading coefficient of the Gaussian function, which is defined experimentally (Meshram et al. 2020). The model output is estimated from the following equation:

\[
y_k = \sum_{j=1}^{N} \omega_j \phi_{kj}(\|x_k - c_j\|) \quad j = 1, \ldots, N \quad \text{and} \quad k = 1, \ldots, M
\]  

(9)

Where \( \omega \) shows connection weight, \( N \) is the number of neurons in the hidden layer, \( c \) denotes the center, \( \|x - c\| \) is Euclidean distance between the center of the radial function and the input data (Menad et al. 2019).

### 3.4 Generalized regression neural network (GR)

Generalized regression neural network (GR) is a memory neural network that is a variation to radial basis function neural network (RBF) based on the statistical technique of kernel regression with a dynamic network structure with powerful nonlinear mapping and robustness (Specht 1991). GR has a high learning speed and is very useful for function approximation problems. For small sample data, the prediction effect is excellent, and also unstable data can be processed. GR does not have RBF accuracy but has a major advantage in classification and fit, especially when data accuracy is inappropriate. The GR model has a two-layer structure, the first layer being the hidden RBF layer and the second layer being the linear output layer (Tayfur et al. 2020). For the hidden layer RBF, the number of elements is equal to the number of training samples. The weight function of the
hidden layer RBF is dist, which is used to determine the distance between the network input and
the weight value $IW_{11}$ of this layer and is calculated from the following equation:

\[
\|dist\|_j = \sqrt{\sum_{i=1}^{R} (x_i - IW_{ji})^2}, \quad (j = 1, 2, \ldots, M)
\]  

(10)

In the hidden layer, network product function $netprod$ multiplies the threshold $b_1$ and $\|dist\|$ output
to get net input $n^1$. The net input $n^1$ is passed to transfer function $radbas$. For the GR model, the
Gaussian function is used as the transfer function, namely,

\[
a_j^1 = \text{radbas} \left(\text{netprod}\left(\|dist\|_j b_{1,j}\right)\right)
\]

\[
= \exp \left[ - \frac{(n_j^1)^2}{2\sigma_j^2} \right], \quad (j = 1, 2, \ldots, M)
\]

(11)

In the above equation, $\sigma_j$ is a smoothing factor, also called the spread parameter, which calculates
the shape of RBF in the $j$th hidden layer.

The normalized weight product function is used as the weight function in the linear output layer,
making the former layer's output with the weight value $IW_{21}$ in this layer as the weight output. The
Purelin function is used as the transfer function for the output of the passed weight. The network
output is calculated from the following equation:

\[
y_k = \sum_{j=1}^{m} IW_{kj} a_j^1 \quad (k = 1, 2, \ldots, S)
\]

(12)

For the GR model, only the spread parameter should be specified. Due to the significant effect of
this parameter on the model’s performance, the optimal value of this parameter must be determined.
3.5 Committee machine intelligent system (CMIS)

To achieve the desired goals, different models are developed, and the best models are chosen as a candidate, and other models are discarded. Under these circumstances, the cost incurred for the discarded models is wasted. For this purpose, a committee machine can be built by combining intelligent models and using the features of each of these models. This committee machine intelligent system (CMIS) was introduced by Nilsson in 1965; it is a kind of artificial neural network and uses division and conquers to solve problems.

In the committee machine method, the models are combined to provide a more accurate solution (Hemmati-Sarapardeh et al. 2020). To combine the models, the linear combination method can be used using simple averaging or weighted averaging (Hashem and Schmeiser 1970). Because all models’ contribution in the simple averaging method is the same, a satisfactory answer is not obtained from this method because a more precise solution must contribute to the final model more. For this purpose, the weighted average that combines models based on their precision, and the sum of coefficients, the linear composition of the unit, can be used. In this research, added a bias term to the equation and an improved weighted average is used. In the final model, any model’s contribution corresponds to that model’s coefficient in the linear equation of the committee machine intelligence system model.

4. Optimization techniques

4.1 Levenberg-Marquardt algorithm (LM)

The most common tool used to optimize the weight and bias of multilayer perceptron (MLP) and cascade forward (CF) is the Levenberg-Marquardt (LM) algorithm, also known as the damped least-squares method. This algorithm is used to solve nonlinear least-squares problems that find local
minimums. This method does not require to calculate of the Hessian matrix and the gradient is calculated from the following equation (Kisi and Uncuoğlu 2005):

\[ H = J^T J \]  
\[ g = J^T e \]  

(13)  
(14)

Where \( e \) stands a vector of network errors, and \( J \) expresses a Jacobian matrix. In the following relation of updating the LM algorithm, the mentioned approximation with the Hessian matrix is used:

\[ x_{k+1} = x_k - (J^T J - \eta I)^{-1} J^T e \]  

(15)

\( \eta \) is a constant, and \( x \) denotes connection weights. \( \eta \) increases when an experimental step enlarges the efficiency function.

