Groundwater Level Simulation Using Soft Computing Methods with Emphasis on Major Meteorological Components

Meysam Vadiati (meysam.vadiati@gmail.com)
UC Davis: University of California Davis https://orcid.org/0000-0003-0841-6919
Saeideh Samani
WRI: Water Research Institute
Farahnaz Azizi
Kohgiluyeh and Boyerahmad Regional Water Company
Efat Zamani
Iran water resources management
Ozgur Kisi
Ilia State University

Research Article

Keywords: Soft computing, groundwater level prediction, least-square support vector machine (LSSVM), meteorological component

Posted Date: February 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1337037/v1

Abstract

Precise and accurate estimates of groundwater level might be of great importance for attaining sustainable development goals and integrated water resources management. Compared with alternative numerical models, soft computing methods are promising tools for groundwater level simulation and prediction, which need more hydrogeological and aquifer characteristics. The central aim of this research is to explore the performance of such well-accepted data-driven models to simulate the groundwater level (GWL_t) with emphasis on major meteorological components, including; precipitation (P), temperature (T), evapotranspiration (ET) dataset on a monthly interval. Artificial neural network (ANN), fuzzy logic (FL), adaptive neuro-fuzzy inference system (ANFIS), group method of data handling (GMDH), and least-square support vector machine (LSSVM) are used to predict one-, two-, and three-month ahead groundwater level in an unconfined aquifer. The main meteorological components (T_t, ET_t, P_t, P_{t-1}) and GWL for one, two, and three lag-time (GWL_{t-1}, GWL_{t-2}, GWL_{t-3}) were used as input for different scenarios to attain precise and accurate prediction. The results showed that all models could have the best simulation for one month ahead in scenario 5, comprising inputs of GWL_{t-1}, GWL_{t-2}, GWL_{t-3}, T_t, ET_t, P_t, T_{t-1}, ET_{t-1}, P_{t-1}. Based on different evaluation criteria, all employed models could simulate the groundwater level with a desirable accuracy and precision, and the results of LSSVM were the superior one.

1. Introduction

Precise and accurate groundwater level (GWL) prediction is vital in developing water resources management strategies since they provide reliable quantitative information (Thomas and Gibbons 2018; Wunsch et al., 2020). Recently, numerous studies have explored the GWL prediction using different numerical and data-driven models (Malekzade et al., 2019; Moghaddam et al., 2019; Adedeji et al., 2020; Almuhaylan et al., 2020; Chakraborty 2020; Rezaei et al., 2021). Needing an extensive and uncertain dataset, including hydrogeological, water budget, geophysical, is a drawback for phsics-based models. Such limitations have pushed engineers and researchers to apply data-driven methods in practice. Data-driven modeling utilizes real-time tolerance to model the hydrological events in an inaccurate and uncertain environment (Kisi et al., 2017; Dehghani and Torabi, 2021; Ghazi et al., 2021).

Hydrological and meteorological time series exhibit nonlinear time-dependent behavior, which are too complicated to solve with standard numerical and statistical models (Rajaee et al., 2019). Recently, artificial intelligence (AI)-based methods such as artificial neural networks (ANNs), group method of data handling (GMDH), gene expression programming (GEP), least-square support vector machine (LSSVM), fuzzy logic (FL), adaptive neuro-fuzzy inference system (ANFIS), model tree (MT), multivariate adaptive regression splines (MARS), and evolutionary polynomial regression have been widely employed to simulate GWL (Shiri et al., 2013; Suryanarayana et al., 2014; Kisi et al., 2017; Mohammadrezapour et al., 2018; Rajaee et al., 2019; Roshni et al., 2019, Adedeji et al., 2020; Afzaal et al., 2020; Roshni et al., 2020; Ghazi et al., 2021). These well-accepted models can cope with the complexity of GWL prediction and could relatively provide a better accuracy than numerical models.

Comparison of the several AI-based models in GWL prediction still is highly demanded. In a study, Moghadam et al. (2021) used combinations of parameters including GWL, groundwater withdrawal, recharge, precipitation (P), evapotranspiration (ET), and temperature (T) in 10 scenarios to predict GWL for five observation wells. Results pointed out that the GMDH had a better outcome Bayesian network, and ANN models in evaluating the influential variables in GWL prediction. Yin et al. (2021) assessed GWL in San Joaquin's aquifers using ANN, response surface regression, support vector machine (SVM) models and Bayesian model averaging machine learning ensemble models. Their results showed that ensemble model performance was better than stand-alone AI-based models. Also, the authors figure out that groundwater extraction for agricultural usage is the main driving force for aquifer storage changes. In another study, Shiri et al. (2020) used six AI-based models, ANN, BT, MARS, RF, GEP, and SVM, in a coastal aquifer to forecast GWL, and they figured out that GEP's outcomes were the superior one. Osman et al.'s (2021) study showed that the Xgboost model had the best results among other used AI-based models such as ANN and support vector regression to predict GWL. A brief detail of studies regarding applying the AI-based models for GWL prediction is presented in Table 1. The ANN model is the most common AI-based model for GWL prediction based on Table 1,

Table 1
Earlier research of AI-based methods for GWL prediction.

Reference	Models						Input Variables	Lead Time	Time Interval
	ANN-based	FI	ANFIS	GMDH	SVM	Other methods			
Gong et al. (2015)	✓		✓		✓		GWL,SWL, P, T	GWL_{t+1}, GWL_{t+2}, GWL_{t+3}	Monthly
Nourani and Mousavi (2016)	✓		✓				GWL, P, T, Q	GWL_{t+1}	Monthly
Wen et al. (2017)	✓				✓	Wavelet-ANN	GWL, P, E, T	GWL_{t+2}, GWL_{t+3}	Monthly
Huang et al. (2017)	✓						GWL	GWL	Daily, Weekly, Monthly
Kouziokas et al. (2018)	✓						P, T, M	GWL	Daily
Roshni et al. (2019)	✓						GWL, P, T, Q	GWL_{t+6}, GWL_{t+12}	Weekly
Derbela and Nouiri (2020)	✓						GWL, P, E	GWL_{t+2}	Monthly
Naganna et al. (2020)			✓	✓		GTB	GWL	GWL_{t+1}	Monthly
Sahu et al. (2020)	✓						GWL, P, T, SWL	GWL_{t+1}, GWL_{t+2}	Monthly
Moosavi et al. (2021)				✓		Wavelet-GMDH	GWL, P, T, SWL, D	GWL_{t+2}	Monthly
Razzagh et al. (2021)	✓	✓	✓				GWL, P, T, D	GWL_{t+1}	Monthly
Current study	✓	✓	✓	✓	✓		GWL, P, T, Et	GWL_{t+1}, GWL_{t+2}, GWL_{t+3}	Monthly
Models									
FL: Fuzzy Logic; **GTB**: Gradient Tree Boosting; **ANFIS**: Adaptive Neuro Fuzzy Inference System; **ANN**: Artificial Neural Network; **GMDH**: Group Method of Data Handling; **LSSVM**: Least Square Support Vector Machines.									
Input Variables: GWL: Groundwater Level; **Q**: River Flow; **SWL**: Surface Water Level; **P**: Precipitation; **T**: Temperature; **ET**: Evapotranspiration; **M**: Moisture; **E**: Evaporation; **D**: Exploitation Well Discharge.									

