



De-noising groundwater level modeling using data decomposition techniques in combination with artificial intelligence (case study Aspas aquifer)

Maryam Shahbazi¹ · Heidar Zarei¹ · Abazar Solgi¹

Received: 16 April 2022 / Accepted: 31 January 2023 / Published online: 6 March 2023
© The Author(s) 2023

Abstract

Considering the recent significant drop in the groundwater level (GWL) in most of world regions, the importance of an accurate method to estimate GWL (in order to obtain a better insight into groundwater conditions) has been emphasized by researchers. In this study, artificial neural network (ANN) and support vector regression (SVR) models were initially employed to model the GWL of the Aspas aquifer. Secondly, in order to improve the accuracy of the models, two preprocessing tools, wavelet transform (WT) and complementary ensemble empirical mode decomposition (CEEMD), were combined with former methods which generated four hybrid models including W-ANN, W-SVR, CEEMD-ANN, and CEEMD-SVR. After these methods were implemented, models outcomes were obtained and analyzed. Finally, the results of each model were compared with the unit hydrograph of Aspas aquifer groundwater based on different statistical indexes to assess which modeling technique provides more accurate GWL estimation. The evaluation of the models results indicated that the ANN model outperformed the SVR model. Moreover, it was found that combining these two models with the preprocessing tools WT and CEEMD improved their performances. Coefficient of determination (R^2) which indicates model accuracy was increased from 0.927 in the ANN model to 0.938 and 0.998 in the W-ANN and CEEMD-ANN models, respectively. It was also improved from 0.919 in the SVR model to 0.949 and 0.948 in the W-SVR and CEEMD-SVR models, respectively. According to these results, the hybrid CEEMD-ANN model is found to be the most accurate method to predict the GWL in aquifers, especially the Aspas aquifer.

Keywords Wavelet transform · CEEMD · Groundwater level modeling · Aspas aquifer

Abbreviations

GWL	Groundwater level	W-ANN	Wavelet artificial neural network
SVR	Support vector regression	ARIMA	Autoregressive integrated moving average
ANN	Artificial neural network	W-ANFIS	Wavelet transform-adaptive neuro-fuzzy inference system
WT	Wavelet Transform	W-SVR	Wavelet transform-support vector regression
EMD	Empirical mode decomposition	ANFIS	Adaptive neuro-fuzzy inference system
CEEMD	Complementary ensemble empirical mode decomposition	SVM	Support vector machine
		GEP	Gene expression programming
		W-GEP	Wavelet transform-gene expression programming
		W-M5	Wavelet transform-M5
		SE	Sample entropy
		HS	Harmony search
		KELM	Extreme learning machine with the kernel
		ELM	Extreme learning machine
		PV	Photovoltaic
		RF	Random forest
		IGIVA	Improved gray ideal value approximation

✉ Heidar Zarei
zareih@scu.ac.ir
Maryam Shahbazi
shahbazi_jm@yahoo.com
Abazar Solgi
abazar_solgi@yahoo.com

¹ Department of Hydrology and Water Resources, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran

DIFPSO	The particle swarm optimization algorithm based on dynamic inertia factor
BPNN	Backpropagation neural network
Trainlm	Levenberg–Marquardt
Trainbfg	BFGS quasi-newton
Trainrp	Resilient backpropagation
Trainscg	Scaled conjugate gradient
Trainsgb	Conjugate gradient with Powell/Beale restarts
Traincgf	Fletcher–Powell conjugate gradient
Traincgp	Polak–Ribière conjugate gradient
Trainoss	One step secant
Trainidx	Variable learning rate gradient descent
Trainbr	Bayesian regularization
Traingdm	Gradient descent with momentum
Traingd	Gradient descent
IMF	Intrinsic mode function
EEMD	Ensemble empirical mode decomposition
R^2	Coefficient of determination
RMSE	Root mean squared error
AIC	Akaike information criterion
Poly	Polynomial
Lin	Linear

Introduction

The population growth, industrial and agricultural development, and occurrence of droughts in recent years have exerted pressure on groundwater in Iran, such that most aquifers have experienced medium to high groundwater level drops. Consequently, many aqueducts have been dried out, and most of permanent springs have experienced remarkable reductions in their water yield. More than 50% of the need for drinking water in Iran is supplied through groundwater. Accordingly, investigating the current condition of groundwater sources can assist managers in terms of more efficient planning and decision-making. The literature review of the field reveals that the most of the studies performed on groundwater have used intelligence models to estimate groundwater level (GWL). Among these intelligence models, the support vector regression (SVR) and artificial neural network (ANN) models have shown to have a good performance. In the following, some of the studies on groundwater using these models are presented.