### 4.2 Bayesian Regularization algorithm (BR)

The Bayesian Regularization (BR) training algorithm, according to Levenberg-Marquardt (LM) optimization, updates weights and biases by minimizing a combination of squared errors and weights (MacKay 1992). Afterward, BR calculates the right combination to develop a network with superior generalization. Network weights are expressed as a training cost function by the BR algorithm using the following equation:

\[ F(\omega) = \alpha E_\omega + \beta E_D \]  

(16)

In which \( E_D \) and \( E_\omega \) are the sum of the network errors and the sum of the squared network weights, respectively, \( F(\omega) \) denotes the objective function. In the BR optimizer, the network weights are random variables in which the network weights and the training sets have a Gaussian distribution. \( \alpha \) and \( \beta \) Factors are objective function parameters that are clarified based on Bayes’ theorem.
4.3 Scaled Conjugate Gradient algorithm (SCG)

One of the basic features of the backpropagation algorithm is to reach the most negative gradient, and it uses the adjustment of weights in the steepest descending direction. Along such a direction, a decrease in function performance is observed faster but does not cause faster convergence. In this direction, a search is performed in the conjugate gradient (CG) method, which leads to faster convergence than the steepest descending direction, and error minimization is maintained in previous steps (Kisi and Uncuoğlu 2005).

\[ P_0 = -g_0 \]  

(17)

\( P \) is search direction, and \(-g_0\) denotes the steepest descent direction in the first iteration. This direction is called conjugate direction, commonly used by conjugate gradient algorithms with the search line. To evaluate the optimal distance to move in the current search direction, the step size is determined by a line search technique (Kisi and Uncuoğlu 2005), which is shown by the following equation:

\[ x_{k+1} = x_k + \alpha_k g_k \]  

(18)

In other words, the proper search direction is calculated in a way that conjugates with the previous search direction.

\[ P_k = -g_k + \beta_k P_{k-1} \]  

(19)

Different versions of the conjugate algorithm are distinguished in the way that \( \beta \) is calculated (Kisi and Uncuoğlu 2005).

4.4 Resilient Backpropagation algorithm (RB)
The most widely used transfer functions in multilayer perceptron neural networks (MLP) are Sigmoid and Tansig, which compress an infinite input range into a finite output. When using the steepest descent to train the network using these activation functions, the slope is small when an extensive input enters the function, leading to slight changes in weights and biases. The Resilient backpropagation (RB) method is used to remove the adverse effect of the partial derivatives, which is specified by the derivatives only for the direction of updating weights (Riedmiller and Braun 1993).

5. Results and discussion

5.1 Models development

In this research, 1542 data points of five piezometric wells have been used to model and predict the groundwater level. The data used are divided into spatial (sediments and bedrock) and temporal (March 21, 2019, to July 2, 2020). To consider the complexity of intelligent methods, 12 spatial parameters (longitude, latitude, hydraulic conductivity of sediments, effective porosity of sediments, the electrical resistivity of sediments, depth of sediments, surface level, hydraulic conductivity of bedrock, effective porosity of bedrock, the electrical resistivity of bedrock, bedrock level, and fault) and 6 time series parameters (day, month, year, drainage, evaporation, and rainfall) have been used (Rajaee et al. 2019). 1377 data points have been used to develop four multilayer perceptron neural network (MLP) models, four cascade forward neural network (CF) models, one radial basis function neural network (RBF) model, and one general regression neural network (GR) model. MLP and CF models are each optimized by Levenberg-Marquardt (LM), Bayesian Regularization (BR), Scaled Conjugate Gradient (SCG), and Resilient Backpropagation (RB) methods. In all the mentioned models, 80% of the data were used for model training, and 20% of
the data were used for model validity. The data are divided into testing and training sets using random distributions to prevent the local accumulation of data.

The number of hidden layers, the type of transmission function, and the number of neurons in each layer affects the efficiency of a developed model. Trial and error can be used to identify these parameters. In this research, two hidden layers were used for multilayer perceptron neural network (MLP) and cascade forward neural network (CF) methods, and one hidden layer was used for radial basis function neural network (RBF) and general regression neural network (GR) methods. Table 4 shows the function used and the best architecture in terms of the number of neurons for each model. Transfer functions are designed to correctly model the complex behavior of nonlinear input and output datasets. The architecture for the MLP model consists of four numbers, the first and last number being the number of inputs and outputs of the model, and the second and third numbers being the number of neurons in the first and second hidden layers. The architecture for the CF model consists of five numbers, the first and last number being the number of inputs and outputs of the model, the third and fourth numbers being the number of hidden layer neurons in the first and second. The second number is the number of neurons in the connections layer between input and output. The architecture for the RBF model consists of 3 numbers, the first and last number being the number of inputs and outputs of the model, and the second number being the number of maximum neurons in the hidden layers. In Table 3, to compare the performance of the models, a summary of statistical parameters including average absolute relative error (AARE, %), average relative error (ARE, %), root mean square error (RMSE), standard deviation (SD), and model run time for one iteration (T, Sec) for all developed models has been calculated. The mentioned parameters are expressed as:
Based on the information in Table 3, the multilayer perceptron neural network (MLP) and cascade forward neural network (CF) created a model based on Levenberg-Marquardt (LM) and Bayesian Regularization (BR) have more precise results than Resilient Backpropagation (RB) and Scaled Conjugate Gradient (SCG) optimizers. The higher LM and BR optimizer accuracy possibly due to using nonlinear least squares to find local minimums. Also, due to the inaccuracy of SCG and RB methods compared to LM and BR methods, it is concluded that these optimizers are not suitable for regression analysis. Since the efficiency of the created models is strongly influenced by the initial biases and weights, the training of artificial neural networks with each optimizer using trial and error was executed more than 50 times with dissimilar initial biases and weights, and the most satisfactory results were chosen.