[Please insert Table 1 **here]**

The ANN has been recently utilized for GWL prediction (e.g., Zare and Kooh 2018; Banadkooki et al., 2020; Dehghani and Torabi 2021). Likewise, ANFIS and SVM have been applied to predict GWL and indicated an improvement in accuracy and precision compared to ANN in GWL prediction (Emamgholizadeh et al., 2014; Kasiviswanathan et al., 2016; Khedri et al., 2020). Even though the ANNs, SVMs, and ANFIS have been commonly employed in GWL simulation whereas the usage of the GMDH model has seldomly been investigated on groundwater modeling procedures. However, this method has been successfully applied in civil engineering, water quality management, and soil science (Najafzadeh et al., 2013; Rahmati 2017; Mehri et al., 2019; Tayebi et al., 2019; Lin et al., 2020). One of the motivation of this study is to assess the viability of GMDH model in the GWL prediction.

The present study evaluates the ability of various AI-based models in GWL prediction. The main aim to conduct this research could be summarized as a) modeling behavior of an rising GWL in a monitoring well while the other parts of the aquifer demonstrate a severe declining GWL; b) predicting GWL at the aquifer scale using monthly GWL, P, T, ET dataset; and c) comparing efficiency of the FL, ANFIS, ANN, GMDH and LSSVM models in GWL prediction. The present study sheds light on the GWL modeling in aquifers with poor hydrological and hydrogeological datasets. The outcomes of this sort of AI-based models provide a reliable perspective for decision-makers to attain sustainable water resources management goals. Figure 1 shows the procedural outline of the applied AI-based models.

[Please insert Fig. 1 **here]**

2. Methods

2.1. Artificial Neural Network (ANN)

The ANN computational approach is biologically inspired by the human brain (Patel et al., 2022). This model seems like the brain in two phases: (a) knowledge is obtained by the network from its environment as a result of a learning procedure, and (b) interneuron connection strengths are used to collect the obtained knowledge (Haykin 2004). The ANN design procedure comprises five stages: selecting inputs, selecting an appropriate architecture, neural network construction, training and testing procedure, and finally evaluating the developed model (Sahoo and Jha 2013). The multilayer perceptron (MLP) ANN, as the most widely used kind of ANN in hydrological studies, was used in this study (McGarry et al., 1999). The general framework of MLP comprises three layers (input, hidden and output). The number of layers and consequently the number of neurons in each layer is essential to reach an optimum model structure. One hidden layer, which is sufficient based on previous studies for GWL prediction, was used in the ANN model.

Krishna et al. (2008) evaluated various training algorithms for GWL prediction in a coastal aquifer. They concluded that the Levenberg-Marquardt (LM) algorithm has been the best learning algorithm compared to Bayesian regularization and scaled conjugate gradient. The LM, the fast and most popular algorithm in GWL prediction, was used in the present study (Adamowski and Karapataki 2010; Khaki et al., 2015). MATLAB® (Mathworks 2014) software was employed to develop AI-based models in this study. The overall framework of AI-based models is presented in Fig. 2.

[Please insert Fig. 2 **here]**

2.2. Fuzzy Logic (FL)

FL models can overcome the intrinsic uncertainty between defined sets in mathematical form (Zadeh 1965). A fuzzy controller comprises three basic processes: fuzzification, inference, and defuzzification (Bai et al., 2006). The fuzzification step involves transforming a crisp dataset into a fuzzy dataset or membership function (MF). The fuzzy inference process (FIS) combines MFs and fuzzy if-then rules to achieve the fuzzy output. The most useful FISs in water resources, Mamdani, Sugeno, and Tsukamoto, differ in aggregation and defuzzification. The defuzzification procedure converts the fuzzy outputs to crips results based on a fuzzy rule base system. In this study, genfis-2 was applied to develop the FL model, which generates a Sugeno-type FIS structure using subtractive clustering and requires clustering radius as input parameters.The clustering radius for genfis-2 fuzzy logic was investigated based on trial and error to optimize the FIS structure, which varies between 0.2 and 0.9. This parameter determines the number of clusters and fuzzy inference system rules. The smaller radius number gives the model fewer clusters and rules (Chiu 1994). Therefore, the optimum cluster number optimizes the fuzzy model structure.

2.3. Adaptive Neuro Fuzzy Inference System (ANFIS)

The ANFIS is a single structure that can capture the benefits of the adaptive neural network and the fuzzy inference system (FIS) methods (Jang, 1993). The ANFIS is an AI-based model with a flexible statistical structure that can identify complicated nonlinearity and uncertainties due to vagueness and randomness between variables without trying to achieve an insight into the nature of the events. FIS can be categorized into Mamdani or Sugeno systems, while ANFIS models are based on the Sugeno system. Various Sugeno FIS models could be developed using subtractive clustering (SC), grid partitioning, and C–means clustering approaches for a given input-output dataset.

The ANFIS structure utilized in this present study is based on the Sugeno fuzzy model that consists of five layers (Fig. 2):

Layer 1: **Fuzzy Membership**, The most frequently applied MFs are: Triangular, Trapezoidal, Gaussian, Two-sided Gaussian, Generalized Bell, and Sigmoidal Z- and S-functions (Nguyen et al., 2003). There is no typical rule to find the optimum number of MFs in the ANFIS model, and the large number of MFS is commonly avoided due to increasing calculation time (Keskin et al., 2004). According to Shiri and Kisi (2011), two, three, or four MFs are enough to predict GWL.

Layer 2: **Fuzzification**, this layer utilizes a fuzzification interface to convert the crisp input dataset into levels of belongingness with linguistic values,

Layer 3: **Normalization**,

Layer 4: **Defuzzification**, converts the fuzzy outputs of the interface to a crisp output, and

Layer 5: **Output** (Jang 1993).

The SC is used in this research paper to divide an input space into n-subdivided particular areas by evaluating n-dimensional input data to produce clusters. The cluster radius, which is a number between 0 and 1, optimizes the range of influence from the cluster centroid. The number of clusters and then the number of if-then fuzzy rules increased when the cluster radius was set small since the size of the clusters came to be small (Chiu 1994; Yager and Filev 1994). Identifying cluster radius is a critical element in determining the number of clusters. So, choosing a proper influential radius is essential to cluster the dataset. The optimum cluster radius for the ANFIS method in this paper was chosen by trial and error approach for every single input-output combination.

2.4. Group Method of Data Handling (GMDH)

Ivakhnenko (1968) proposed GMDH to solve complicated and nonlinear problems. This algorithm generates a self-organizing model (SOM) to solve classification, prediction, and other system questions. The number of neurons, hidden layers, influential input variables, and network framework is necessarily defined in the GMDH model. GMDH as a polynomial neural network is so similar to ANNs. Based on Muller et al. (1998), ANNs, statistical analysis, and statistical neural networks are deductive techniques that cannot detect complicated objects since they require a sizeable amount of a priori info. Instead, GMDH algorithms are considered a regression-based technique that combines the best of both neural networks and statistic analysis while embedding the additional fundamental property of induction (Lemke 1997). Hence, GMDH algorithms can overcome the shortfalls of neural networks, while statistical neural networks can somewhat resolve them. Based on the GMDH methodology, all model structure features like the neuron and layer numbers can be defined by default. The detailed information about the GMDH method could be obtained from Nariman-Zadeh et al. (2002).

2.5. Least Square Support Vector Machine (LSSVM)

Fundamental concepts of SVM and its theory have been proposed by Vapnik (1998). The broad overview capability of the SVM is deemed better than ANN because it is on the basis of structural risk minimization while the ANN uses experimental risk minimization. The primary procedure of the SVM model comprises support vectors selections that support the model framework and define their weights. A complete mathematical outline of SVM was proposed by Vapnik (1998). The LSSVM model was established by Suykens and Vandewalle (1999), based on the SVM model. It is a robust technique for resolving function estimation, nonlinear classification, and density estimation problems. LSSVM resolves one of the linear programming problems by adjusting inequality constraints in the SVM method to equality constraints (Kumar and Kar 2009; Kisi 2013). Furthermore, the LSSVM model demonstrates a superior outcome than the SVM in running fast training (Gu et al., 2010).