Radial basis function neural network (RBF) models consist of two key parameters: the number of neurons and the spread coefficient. To determine these two parameters, trial and error have been used. In this research, the number 5 for the coefficient of expansion and the number 30 for the number of neurons have been used. As shown in Table 3, the RBF model is less accurate than other
models but requires less run time the program, which can be used for models with many input data.

To identify the optimal value for the coefficient of expansion and the number of neurons, the RBF model was implemented more than 100 times, and the best results were stored. The general regression neural network (GR) model, like the RBF model, has the spread coefficient parameter, which is 0.5 for this parameter by trial and error. Based on the information in Table 3, the developed GR model has a value of 0.035% average absolute relative error (AARE, %). This model like RBF models requires less time to run than other methods. RBF and GR methods are faster than the multilayer perceptron neural network (MLP) and cascade forward neural network (CF) methods due to their monolayer. The results of Table 3 have been used to rank the proposed models based on having the highest accuracy:

\[\text{MLP - LM} > \text{CF - BR} > \text{MLP - BR} > \text{MLP - RB} > \text{CF - LM} > \text{CF - RB} > \text{CF - SCG}\]

> MLP - SCG > GR > RBF

After developing different models using 1377 data points in the previous step, 165 extra validation data points have been used to identify the best models and their combination and to develop the proposed committee machine intelligence system (CMIS) model. The statistical results of the developed models using 165 extra validation data points are shown in Table 4. The results show high accuracy of radial basis function neural network (RBF), multilayer perceptron neural network using Levenberg-Marquardt optimizer (MLP-LM), and cascade forward neural network using Bayesian Regularization optimizer (CF-BR) compared to other models. The three models developed with the lowest average absolute relative error (AARE, %) value are combined with a CMIS. To detect the optimal coefficients of this model, multiple linear regression is used, which is shown in the next equation:
\[ \mu_{\text{CMIS}} = \alpha_1 \mu_{\text{MLP-LM}} + \alpha_2 \mu_{\text{CF-BR}} + \alpha_3 \mu_{\text{RBF}} + \alpha_4 \] (24)

In the above equation, \( \alpha_1 \) to \( \alpha_4 \) are as follows:

\[ \alpha_1 = 0.525543; \alpha_2 = 0.522345; \alpha_3 = -0.04743; \alpha_4 = -0.7539; \]

CMIS model proposed in this research showed in Fig. 3.

5.2 Evaluation of the validity and precision of the developed models

Statistical parameters such as average absolute relative error (AARE, \%), average relative error (ARE, \%), root mean square error (RMSE), standard deviation (SD) for developed models are shown in Table 4. Some models, such as radial basis function neural network (RBF), which has the highest AARE value in Table 3, but here with one of the lowest error values, is one of the best models to predict, indicating accurate learning between input and output relationships. Other models, such as the cascade forward neural network using Scaled Conjugate Gradient optimizer (CF-SCG), which do not have a good prediction of groundwater level from extra validation data, probably due to overtraining and memorizing input-output rules or maybe stuck in local optimizations. AARE value of about 0.11% indicates that the proposed committee machine intelligence system (CMIS) model is the most precise model for groundwater level forecasting, among other developed models. In addition to the CMIS model, RBF, cascade forward neural network using Bayesian Regularization optimizer (CF-BR), and multilayer perceptron neural network using Levenberg-Marquardt optimizer (MLP-LM) models alone have an acceptable ability to predict groundwater levels. To visually confirm the precision of the developed CMIS model using 1377 data points, the cross plot of the experimental data versus predicted relative groundwater level are plotted in Fig. 4. and the relative error distribution curve for testing and training dataset is drawn in Fig. 5. The high concentration of test and train data around the unit slope indicates the
high accuracy of the CMIS model prediction. In Fig. 5, the maximum relative error of the predicted
and experimental data values is 0.15%. Most of the dataset points are located about the zero-error
line for the tentative relative groundwater level. This figure confirms the stability between the
experimental data and the CMIS model prediction. Fig. 6 shows the actual groundwater level and
the prediction of groundwater level for 165 days ahead by the developed models with the least
AARE value.

The lowest average absolute relative error (AARE, %), average relative error (ARE, %), root mean
square error (RMSE), standard deviation (SD) values indicate an acceptable estimate by the
developed committee machine system intelligence (CMIS) model compared to other developed
models. A detailed error analysis has been used to more compare the efficiency of the proposed
CMIS model and other developed models for groundwater level prediction. The AARE (%) in
Fig.7-11 is drawn as a function of the input parameters that have the most significant impact on the
groundwater surface. According to Table 1, the highest relevancy is related to the electrical
resistivity of sediments, depth of sediments, the electrical resistivity of bedrock, hydraulic
conductivity of bedrock, and effective porosity of sediments. For this purpose, error analysis was
performed by dividing the entire data into four groups to indicate the models’ precision in various
ranges of crucial parameters.

Fig. 7 compares the average absolute relative error (AARE, %) of dissimilar models in four ranges
of electrical resistivity of sediments. The maximum AARE value of 0.021% for the proposed
committee machine intelligence system (CMIS) model indicates its very high capability to predict
the groundwater level in the whole range of electrical resistivity of sediments. As shown from this
figure, developed models for the range of 21 to 32 (ohm-m) have a lower AARE error than other
values. Fig. 8 shows the AARE (%) of the developed models in four ranges of the depth of
sediments. This figure clearly shows a lower error of the CMIS model than all models. Close to the surface in the range of 130 to 150 meters, the AARE value is the lowest. Fig. 9. presents the efficiency of the developed models over dissimilar ranges of the electrical resistivity of bedrock. As can be easily seen, as in Figure 7, the lowest AARE value here is in the range of 21 to 30 (ohm-m). Besides, the CMIS model has the lowest AARE in all four ranges. Fig. 10. depicts the AARE of the developed models for dissimilar ranges of hydraulic conductivity of bedrock. Again, the proposed CMIS is accurate in the entire range, especially in the range of 18 to 19 (m/day). The effect of porosity of sediments in Fig. 11. in four ranges shows a low AARE value for all models, mainly CMIS. In all ranges, the AARE values are close to each other models and below 0.05%.

To better compare the models, the cumulative frequency of average absolute relative error for the developed committee machine intelligence system (CMIS) model and all other models is shown in Fig. 12. According to this figure, about 85.5% of the groundwater level predicted by the developed CMIS model has an average absolute relative error (AARE, %) of less than 0.03%. Other developed multilayer perceptron neural network using Levenberg-Marquardt optimizer (MLP-LM), multilayer perceptron neural network using Bayesian Regularization optimizer (MLP-BR), and cascade forward neural network using Bayesian Regularization optimizer (CF-BR) models have 82.4%, 83.1%, and 82.9% errors. The results of this figure further indicate the success of the proposed CMIS method compared to other developed methods for groundwater level prediction.

5.3 Trend analysis of the developed model

Trend analysis in a hydrogeological time series can be a practical tool to study groundwater level fluctuations (Halder et al. 2020). By examining the groundwater level predicted in Figure 6, significant changes are seen on days 40 to 110, indicating the pumping well in the southeastern part
of the mine. For this reason, considering that the drainage time series is a parameter affecting the groundwater level, the physically expected trends of the groundwater level by changing the drainage is investigated by the models developed in Fig 13. As can be seen in this figure, the relative groundwater level decreases with the drainage increases. All developed models can record the expected trend with drainage changes. This figure also shows the accuracy of the proposed committee machine intelligence system (CMIS) model compared to other methods.

5.4 Sensitivity analysis on models’ inputs

Sensitivity analysis is important from the perspective that uncertainty in model inputs affects uncertainty in model output. Sensitivity analysis can be used to evaluate the correlation between model its inputs and outputs, to search for errors in the model structure, and to simplify a developed model by removing its inputs that do not affect its output. To determine the effect of a model’s input parameter on its output, the reliable method of relevancy factor analysis can be used. The relevancy factor measures the effect of input parameter on the output and is denoted by $r$. The higher $r$ value, the greater the effect of the input on the output. The relevancy factor is calculated from the following equation:

$$r(\text{Inp}_k, \mu) = \frac{|\sum_{i=1}^{n}(\text{Inp}_{k,i} - \text{Inp}_{\text{ave,k}})(\mu_i - \mu_{\text{ave}})|}{\sqrt{\sum_{i=1}^{n}(\text{Inp}_{k,i} - \text{Inp}_{\text{ave,k}})^2 \sum_{i=1}^{n}(\mu_i - \mu_{\text{ave}})^2}}$$

(25)

Where $\text{Inp}_{k,i}$ shows the $i$th value, and $\text{Inp}_{\text{ave,k}}$ is the average value of $k$th input, respectively ($k$ represents the model inputs). $\mu_i$ and $\mu_{\text{ave}}$ show the $i$th value and the average value of the predicted output (Hemmati-Sararpardeh et al. 2020). The Fig. 14. shows the value of the relevancy factor. The obtained numbers show that spatial parameters have a more significant effect on groundwater surface than temporal parameters. The most influential spatial parameters on the groundwater level
are the electrical resistivity of sediments, depth of sediments, the electrical resistivity of bedrock, hydraulic conductivity of the bedrock, and effective porosity of sediments. Among the temporal parameters, drainage volume has the most significant impact on groundwater level due to pumping wells in the pit mine’s southeastern region.

5.5 Outlier diagnosis and applicability domain of the model

Outliers commonly appear in a comprehensive set of experimental data and differ from the bulk of the data. Because such a dataset can affect the reliability and precision of experimental models, thus finding this type of dataset is essential in developing models. One of the practical advantages of this analysis is that by carefully examining the outlier data, a good view of the model constraint can be provided, which may be due to the ignoring of some justifiable effects. In this research, the leverage approach has been used to determine and eliminate outliers, including calculating the model deviation from the relevant experimental data. The computed model deviations are placed in a Hat matrix and are called "standardized cross-validated residuals" (Rousseeuw and Leroy 1987).