Various algorithms have been suggested to resolve the dual optimization problem of SVMs. The recent SVM learning algorithm is known as Sequential Minimal Optimization (SMO). SMO utilizes an analytical QP phase (Platt 1999) and as a straightforward algorithm, an SMO algorithm is able to instantly resolve the SVM problem without the necessity to use a quadratic optimizer and without any additional matrix space, which was utilized in this study.

The outcome of the LSSVM model depends firmly on the suitable choice of the kernel function and adjusting correct C and γ parameters. The present study used the polynomial kernel function for the LSSVM model because of its superior results in GWL prediction based on the used dataset in the study area. The trial-and-error procedure was applied to get the optimal parameters of the SVM model (Suryanarayana et al., 2014). The LSSVM modeling procedures in the present study were applied by Support Vector Machines (LIBSVM) library codes presented by Chang and Lin (2011).

3. Study Area

Qazvin Aquifer at the east of Ghazvin province, one of Iran's most important aquifers for agricultural purposes, was chosen as the study area. The absence of a stable river in this study area has been created a significant reliance on groundwater resources for supplying water demands in agricultural, domestic, and industrial usages and has been the leading reason for groundwater level drop rating to one meter yearly. Declining GWLs in the Qazvin Aquifer have caused a negative water budget balance of 300 million cubic meters and deteriorating groundwater quality.

[Please insert Fig. 3 **here]**

The study area map is presented in Fig. 3. Mean annual precipitation and temperature of the study area are 330 mm, and 12° C, respectively, and elevation varies between 1000 and 3000 m above mean sea level.

4. Model Development

To predict the GWL for one-, two-, and three-month ahead over Ghazvin Aquifer, monthly GWL, average monthly temperature (T), monthly precipitation (P), and monthly evapotranspiration (ET) data were considered. In the present study, GWL changes were explored in an observation well with a rising groundwater trend, showing a different behavior related to severe declining groundwater levels in the entire aquifer. The Ghazvin Regional Water Authority issued the monthly groundwater levels for 15 years from 2005 to 2020. A monthly time interval has been deemed the most appropriate interval for GWL prediction (Nourani and Mousavi 2016). To assess the potential of the used models in predicting the GWL over the Ghazvin Aquifer, the dataset is split into training and testing datasets (70% and 30% of total data, respectively).

5. Model Implementation

The input-output dataset undertook a normalization procedure to leave out dimension effects. GWL changes in the Qazvin Aquifer strongly rely on hydro-meteorological changes. Hence, meteorological parameters as an auxiliary dataset along with GWL data were utilized to predict GWL in Gazvin Aquifer.

Various input combinations are assessed utilizing the predictive variables with various lag intervals from one-month "GWL_{t-1}" to three-month prior "GWL_{t-3}" for GWL prediction with different lead times (one- to three-month ahead). In the present study, optimal input combinations chosen based on correlations among the inputs and GWL are given as follows:

(1) GWL_{t-1}, GWL_{t-2}, GWL_{t-3};

(2) GWL_{t-1}, GWL_{t-2}, GWL_{t-3}, T_t, ET_t, P_t;

(3) GWL_{t-1}, GWL_{t-2}, GWL_{t-3}, T_t, ET_t, P_t, P_{t-1};

(4) GWL_{t-1}, GWL_{t-2}, GWL_{t-3}, T_t, ET_t, P_t, P_{t-1}, ET_{t-1};

(5) GWL_{t-1}, GWL_{t-2}, GWL_{t-3}, T_t, ET_t, P_t, P_{t-1}, ET_{t-1}, T_{t-1}.

GWL_{t-1}, GWL_{t-2}, GWL_{t-3} are GWLs with various lag time from the one-month "t-1" to three-month "t-3"; T_t, ET_t and P_t are the temperature, evapotranspiration and precipitation at the current month and vice versa. The mentioned combinations are employed to achieve the most optimum prediction for every lead time (GWL_{t+1}, GWL_{t+2}, GWL_{t+3}).

6. Efficiency Criteria

Different criteria can assess the ability of AI-based models. In this study, various statistical criteria were utilized to assess the effectiveness of the methods, comprising the Correlation Coefficient (R), Nash–Sutcliffe efficiency (NSE), mean absolute error (MAE), and root means squared error (RMSE). The closer the value of R and NSE to one, the higher the estimation capability of the model will be, and vice versa. The values of MAE and RMSE close to zero indicate better model efficiency.

7. Results And Discussion

7.1. Results of the ANN Model

The structural design of the ANN model is the significant and critical stage of the modeling since an improper model's structure can cause under/over-fitting and computational overload problems. Three-layered ANN containing input, hidden, and output layers were considered to predict the study area's seasonal GWL. Based on different combinations, designed ANN models were trained and then tested to predict GWL_{t+1}, GWL_{t+2}, and GWL_{t+3}. Preliminary findings of the present study showed that one hidden layer was enough to find a relationship between GWL and the other predictor inputs. Overall, a trial and error process was used to define the neuron numbers in the hidden layer. The number of neurons for the hidden layer was changed where considerable progress was no longer seen to select the suitable neuron number for hidden layer. For instance, the optimum number of nodes in the hidden layer for input combinationfive was identified as 6 (Table 2).

Table 2
The results of the various combinations utilizing the AI-based methods in predicting GWL_{t+1}, GWL_{t+2}, GWL_{t+3}.