William's plot is plotted in Fig. 15. for the resulting by the developed committee machine intelligence system (CMIS) model for 1542 groundwater level predictions data. Due to the location of most of the predicted points in the feasibility domain of the proposed model (0 ≤ hat ≤ 0.0414 and −3 ≤ R ≤ 3), statistical validity and high reliability of the developed CMIS model are shown. About 2.65% of the points are outside the acceptable range of the model, which can be ignored due to the number of data points used in the model’s development. Points in the range of R < −3 or R > 3 are defined as "Bad High Leverage" regardless of their hat value compared to hat* (Hemmati-Sarapardeh et al. 2016). These data may be well predicted, but due to the difference with a large amount of data, but, are outside the acceptable range of the model.
6. Conclusions

In this research, data from five piezometric wells includes 1542 data points around Gohar Zamin Iron Ore Mine, located in Sirjan, Iran, have been used to predict the groundwater level. 1377 spatial (sediment and bedrock) and temporal data points (276 days) have been used to develop ten supervised learning models. 165 extra validation data points have been used to identify the best models with less average absolute relative error (AARE, %) value, combine them into a single model, and develop the proposed committee machine intelligence system (CMIS) model. The results of this research can be concluded:

1. Almost all developed models with 1377 data points can predict groundwater levels with acceptable AARE values in the range of 0.017 to 0.043%. The lowest error is related to multilayer perceptron neural network using Levenberg-Marquardt optimizer (MLP-LM), cascade forward neural network using Bayesian Regularization optimizer (CF-BR), and multilayer perceptron neural network using Bayesian Regularization (MLP-BR) models, all of which have average absolute relative error (AARE, %) values below 0.018%.

2. The running time of different algorithms with similar conditions shows that general regression neural network (GR) and radial basis function neural network (RBF) algorithms are significantly faster due to these models’ monolayer.

3. Developed models using 1377 data points to predict groundwater levels for 165 extra validation data points showed almost similar results, except that the RBF model has the lowest average absolute relative error (AARE, %). A lower AARE value for the radial basis function neural network (RBF) model indicates acceptable learning of the input-output relationships. Developed models predict groundwater levels for these 165 extra validation data points with AARE values
between 0.12 and 0.84%. A combination of three models, RBF, cascade forward neural network using Bayesian Regularization optimizer (CF-BR), and multilayer perceptron neural network using Levenberg-Marquardt optimizer (MLP-LM), is used to develop the committee machine intelligence system (CMIS) model.

4. The developed committee machine intelligence system (CMIS) model has a better performance than all other models and can predict the groundwater level with an average absolute relative error (AARE, %) value of about 0.11%.

5. About 85.5% of the prediction of groundwater level has an average absolute relative error (AARE, %) value below 0.03%, which indicates the statistical validity of this method. Data that are out of the applicability domain of the proposed committee machine intelligence system (CMIS) are about 2.65%, which indicates the high accuracy of this method.

6. The electrical resistivity of sediments, Depth of sediments, the electrical resistivity of bedrock, hydraulic conductivity of bedrock, and effective porosity of sediments of the 18 input parameters are the most influential input parameters on the groundwater level, respectively. Also, trend analysis for drainage parameters shows a decrease in groundwater level with increasing drainage due to pumping wells’ activity in the southeastern part of the mine.

Declaration

Ethical Approval

All work is compliance with Ethical Standards.

Consent to Participate
Authors give their permission.

Consent to Publish

Authors give their permission.

Authors Contributions

All authors have contributed to the paper. Amirhossein Najafabadi-pour has written the paper and Gholamreza Kamali and Hossein Nezamabadi-pour reviewed and improved the manuscript several times.

Funding

Not applicable

Competing Interests

Not applicable

Availability of data and materials

Not applicable
References

Brawner CO (1986) Groundwater and coal mining Mining Science and Technology 3:187-198
doi:https://doi.org/10.1016/S0167-9031(86)90331-2

Broomhead D, Lowe D (1988) Radial basis functions, multi-variable functional interpolation and adaptive
networks ROYAL SIGNALS AND RADAR ESTABLISHMENT MALVERN (UNITED
KINGDOM) RSRE-MEMO-4148

Cybenko G (1989) Approximation by superpositions of a sigmoidal function Mathematics of Control,
Signals and Systems 2:303-314 doi:10.1007/BF02551274

Fahlman S, Lebiere C (1989) The cascade-correlation learning architecture Advances in neural information
processing systems 2:524-532

Golestani Kermani S, Sayari S, Kisi O, Zounemat-Kermani M (2019) Comparing data driven models versus
numerical models in simulation of waterfront advance in furrow irrigation Irrigation Science
37:547-560 doi:10.1007/s00271-019-00635-5

Halder S, Roy MB, Roy PK (2020) Analysis of groundwater level trend and groundwater drought using
Standard Groundwater Level Index: a case study of an eastern river basin of West Bengal, India
SN Applied Sciences 2:507 doi:10.1007/s42452-020-2302-6