	Input	Hidden Layers	Step	Ht + 1				Ht + 2				Ht + 3			
				R	RMSE	MAE	NS	R	RMSE	MAE	NS	R	RMSE	MAE	NS
ANN	Combination 1	2	Training	0.92	1.11	0.89	0.82	0.86	1.32	1.00	0.73	0.77	1.62	1.26	0.59
			Test	0.92	0.70	0.58	0.76	0.88	0.74	0.63	0.75	0.76	1.04	0.88	0.55
	Combination 2	3	Training	0.93	0.93	0.64	0.87	0.88	1.25	0.95	0.77	0.80	1.53	1.25	0.64
			Test	0.93	0.54	0.45	0.86	0.88	0.74	0.57	0.76	0.85	0.83	0.64	0.71
	Combination 3	4	Training	0.95	0.83	0.60	0.90	0.89	1.16	0.88	0.80	0.85	1.34	1.00	0.72
			Test	0.91	0.61	0.45	0.82	0.86	0.79	0.63	0.72	0.76	1.05	0.85	0.54
	Combination 4	5	Training	0.95	0.83	0.63	0.90	0.90	1.13	0.88	0.81	0.85	1.34	1.03	0.72
			Test	0.92	0.56	0.41	0.85	0.86	0.82	0.64	0.70	0.81	0.96	0.80	0.62
	Combination 5	6	Training	0.96	0.75	0.56	0.92	0.92	1.01	0.79	0.84	0.88	1.19	0.91	0.78
			Test	0.92	0.56	0.44	0.85	0.83	0.84	0.70	0.68	0.75	1.06	0.86	0.53
FL	Combination 1	0.8	Training	0.94	0.84	0.62	0.89	0.87	1.28	0.95	0.75	0.82	1.46	1.12	0.67
			Test	0.94	0.56	0.44	0.84	0.86	0.78	0.66	0.69	0.69	1.07	0.85	0.44
	Combination 2	0.8	Training	0.96	0.75	0.55	0.91	0.90	1.10	0.86	0.82	0.87	1.25	1.01	0.76
			Test	0.91	0.70	0.49	0.75	0.81	0.97	0.74	0.51	0.69	1.29	0.99	0.19
	Combination 3	0.9	Training	0.96	0.72	0.53	0.92	0.91	1.06	0.79	0.82	0.87	1.20	0.90	0.76
			Test	0.83	0.86	0.61	0.63	0.79	1.07	0.78	0.45	0.69	1.41	1.05	0.12
	Combination 4	0.9	Training	0.96	0.73	0.54	0.92	0.91	1.01	0.77	0.84	0.89	1.14	0.88	0.79
			Test	0.86	0.77	0.54	0.70	0.79	1.05	0.76	0.48	0.62	1.54	1.16	-0.05
	Combination 5	0.9	Training	0.97	0.64	0.48	0.94	0.94	0.89	0.67	0.88	0.92	0.99	0.77	0.84
			Test	0.86	0.72	0.55	0.72	0.76	1.07	0.85	0.41	0.66	1.46	1.23	0.05
ANFIS	Combination 1	0.8	Training	0.94	0.84	0.62	0.89	0.87	1.25	0.93	0.75	0.79	1.53	1.17	0.62
			Test	0.94	0.56	0.44	0.84	0.89	0.80	0.68	0.69	0.80	1.08	0.91	0.48
	Combination 2	0.8	Training	0.96	0.75	0.55	0.91	0.90	1.08	0.83	0.81	0.86	1.25	0.98	0.74
			Test	0.91	0.69	0.50	0.76	0.87	0.89	0.65	0.63	0.65	1.47	1.07	0.04
	Combination 3	0.9	Training	0.96	0.72	0.53	0.92	0.91	1.04	0.78	0.83	0.89	1.14	0.86	0.79
			Test	0.84	0.84	0.60	0.64	0.79	1.08	0.78	0.45	0.62	1.52	1.14	-0.02
	Combination 4	0.9	Training	0.96	0.73	0.54	0.92	0.91	1.02	0.77	0.83	0.88	1.15	0.89	0.78
			Test	0.87	0.76	0.55	0.70	0.79	1.04	0.76	0.48	0.70	1.35	1.03	0.19
	Combination 5	0.9	Training	0.97	0.60	0.44	0.94	0.95	0.81	0.59	0.89	0.93	0.92	0.71	0.86
			Test	0.88	0.74	0.58	0.72	0.65	1.33	1.03	0.16	0.51	2.33	1.69	-1.41
GMDH	Combination 1	15,15,1	Training	0.94	0.79	0.59	0.88	0.87	1.16	0.86	0.76	0.72	1.62	1.24	0.52
			Test	0.93	1.00	0.70	0.85	0.85	1.22	0.95	0.72	0.56	1.96	1.55	0.14
	Combination 2	15,15,1	Training	0.94	0.78	0.57	0.89	0.87	1.11	0.85	0.76	0.84	1.31	1.07	0.70
			Test	0.94	0.85	0.62	0.88	0.87	1.30	1.05	0.76	0.80	1.33	1.03	0.63
	Combination 3	15,15,1	Training	0.95	0.75	0.54	0.91	0.89	1.05	0.83	0.80	0.84	1.26	0.93	0.71
			Test	0.89	0.83	0.59	0.80	0.83	1.40	1.09	0.63	0.78	1.59	1.24	0.55
	Combination 4	15,15,1	Training	0.95	0.76	0.53	0.91	0.89	1.08	0.83	0.80	0.87	1.15	0.84	0.75
			Test	0.93	0.78	0.58	0.86	0.81	1.40	1.06	0.63	0.75	1.62	1.30	0.56

	Combination 5	15,15,1	Training	0.95	0.73	0.53	0.91	0.89	1.05	0.76	0.80	0.87	1.17	0.88	0.75
			Test	0.94	0.74	0.52	0.88	0.89	1.12	0.86	0.79	0.81	1.38	1.12	0.66
LSSVM	Combination 1	sig2 = [0.2,1], 5	Training	0.98	0.50	0.35	0.96	0.94	0.82	0.62	0.89	0.87	1.23	0.94	0.75
			Test	0.97	0.35	0.23	0.94	0.94	0.55	0.42	0.86	0.90	0.76	0.61	0.74
	Combination 2	sig2 = [0.2,1], 5	Training	0.98	0.50	0.34	0.96	0.95	0.79	0.57	0.90	0.89	1.12	0.84	0.79
			Test	0.97	0.35	0.22	0.94	0.94	0.53	0.36	0.87	0.88	0.79	0.58	0.72
	Combination 3	sig2 = [0.2,1], 5	Training	0.98	0.49	0.34	0.96	0.95	0.75	0.55	0.91	0.90	1.06	0.80	0.82
			Test	0.98	0.32	0.22	0.95	0.95	0.51	0.40	0.88	0.88	0.80	0.67	0.71
	Combination 4	sig2 = [0.2,1], 5	Training	0.98	0.49	0.34	0.96	0.95	0.75	0.55	0.91	0.90	1.06	0.80	0.82
			Test	0.97	0.37	0.26	0.93	0.95	0.51	0.40	0.88	0.88	0.80	0.67	0.71
	Combination 5	sig2 = [0.2,1], 5	Training	0.98	0.47	0.34	0.97	0.96	0.72	0.54	0.92	0.91	1.02	0.77	0.83
			Test	0.98	0.32	0.22	0.95	0.93	0.56	0.45	0.85	0.83	0.93	0.75	0.62

[Please insert Table 2 **here]**

Statistical assessment of the ANN model's results demonstrates that Combination 5 as input and GWL_{t+1} as output is reasonable (Table 2). In fact, RMSE and MAE are low, R and NSE are close to 1 for the observation well dataset. This might be owing to the increasing input variables in combination 5, improving model performance. Performance of ANN deteriorated after 1-month ahead of prediction. For the best combination of ANN, FL, ANFIS, GMDH, and LSSVM methods, the scattering curves and time-variation charts among the observed and simulated data were employed to compare different models function (Fig. 4).

[Please insert Fig. 4 **here]**

7.2. Results of the Fuzzy Logic (FL) Model

FL Model was applied for all five combinations to predict GWLs in the observation well in training and testing steps with one-, two- and three-month ahead (Table 2). The range of the radius parameter altered between 0.2 and 0.9 by trial-and-error approach to find the minimum RMSE and MAE between observed and simulated GWL. The optimal parameter radius was 0.8 for combinations 1 and 2 and it was 0.9 for combination 3, 4 and 5.

The result of the FL models shows that combination 5 indicated high ability in the training step but, the model ability was not reasonable in the testing step based on the values of RMSE, MAE, R, and NSE (see Table 2).

7.3. Results of the ANFIS Model

The ANFIS model was also employed for the temporal GWL prediction using different input combinations. The structure of the ANFIS model was chosen by trial and error methodology for every input combination. An appropriate cluster radius as a predesigned interior ANFIS component was efficient, making it feasible for the ANFIS method to attain the efficiency objective. Commonly, smaller radii cause many small clusters and have numerous rules; however, large radii results in a few large clusters in the dataset (having fewer rules) (Sanikhani and Kisi 2012). Table 2 illustrates the evaluation criteria of GWL prediction for the ANFIS method. Based on the ANFIS results, the model with combination five as input and GWL_{t+1} as output in the training step could present the best results for GWL prediction than other combinations. However, the model ability is not reasonable in the testing step based on the values of RMSE, MAE, R, and NSE (see Table 2).

7.4. Results of the GMDH model

The GMDH model, as an intelligent tool, showed promising results to predict fluctuations in GWL for one-, two- and three-time horizons utilizing five different combinations. Based on the model results, a GMDH model structure with four layers and 15 neurons was considered for GWL prediction. The model results indicated that the model attained the desired outcomes in the fourth layer, with ten neurons in the first layer, 15 neurons in the second and third layers, and one in the fourth layer. Table 2 illustrates the evaluation criteria of GWL prediction using the GMDH model for the selected observation well. In this method, combination five was indicated as a suitable input dataset to predict the GWL_{t+1}. Figure 4 displays observed and simulated GWL results produced by the GMDH model.