Hashem S, Schmeiser B (1970) Approximating a Function and its Derivatives Using MSE-Optimal Linear
Combinations of Trained Feedforward Neural Networks 1

Hayder G, Solihin MI, Mustafa HM (2020) Modelling of River Flow Using Particle Swarm Optimized
Cascade-Forward Neural Networks: A Case Study of Kelantan River in Malaysia Applied Sciences
10:8670

Hebb DO (1949) The organization of behavior: a neuropsychological theory. J. Wiley; Chapman & Hall,

Hemmati-Sarapardeh A, Ameli F, Dabir B, Ahmadi M, Mohammadi AH (2016) On the evaluation of
asphaltenes precipitation titration data: Modeling and data assessment Fluid Phase Equilibria
415:88-100
Hemmati-Sarapardeh A, Larestani A, Nait Amar M, Hajirezaie S (2020) Chapter 1 - Introduction. In: Hemmati-Sarapardeh A, Larestani A, Nait Amar M, Hajirezaie S (eds) Applications of Artificial Intelligence Techniques in the Petroleum Industry. Gulf Professional Publishing, pp 1-22. doi:https://doi.org/10.1016/B978-0-12-818680-0.00001-1

Hemmati-Sarapardeh A, Varamesh A, Husein MM, Karan K (2018) On the evaluation of the viscosity of nanofluid systems: Modeling and data assessment Renewable and Sustainable Energy Reviews 81:313-329 doi:https://doi.org/10.1016/j.rser.2017.07.049

Hong H et al. (2020) Radial basis function artificial neural network (RBF ANN) as well as the hybrid method of RBF ANN and grey relational analysis able to well predict trihalomethanes levels in tap water Journal of Hydrology 591:125574 doi:https://doi.org/10.1016/j.jhydrol.2020.125574

Kisi O, Uncuoğlu E (2005) Comparison of three back-propagation training algorithms for two case studies Indian Journal of Engineering and Materials Sciences 12

Kouziokas GN, Chatzigeorgiou A, Perakis K (2018) Multilayer Feed Forward Models in Groundwater Level Forecasting Using Meteorological Data in Public Management Water Resources Management 32:5041-5052 doi:10.1007/s11269-018-2126-y

MacKay DJ (1992) Bayesian interpolation Neural computation 4:415-447

Malekpour MM, Mohammad Rezapour Tabari M (2020) Implementation of supervised intelligence committee machine method for monthly water level prediction Arabian Journal of Geosciences 13:1049 doi:10.1007/s12517-020-06034-x

Menad NA, Noureddine Z, Hemmati-Sarapardeh A, Shamshirband S (2019) Modeling temperature-based oil-water relative permeability by integrating advanced intelligent models with grey wolf optimization: Application to thermal enhanced oil recovery processes Fuel 242:649-663 doi:https://doi.org/10.1016/j.fuel.2019.01.047

Meshram SG, Singh VP, Kisi O, Karimi V, Meshram C (2020) Application of Artificial Neural Networks, Support Vector Machine and Multiple Model-ANN to Sediment Yield Prediction Water Resources Management 34:4561-4575 doi:10.1007/s11269-020-02672-8
Nilsson N (1965) Learning machines

Pal J, Chakrabarty D (2020) Assessment of artificial neural network models based on the simulation of groundwater contaminant transport Hydrogeology Journal 28:2039-2055 doi:10.1007/s10040-020-02180-4

Parisouj P, Mohebzadeh H, Lee T (2020) Employing Machine Learning Algorithms for Streamflow Prediction: A Case Study of Four River Basins with Different Climatic Zones in the United States Water Resources Management 34:4113-4131 doi:10.1007/s11269-020-02659-5

Rajaee T, Ebrahimi H, Nourani V (2019) A review of the artificial intelligence methods in groundwater level modeling Journal of Hydrology 572:336-351 doi:https://doi.org/10.1016/j.jhydrol.2018.12.037

Riedmiller MA, Braun H (1993) A direct adaptive method for faster backpropagation learning: the RPROP algorithm IEEE International Conference on Neural Networks:586-591 vol.581

Rousseeuw PJ, Leroy AM (1987) Robust regression and outlier detection. John Wiley & Sons, Inc.,

Santos Finck J, Correa Pedrollo O (2021) Facing Losses of Telemetric Signal in Real Time Forecasting of Water Level using Artificial Neural Networks Water Resources Management doi:10.1007/s11269-021-02782-x

Singh RN, Atkins AS (1985) Analytical techniques for the estimation of mine water inflow International Journal of Mining Engineering 3:65-77 doi:10.1007/BF00881342

Specht DF A general regression neural network

Tayfur G, Aksoy H, Eris E (2020) Prediction of rainfall runoff-induced sediment load from bare land surfaces by generalized regression neural network and empirical model Water and Environment Journal 34:66-76 doi:https://doi.org/10.1111/wej.12442

Xu X, Huang G, Qu Z, Pereira LS (2011) Using MODFLOW and GIS to Assess Changes in Groundwater Dynamics in Response to Water Saving Measures in Irrigation Districts of the Upper Yellow River Basin Water Resources Management 25:2035-2059 doi:10.1007/s11269-011-9793-2
Fig. 1. Study area and location of Gohar-Zamin Iron Ore Mine.