7.5. Results of the LSSVM Model

Similarly, employing different input combinations, the LSSVM model was likewise used to predict three-month ahead GWL. In the LSSVM modeling, the appropriate choice of the kernel function and parameter values (C and γ) is significant. In this study, the C= [0.2, 1] and γ = 5 are determined as optimum parameters by trial and error procedure and considering the least possible prediction error. The polynomial kernel function was chosen to represent the resemblance of vectors in the training dataset in a feature space over polynomials of the initial dataset.

Table 2 shows the MSE, MAE, R, and NSE for various LSSVM model structures. This model shows that combination five can achieve accurate and reliable prediction results for one-month ahead GWL. Figure 4 illustrates the observed and simulated GWL using the LSSVM model.

7.6. Comparison of the Different Models

The statistical criteria for the optimum input combination (i.e., combination five) and one-month lead time were assessed to evaluate models' performance and explore the best method. Performance measures of these black-box methods indicate that the values of the evaluation criteria did not vary significantly, and the whole method demonstrated satisfactory results in GWL prediction in Qazvin Plain.

A model is supposed to be ideal with the optimized results if the NSE criterion on the estimated values is very close to 1 or the value of NSE is more significant than 0.8 (Moriasi et al., 2015). Based on Table 2, it is apparent that all methods at the training step provide enough precision for GWL prediction with NSE greater than 0.8. However, the superior performance prediction criteria are seen in the LSSVM model based on NSE values. Based on RMSE and R values, the LSSVM method also demonstrates the best groundwater prediction precision. In fact, the low RMSE and high R values in the LSSVM model represent that the GWL prediction using the LSSVM model is precise for the study area.

The models' accuracy is further compared graphically in Fig. 4 in the form of a time variation graph and scatterplot. From the graphs provided in the first columns, we can see the detailed variation of models' predictions and observed ones, and the graphs given in the second column shows how each models' predictions are scattering and fit line equations, and R2 values give information about the fitting accuracy of the models. From the hydrographs and scatterplots, it is apparent that the simulations of LSSVM models are closely following the observed GWL values and less scattered than the other four models. The deviations between simulations and observed values are clearly seen for the ANN, FL, ANFIS and GMDH models. LSSVM can not also catch some extreme GWL values, and this can be explained by the limited amount of samples since we study with monthly time intervals.

Additionally, one of the significant attributes of the applied models in GWL prediction is providing the most important statistics of the observed GWL, i.e., minimum, maximum, mean, median, upper and lower quintiles. In Fig. 5, the box plots for the GWL changes are presented. The chart for one month ahead and combination 2 (Fig. 5a) indicates that the GMDH model is consistent with the observed maximum groundwater level fluctuation.

Likewise, the FL methods have the least compatibility. A similar inference can be drawn for the minimum changes of the observed GWL. The results indicate that the GMDH method cannot be suitable enough to predict maximum and minimum values for two and three-month ahead. For two and 3-month ahead, the LSSVM method outperformed other methods in predicting the main statistics for combination 2 (Fig. 5b,c).

[Please insert Fig. 5 **here]**

In brief, the present study indicates the superiority of the LSSVM method in terms of all evaluation criteria to other methods. However, this study demonstrated that all models can predict short-term GWL and that the higher the number of influential dependent variables, the better the network's performance. These results reinforce earlier outcomes in different studies (Miraki et al., 2019; Mirarabi et al., 2019; Nadiri et al., 2017; Guzman et al., 2019; Yin et al., 2021).

Conclusions

In the present study, a well-accepted range of AI-based models was used to predict GWL with compelling precision and accuracy. The methodology assumed that groundwater dynamics are generally dominated by hydrogeological and meteorological factors like monthly groundwater level, precipitation, temperature, and evapotranspiration. The models were trained (calibrated) and tested (verified) using an observation GWL in the Qazvin Aquifer to assess the performance of the developed methods. Different combinations with 3, 6, 7, 8, and 9 antecedent inputs comprising GWL_{t-1}, GWL_{t-2}, GWL_{t-3}, T_t, ET_t, P_t, T_{t-1}, ET_{t-1}, P_{t-1} were explored for GWL prediction with different lead times (one to three months ahead). The performances of the various methods were explored through statistical indices (R, RMSE, MAE, and NSE) to recognize the superior method that can simulate the increasing trend of the GWL and provide a reasonable prediction. Four statistical indicators related to predictive efficacy showed that the LSSVM methods had the best precision in the GWL prediction, although all methods can yield convincing results to predict GWL. The present study's findings showed that all the models achieved satisfactory results for one- and two-month ahead. However, three months ahead, the performance of the models was not satisfactory enough. The results also showed that increasing the number of input variables from 3 to 9 considerably increased the accuracy and precision of the model's results.

Abbreviations

AI: Artificial Intelligence

ANFIS: Adaptive Neuro-Fuzzy Inference System

ANN: Artificial Neural Network

FL: Fuzzy Logic

GEP: Gene Expression Programming

GMDH: Group Method of Data Handling

GTB: Gradient Tree Boosting

GWL: Groundwater Level

K-NN: k-nearest neighbor

LM: Levenberg-Marquardt

LSSVM: Least Square Support Vector Machine

MAE: Mean Absolute Error

MARS: Multivariate Adaptive Regression Splines

MF: Membership Functions

MLP: Multilayer Perceptron

MT: Model Tree

NSE: Nash–Sutcliffe efficiency

QP: quadratic programming

R: Correlation Coefficient

RF: Random Forest

RMSE: Root Means Squared Error

SCG: Scaled Conjugate Gradient

SMO: Sequential Minimal Optimization

Declarations

Acknowledgments

The authors acknowledge the Qazvin Regional Water Authority for providing part of the data.

Ethical Approval

Not applicable because this article does not contain any studies with human or animal subjects.

Consent to Participate

Not applicable

Consent to Publish

We, the undersigned, give us consent to publish identifiable details, including text, figures, and tables in the water resource management journal.

Authors Contributions

S.Smani and M. Vadiati analyzed and interpreted data and contributed to writing the manuscript. E.Zamani and Farahnaz Azizi collected data and had contributed to drafting manuscript preparation. O.Kisi was involved in revising the manuscript critically for important intellectual content. All authors read and approved the final manuscript.

Funding

This research received no specific grant from any funding agency.

Competing Interests

The authors declare that they have no competing interests

Availability of data and materials

The data, models, and codes generated or used during the study are available from the corresponding author by request.

Conflict of interest

The authors declare no conflict of interest

References

1. Adamowski, J., & Karapataki, C. (2010). Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms. Journal of Hydrologic Engineering, 15(10), 729-743.
2. Adedeji, P. A., Akinlabi, S., Madushele, N., & Olatunji, O. O. (2020). Wind turbine power output very short-term forecast: a comparative study of data clustering techniques in a PSO-ANFIS model. Journal of Cleaner Production, 254, 120135.
3. Afzaal, H., Farooque, A. A., Abbas, F., Acharya, B., & Esau, T. (2020). Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning. Water, 12(1), 5.
4. Almuhaylan, M. R., Ghumman, A. R., Al-Salamah, I. S., Ahmad, A., Ghazaw, Y. M., Haider, H., & Shafiquzzaman, M. (2020). Evaluating the Impacts of Pumping on Aquifer Depletion in Arid Regions Using MODFLOW, ANFIS and ANN. Water, 12(8), 2297.
5. Bai, Y., & Wang, D. (2006). Fundamentals of fuzzy logic control—fuzzy sets, fuzzy rules and defuzzifications. In Advanced fuzzy logic technologies in industrial applications (pp. 17-36). Springer, London.
6. Banadkooki, F. B., Ehteram, M., Ahmed, A. N., Teo, F. Y., Fai, C. M., Afan, H. A., ... & 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.
7. Chakraborty, S., Maity, P. K., & Das, S. (2020). Investigation, simulation, identification and prediction of groundwater levels in coastal areas of Purba Midnapur, India, using MODFLOW. Environment, Development and Sustainability, 22(4), 3805-3837.
8. Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. ACM transactions on intelligent systems and technology (TIST), 2(3), 1-27.
9. Chang, J., Wang, G., & Mao, T. (2015). Simulation and prediction of suprapermafrost groundwater level variation in response to climate change using a neural network model. Journal of Hydrology, 529, 1211-1220.
10. Chiu, S. L. (1994). Fuzzy model identification based on cluster estimation. Journal of Intelligent & fuzzy systems, 2(3), 267-278.
11. Dehghani, R., & Torabi Poudeh, H. (2021). Application of novel hybrid artificial intelligence algorithms to groundwater simulation. International Journal of Environmental Science and Technology, 1-18.
12. Emamgholizadeh, S., Moslemi, K., & Karami, G. (2014). Prediction the groundwater level of bastam plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS). Water resources management, 28(15), 5433-5446.
13. Ghazi, B., Jeihouni, E., & Kalantari, Z. (2021). Predicting groundwater level fluctuations under climate change scenarios for Tasuj plain, Iran. Arabian Journal of Geosciences, 14(2), 1-12.
14. Gong, Y., Zhang, Y., Lan, S., & Wang, H. (2016). A comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near Lake Okeechobee, Florida. Water resources management, 30(1), 375-391.
15. Gu, Y., Zhao, W., & Wu, Z. (2010). Least squares support vector machine algorithm [J]. Journal of Tsinghua University (science and technology), 7, 1063-1066.
16. Guzman, S. M., Paz, J. O., Tagert, M. L. M., & Mercer, A. E. (2019). Evaluation of seasonally classified inputs for the prediction of daily groundwater levels: NARX networks vs support vector machines. Environmental Modeling & Assessment, 24(2), 223-234.
17. Haykin, S. (2004). Neural Networks: A Comprehensive Foundation. Prentice Hall, New Jersey.
18. Huang, F., Huang, J., Jiang, S. H., & Zhou, C. (2017). Prediction of groundwater levels using evidence of chaos and support vector machine. Journal of Hydroinformatics, 19(4), 586-606.

19. Ibrahem Ahmed Osman, A., Najah, A. A., Chow, M., Feng, H. Y., & El-Shafie, A. (2021). Extreme gradient boosting (Xgboost) model to predict thegroundwater levels in Selangor Malaysia. Ain Shams Eng. J.
20. Ivakhnenko, A. G. (1968). The group method of data of handling; a rival of the method of stochastic approximation. Soviet Automatic Control, 13, 43-55.
21. Ivakhnenko, A. G. (1970). Heuristic self-organization in problems of engineering cybernetics. Automatica, 6(2), 207-219.
22. Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. IEEE transactions on systems, man, and cybernetics, 23(3), 665-685.
23. Kasiviswanathan, K. S., Saravanan, S., Balamurugan, M., & Saravanan, K. (2016). Genetic programming based monthly groundwater level forecast models with uncertainty quantification. Modeling Earth Systems and Environment, 2(1), 27.
24. Khaki, M., Yusoff, I., & Islami, N. (2015). Simulation of groundwater level through artificial intelligence system. Environmental Earth Sciences, 73(12), 8357-8367.
25. 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.
26. Kisi, O. (2013). Least squares support vector machine for modeling daily reference evapotranspiration. Irrigation Science, 31(4), 611-619.
27. Kisi, O., Alizamir, M., & Zounemat-Kermani, M. (2017). Modeling groundwater fluctuations by three different evolutionary neural network techniques using hydroclimatic data. Natural Hazards, 87(1), 367.
28. Kouziokas, G. N., Chatzigeorgiou, A., & Perakis, K. (2018). Multilayer feed forward models in groundwater level forecasting using meteorological data in public management. Water Resources Management, 32(15), 5041-5052.
29. Krishna, B., Satyaji Rao, Y. R., & Vijaya, T. (2008). Modelling groundwater levels in an urban coastal aquifer using artificial neural networks. Hydrological Processes: An International Journal, 22(8), 1180-1188.
30. Kumar, M., & Kar, I. N. (2009). Nonlinear HVAC computations using least square support vector machines. Energy Conversion and Management, 50(6), 1411-1418.
31. Lemke, F. (1997). Knowledge extraction from data using self-organizing modeling technologies. In Proceedings of the SEAM'97 Conference.
32. Lin, L., Li, S., Sun, S., Yuan, Y., & Yang, M. (2020). A novel efcient model for gas compressibility factor based on GMDH network. Flow Measurement and Instrumentation, 71, 101677.
33. Malekzadeh, M., Kardar, S., & Shabanlou, S. (2019). Simulation of groundwater level using MODFLOW, extreme learning machine and Wavelet-Extreme Learning Machine models. Groundwater for Sustainable Development, 9, 100279.
34. Mathworks, M. (2014). Fuzzy logic toolbox. User's Guide, The Mathworks, Massachusetts.
35. McGarry, K. J., Wermter, S., & MacIntyre, J. (1999). Knowledge extraction from radial basis function networks and multilayer perceptrons. In IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339) (Vol. 4, pp. 2494-2497). IEEE.
36. Mehri, Y., Soltani, J., & Khashehchi, M. (2019). Predicting the coefficient of discharge for piano key side weirs using GMDH and DGMDH techniques. Flow Measurement and Instrumentation, 65, 1-6.
37. Miraki, S., Zanganeh, S. H., Chapi, K., Singh, V. P., Shirzadi, A., Shahabi, H., & Pham, B. T. (2019). Mapping groundwater potential using a novel hybrid intelligence approach. Water resources management, 33(1), 281-302.
38. Mirarabi, A., Nassery, H. R., Nakhaei, M., Adamowski, J., Akbarzadeh, A. H., & Alijani, F. (2019). Evaluation of data-driven models (SVR and ANN) for groundwaterlevel prediction in confned and unconfned systems. Environmental Earth Sciences, 78(15), 489.
39. Moghaddam, H. K., 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.
40. 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.
41. 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.
42. Moosavi, V., Mahjoobi, J., & Hayatzadeh, M. (2021). Combining Group Method of Data Handling with Signal Processing Approaches to Improve Accuracy of Groundwater Level Modeling. Natural Resources Research, 30(2), 1735-1754.
43. Moriasi, D. N., Gitau, M. W., Pai, N., & Daggupati, P. (2015). Hydrologic and water quality models: Performance measures and evaluation criteria. Transactions of the ASABE, 58(6), 1763-1785.
44. Mueller, J. A., Ivachnenko, A. G., & Lemke, F. (1998). GMDH algorithms for complex systems modelling. Mathematical and Computer Modelling of Dynamical Systems, 4(4), 275-316.
45. Nadiri, A. A., Gharekhani, M., Khatibi, R., Sadeghfam, S., & Moghaddam, A. A. (2017). Groundwater vulnerability indices conditioned by supervised intelligence committee machine (SICM). Science of the Total Environment, 574, 691-706.
46. Najafzadeh, M., Barani, G. A., & Azamathulla, H. M. (2013). GMDH to predict scour depth around a pier in cohesive soils. Applied ocean research, 40, 35–41.