Fig. 2. The architecture of CF neural network.
Fig. 3. The architecture of the proposed CMIS model.

Fig. 4. Cross plot of predicted relative groundwater level versus tentative relative groundwater level for test and train data using 1377 data points.
Fig. 5. The relative error between the tentative and predicted relative groundwater level versus tentative relative groundwater level for train and test data using 1377 data points.

Fig. 6. Prediction of groundwater level for different models using 165 extra data points.
Fig. 7. AARE (%) for the developed CMIS model and best models for different electrical resistivity of sediments ranges.

Fig. 8. AARE (%) for the developed CMIS model and best models for the depth of sediments ranges.
Fig. 9. AARE (%) for the developed CMIS model and best models for electrical resistivity of bedrock ranges.

Fig. 10. AARE (%) for the developed CMIS model and best models for hydraulic conductivity of bedrock ranges.
Fig. 11. AARE (%) for the developed CMIS model and best models for effective porosity of sediments ranges.

Fig. 12. Cumulative frequency versus absolute relative error for all developed models.
**Fig. 13.** Comparing the relative variation with drainage for the proposed CMIS model and other developed models with real data.

**Fig. 14.** Relevancy factor for temporal and spatial parameters on the groundwater level.
Fig. 15. William's plot for the resulting outputs by the proposed CMIS model.
Table 1
Spatial input parameters of the well piezometer over the Gohar Zamin Iron Ore Mine

| Number of well piezometers | UTM:X (m) | UTM:Y (m) | Hydraulic conductivity of sediments (m/day) | Effective porosity of sediments | Electrical resistivity of sediments (Ohm-m) | Depth of sediments (m) | Surface level (m) | Hydraulic conductivity of bedrock (m/day) | Effective porosity of bedrock | Electrical resistivity of bedrock (Ohm-m) | Bedrock level (m) | Fault |
|---------------------------|-----------|-----------|-------------------------------------------|---------------------------------|-------------------------------------------|----------------------|-----------------|------------------------------------------|---------------------------------|------------------------------------------|----------------|-------|
| 1                         | 333,513   | 3,219,470 | 200                                       | 15                              | 14.9                                     | 130.55               | 1724.47         | 19                                        | 0.09                            | 31.6                                     | 1590.37 | ✗     |
| 2                         | 332,489   | 3,219,947 | 187                                       | 20                              | 29.4                                     | 180                  | 1735.77         | 20                                        | 0.07                            | 17.2                                     | 1645.77 | ✗     |
| 3                         | 332,149   | 3,220,733 | 150                                       | 18                              | 55.8                                     | 212                  | 1747.5          | 22                                        | 0.05                            | 5.58                                     | 1641.5  | ✗     |
| 4                         | 332,861   | 3,221,499 | 160                                       | 14                              | 9.69                                     | 139                  | 1733.47         | 19                                        | 0.06                            | 36.5                                     | 1663.97 | ✗     |
| 5                         | 334,438   | 3,220,764 | 315                                       | 10                              | 22.3                                     | 141.65               | 1729.92         | 17                                        | 0.1                             | 13.5                                     | 1659.1  | ✓      |
### Table 2
The statistical explanation of the temporal input parameters

|                | Drainage (m³/day) | Evaporation (m/day) | Rainfall (m/day) |
|----------------|-------------------|---------------------|------------------|
| **Mean**       | 9366              | 0.007               | 0.02             |
| **Median**     | 9183              | 0.006               | 0.006            |
| **Mode**       | 9504              | 0.004               | 0                |
| Skewness       | 0.119             | 0.257               | 1.37             |
| Kurtosis       | 0.245             | -1.345              | 0.862            |
| **Minimum**    | 0                 | 0.002               | 0                |
| **Maximum**    | 17466             | 0.013               | 0.09             |
Table 3

The statistical parameters of the developed models for prediction of groundwater level using 1377 data points