47. Nariman-Zadeh, N., Darvizeh, A., Darvizeh, M., & Gharababaei, H. (2002). Modelling of explosive cutting process of plates using GMDH-type neural network and singular value decomposition. Journal of Materials Processing Technology, 128(1-3), 80-87.
48. Natarajan, N., & Sudheer, C. (2020). Groundwater level forecasting using soft computing techniques. Neural Computing and Applications, 32(12), 7691-7708.
49. Nguyen, H. T., Prasad, N. R., Walker, C. L., & Walker, E. A. (2002). A first course in fuzzy and neural control. CRC press.
50. Nourani, V., & Mousavi, S. (2016). Spatiotemporal groundwater level modeling using hybrid artifcial intelligencemeshless method. Journal of Hydrology, 536, 10–25.
51. Nourani, V., & Mousavi, S. (2016). Spatiotemporal groundwater level modeling using hybrid artificial intelligence-meshless method. Journal of Hydrology, 536, 10-25.
52. Patel, M. B., Patel, J. N., & Bhilota, U. M. (2022). Comprehensive Modelling of ANN. In Research Anthology on Artificial Neural Network Applications (pp. 31-40). IGI Global.
53. Platt J.C. (1999). Fast training of support vector machines using sequential minimal optimization, advances in kernel methods. Support Vector Learning, 185-208.
54. Rahmati, M. (2017). Reliable and accurate point-based prediction of cumulative infltration using soil readily available characteristics: A comparison between GMDH, ANN, and MLR. Journal of Hydrology, 551, 81–91.
55. Rajaee, T., Ebrahimi, H., & Nourani, V. (2019). A review of the artificial intelligence methods in groundwater level modeling. Journal of hydrology, 572, 336-351.
56. Razzagh, S., Sadeghfam, S., Nadiri, A. A., Busico, G., Ntona, M. M., & Kazakis, N. (2021). Formulation of Shannon entropy model averaging for groundwater level prediction using artificial intelligence models. International Journal of Environmental Science and Technology, 1-18.
57. Rezaei, M., Mousavi, S. F., Moridi, A., Gordji, M. E., & Karami, H. (2021). A new hybrid framework based on integration of optimization algorithms and numerical method for estimating monthly groundwater level. Arabian Journal of Geosciences, 14(11), 1-15.
58. Roshni, T., Jha, M. K., Deo, R. C., & Vandana, A. (2019). Development and evaluation of hybrid artificial neural network architectures for modeling spatio-temporal groundwater fluctuations in a complex aquifer system. Water Resources Management, 33(7), 2381-2397.
59. Sahoo, S., & Jha, M. K. (2013). Groundwater-level prediction using multiple linear regression and artificial neural network techniques: a comparative assessment. Hydrogeology Journal, 21(8), 1865-1887.
60. Sanikhani, H., & Kisi, O. (2012). River flow estimation and forecasting by using two different adaptive neuro-fuzzy approaches. Water resources management, 26(6), 1715-1729.
61. Shiri, J., & Kisi, O. (2011). Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations. Computers & Geosciences, 37(10), 1692-1701.
62. Shiri, J., Kisi, O., Yoon, H., Kazemi, M. H., Shiri, N., Poorrajabali, M., & Karimi, S. (2020). Prediction of groundwater level variations in coastal aquifers with tide and rainfall effects using heuristic data driven models. ISH Journal of Hydraulic Engineering, 1-11.
63. Shiri, J., Kisi, O., Yoon, H., Lee, K. K., & Nazemi, A. H. (2013). Predicting groundwater level fluctuations with meteorological effect implications—A comparative study among soft computing techniques. Computers & Geosciences, 56, 32-44.
64. Suryanarayana, C., Sudheer, C., Mahammood, V., & Panigrahi, B. K. (2014). An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India. Neurocomputing, 145, 324-335.
65. Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural processing letters, 9(3), 293-300.
66. Tayebi, H. A., Ghanei, M., Aghajani, K., & Zohrevandi, M. (2019). Modeling of reactive orange 16 dye removal from aqueous media by mesoporous silica/crosslinked polymer hybrid using RBF, MLP and GMDH neural network models. Journal of Molecular Structure, 1178, 514–523.
67. Thomas, B. F., & Gibbons, A. C. (2018). Sustainable water resources management: Groundwater depletion. In The Palgrave Handbook of Sustainability (pp. 53-77). Palgrave Macmillan, Cham.
68. Wen, X., Feng, Q., Deo, R. C., Wu, M., & Si, J. (2017). Wavelet analysis–artificial neural network conjunction models for multi-scale monthly groundwater level predicting in an arid inland river basin, northwestern China. Hydrology Research, 48(6), 1710-1729.
69. Wunsch, A., Liesch, T., & Broda, S. (2020). Groundwater Level Forecasting with Artificial Neural Networks: A Comparison of LSTM, CNN and NARX. Hydrology and Earth System Sciences Discussions, 2020, 1-23.
70. Yager, R. R., & Filev, D. P. (1994). Generation of fuzzy rules by mountain clustering. Journal of Intelligent & Fuzzy Systems, 2(3), 209-219.
71. Yin, J., Medellín-Azuara, J., Escriva-Bou, A., & Liu, Z. (2021). Bayesian machine learning ensemble approach to quantify model uncertainty in predicting groundwater storage change. Science of The Total Environment, 769, 144715.
72. Zadeh, L. A. (1995). Fuzzy sets. In Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh (pp. 394-432).
73. Zare, M., & Koch, M. (2018). Groundwater level fluctuations simulation and prediction by ANFIS-and hybrid Wavelet-ANFIS/Fuzzy C-Means (FCM) clustering models: Application to the Miandarband plain. Journal of Hydro-environment Research, 18, 63-76

Figures

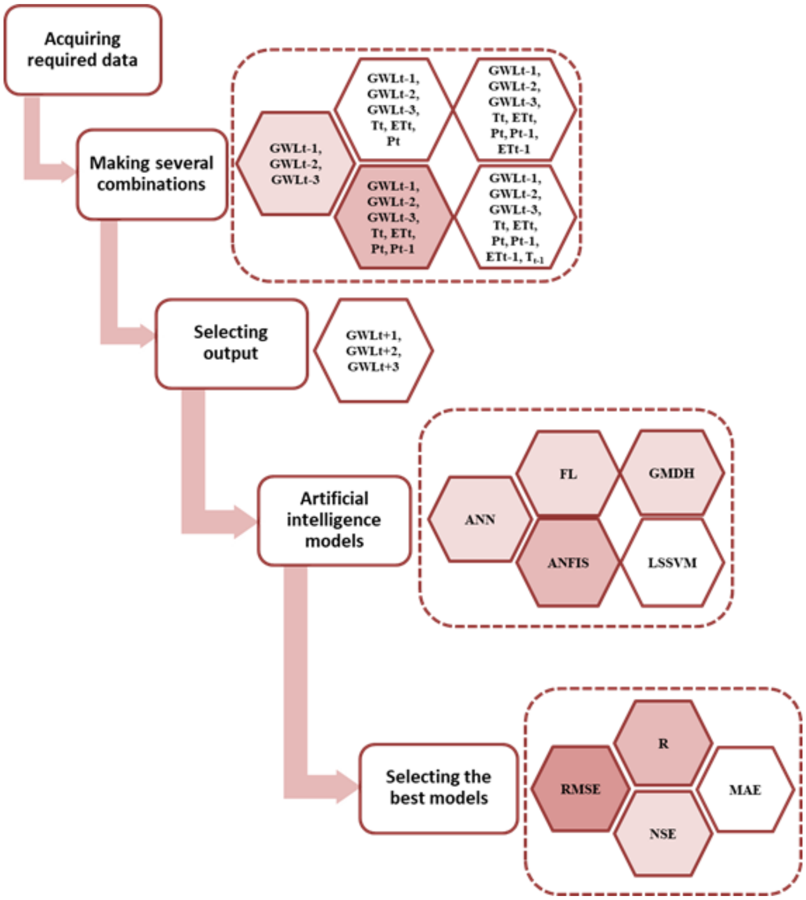

Figure 1

Methodological framework of the proposed groundwater models.

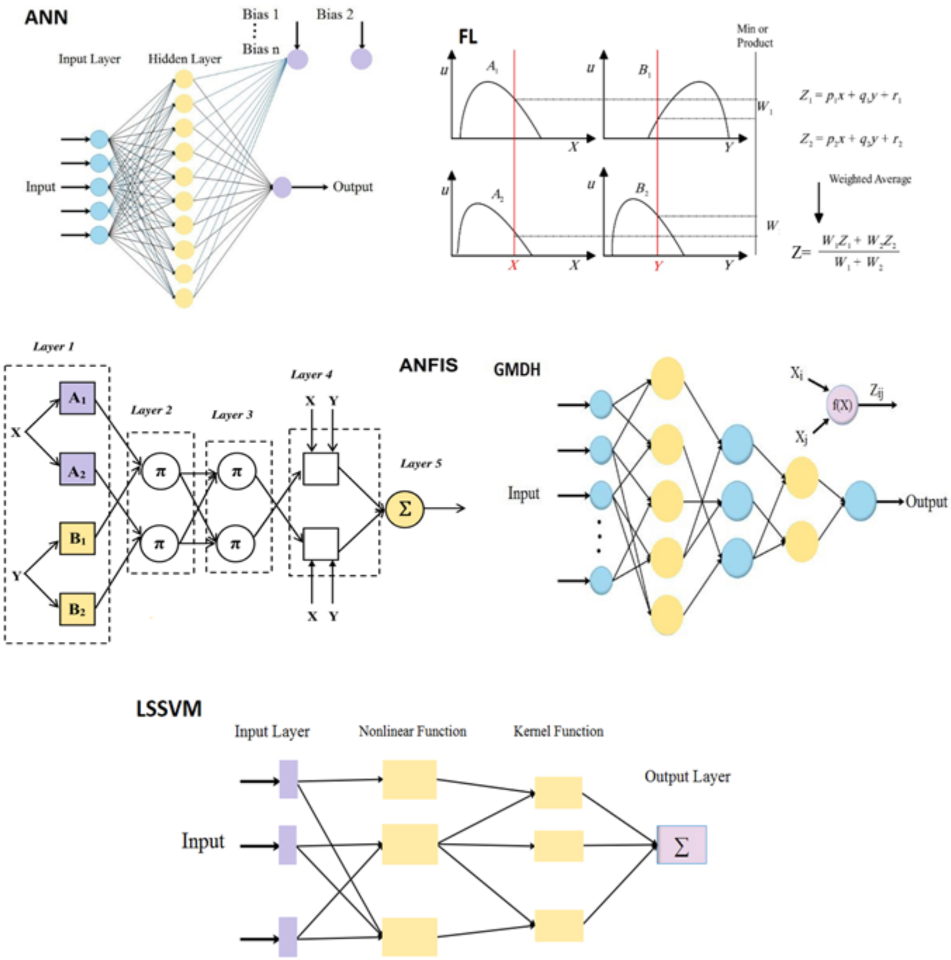

Figure 2

The general structure of AI-based models.

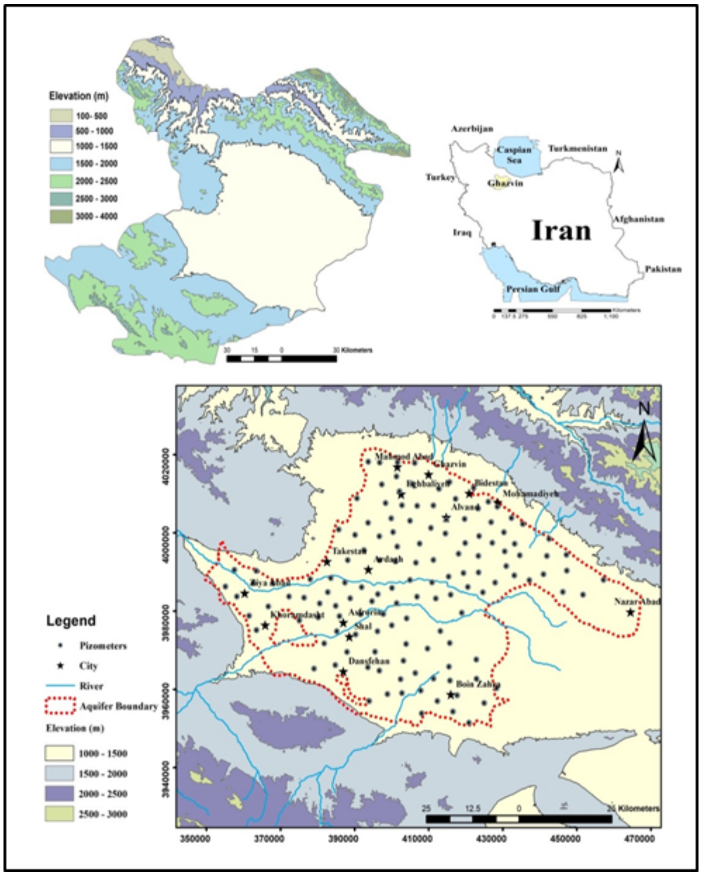

Figure 3

Location map of the study area.

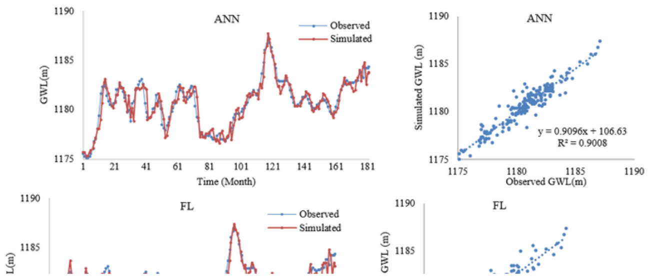

Figure 4

The observed and simulated GWL utilizing the ANN (top panel), FL (top middle panel), ANFIS (middle panel), GMDH (bottom middle panel) model and LSSVM (bottom panel) models in the GWL_{t+1} for the combination 5.

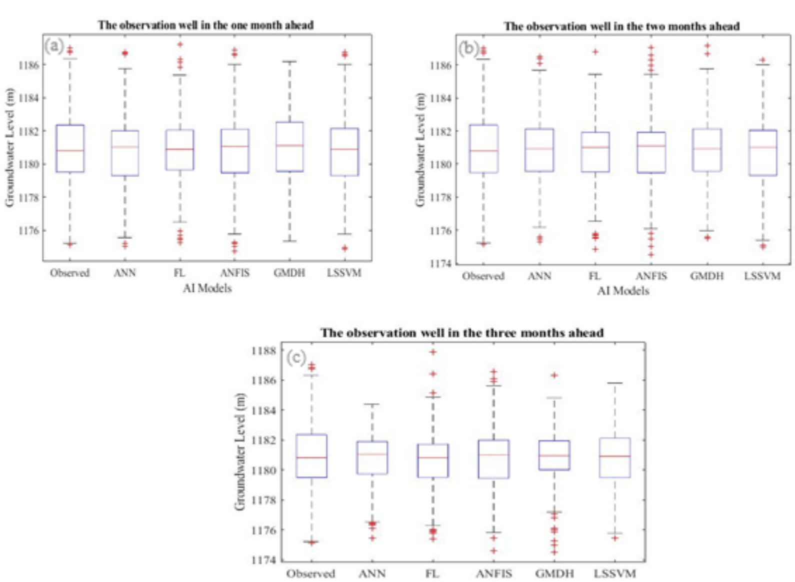

Figure 5

The observed and predicted GWL results using the five models for the combination 2 in the GWL_{t+1} (a), GWL_{t+2} (b), GWL_{t+3} (c) lead months in the observation well.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Highlights.docx