| Statistical parameters | ARE (%) | AARE (%) | SD | RMSE | T (s) |
|------------------------|---------|----------|----|------|-------|
|                        | Training data | Test data | Total | Training data | Test data | Total | Training data | Test data | Total | One iteration |
| MLP-LM                 | -0.00014 | 0.00077  | -0.00095 | 0.01678 | 0.01961 | 0.01751 | 0.00026 | 0.00028 | 0.00027 | 0.43486 | 0.46825 | 0.44701 | 39.8 |
| MLP-BR                 | -0.00054 | 0.00264  | -0.00059 | 0.01785 | 0.01663 | 0.01809 | 0.00029 | 0.00026 | 0.00029 | 0.47691 | 0.43057 | 0.48550 | 25.07 |
| MLP-SCG                | 0.00013  | 0.00372  | -0.0001  | 0.03514 | 0.0389  | 0.03522 | 0.00052 | 0.00053 | 0.00051 | 0.85940 | 0.86466 | 0.84850 | 16.32 |
| MLP-RB                 | -0.00006 | 0.00265  | 0.00046  | 0.01872 | 0.02402 | 0.01992 | 0.00028 | 0.00037 | 0.00030 | 0.46062 | 0.60489 | 0.49080 | 14.36 |
| CF-LM                  | 0.00188  | 0.00609  | 0.00212  | 0.02053 | 0.02591 | 0.02123 | 0.00031 | 0.00040 | 0.00032 | 0.50853 | 0.65183 | 0.52822 | 15.5  |
| CF-BR                  | 0.00046  | -0.00052 | 0.00042  | 0.01741 | 0.02016 | 0.01791 | 0.00026 | 0.00029 | 0.00027 | 0.43614 | 0.48447 | 0.44753 | 16.41 |
| CF-SCG                 | -0.00017 | -0.00414 | -0.00076 | 0.03098 | 0.03671 | 0.03153 | 0.00045 | 0.00052 | 0.00045 | 0.73403 | 0.86044 | 0.74641 | 13.88 |
| CF-RB                  | -0.00001 | 0.00474  | 0.00083  | 0.02404 | 0.02384 | 0.02419 | 0.00036 | 0.00031 | 0.00036 | 0.59830 | 0.52017 | 0.59196 | 17.98 |
| RBF                    | 0.00006  | -0.0182  | -0.00266 | 0.04173 | 0.04892 | 0.04322 | 0.00055 | 0.00063 | 0.00057 | 0.90080 | 1.03346 | 0.93469 | 9.58  |
| GR                     | 0.00077  | 0.004   | 0.00035  | 0.0352  | 0.03805 | 0.03557 | 0.00052 | 0.00054 | 0.00053 | 0.85575 | 0.88188 | 0.86330 | 7.92  |
Table 4

The statistical parameters and information of all developed models for prediction of groundwater level using 165 extra validation data points

| Models  | ARE (%) | AARE (%) | SD      | RMSE   | Function used            | Best architecture |
|---------|---------|----------|---------|--------|--------------------------|------------------|
| MLP-LM  | -0.12983 | 0.16232  | 0.00177 | 2.77969 | Logsig-Purelin-Purelin   | 18-12-8-1        |
| MLP-BR  | -0.37058 | 0.37081  | 0.00396 | 6.20605 | Elliot2sig-Purelin-Purelin | 18-12-8-1       |
| MLP-SCG | -0.08426 | 0.20241  | 0.00230 | 3.60198 | Elliot2sig-Purelin-Purelin | 18-12-8-1       |
| MLP-RB  | -0.30413 | 0.30413  | 0.00338 | 5.29602 | Elliot2sig-Purelin-Purelin | 18-12-8-1       |
| CF-LM   | -0.55160 | 0.55160  | 0.00569 | 8.90091 | Elliot2sig-Purelin-Purelin | 18-6-20-12-1    |
| CF-BR   | -0.08841 | 0.13097  | 0.00160 | 2.50210 | Tansig-Logsig-Purelin     | 18-10-20-12-1   |
| CF-SCG  | -0.14831 | 0.17299  | 0.00195 | 3.05459 | Tansig-Logsig-Purelin     | 18-8-20-12-1    |
| CF-RB   | 0.16633  | 0.18931  | 0.00217 | 3.40278 | Tansig-Logsig-Purelin     | 18-8-11-7-1     |
| RBF     | -0.12491 | 0.12945  | 0.00157 | 2.46439 | Gaussian                 | 18-30-1         |
| GR      | -0.84633 | 0.84632  | 0.00873 | 13.6627 | Gaussian                 | -               |
| CMIS    | -0.10723 | 0.11340  | 0.00141 | 2.22181 | -                        | -               |
Study area and location of Gohar-Zamin Iron Ore Mine. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.
Figure 2

The architecture of CF neural network.

Figure 3

The architecture of the proposed CMIS model.
Figure 4

Cross plot of predicted relative groundwater level versus tentative relative groundwater level for test and train data using 1377 data points.
Figure 5

The relative error between the tentative and predicted relative groundwater level versus tentative relative groundwater level for train and test data using 1377 data points.
Figure 6

Prediction of groundwater level for different models using 165 extra data points.

![Figure 6 graph]

Figure 7

AARE (%) for the developed CMIS model and best models for different electrical resistivity of sediments ranges.

![Figure 7 graph]

Figure 8

AARE (%) for different depths of sediments.
AARE (%) for the developed CMIS model and best models for the depth of sediments ranges.

![Figure 9](image)

**Figure 9**

AARE (%) for the developed CMIS model and best models for electrical resistivity of bedrock ranges.

![Figure 10](image)

**Figure 10**

AARE (%) for the developed CMIS model and best models for hydraulic conductivity of bedrock ranges.
Figure 11
AARE (%) for the developed CMIS model and best models for effective porosity of sediments ranges.

Figure 12
Cumulative frequency versus absolute relative error for all developed models.
Figure 13
Comparing the relative variation with drainage for the proposed CMIS model and other developed models with real data.

Figure 14
Relevancy factor for temporal and spatial parameters on the groundwater level.
Figure 15

William's plot for the resulting outputs by the proposed CMIS model.