Sattari et al. (2017) used the SVR and M5 tree models to predict the GWL in Ardabil city plain. Their results showed that both models had satisfactory performances, but the M5 model was easier to implement and interpret. In a review paper, Rajaei et al. (2019) assessed the artificial intelligence models used in groundwater modeling. Mirarabi et al. (2019) investigated SVR and ANN models for the GWL prediction. The outcomes indicated the better performance of the SVR model than the ANN model.

Preprocessing tools such as wavelet transform (WT) and empirical mode decomposition (EMD) have received attention in different fields over recent years, such that the combination of these tools with different models has resulted in hybrid models with higher accuracy. The EMD is a thoroughly effective method for extracting signals from data and is used to decompose signals in the time–frequency domain (Huang et al. 2009). The complementary ensemble empirical mode decomposition (CEEMD) method is the completed form of EMD. In this method, a pair of positive and negative white noise is added to the main data to create two series of IMF. Therefore, a combination of the main data and the additional noise is obtained in which the sum of IMFs equals the main signal. Sang et al. (2012) employed the EMD method to analyze nonlinear data in hydrology.

In the following, some of the studies on the use of the two preprocessing tools are discussed.

Adamowski and Chan (2011) applied the wavelet artificial neural network (W-ANN) to predict the GWL in the Quebec Province of Canada. The ANN and autoregressive integrated moving average (ARIMA) models were also used. The results indicated the high capability of the W-ANN compared to the other two models. Moosavi et al. (2014) optimized hybrid models of wavelet transform-adaptive neuro-fuzzy inference system (W-ANFIS) and W-ANN using the Taguchi method to predict the GWL in Mashhad, Iran. In this study, different structures were evaluated for both integrated models. According to the results, the W-ANFIS model had better performance than the W-ANN model. Suryanarayana et al. (2014) used the ANN, SVR, wavelet transform-support vector regression (W-SVR) and ARIMA to forecast monthly fluctuations of the GWL in Visakhapatnam, India. This research used monthly data of precipitation, average temperature, maximum temperature, and groundwater depth for a 13-year period (2001–2012). The results showed that the W-SVR had better performance than other models.

Eskandari et al. (2018) simulated the GWL fluctuations of the Borazjan plain using an integration of the support vector machine (SVM) and WT. The outcomes indicated the better performance of the hybrid model than the SVM model. Eskandari et al. (2019) assessed the combination of the adaptive neuro-fuzzy inference system (ANFIS) and WT in modeling and predicting the GWL of the Dalaki basin in Bushehr Province in Iran. Their results indicated that the use of the WT improved the performance of the ANFIS model by 14%. Salehi et al. (2019) predicted the GWL of Firuzabad plain using the combined model of time series-WT. They reported that the combined model outperformed the time series model.

Bahmani et al. (2020) simulated GWL using GEP and M5 tree models and their combination with WT. According to the results, the combined models had better performance

than GEP and M5 models. Moreover, the performance of the two combined models (W-GEP and W-M5) was similar. It was also observed that choosing the proper level of decomposition significantly affected the accuracy of the combined models.

Bahmani and Ouarda (2021) modeled GWL by combining artificial intelligence techniques. They used four observation wells in Delfan plain in Iran. They combined gene expression programming (GEP) and M5 models with WT and CEEMD techniques. The results indicated the better performance of the model combined with GEP.

Awajan et al. (2019) presented a review on empirical mode decomposition in forecasting time series from 1998 to 2017. Wang et al. (2019) provided a wind power short-term forecasting hybrid model based on the CEEMD-SE method in a Chinese wind farm. This method was obtained after integration processing CEEMD, sample entropy (SE), and harmony search (HS) with an extreme learning machine with the kernel (KELM). The results showed the hybrid method (CEEMD-SE-HS-KELM) had higher forecasting accuracy than EMD-SE-HS-KELM, HS-KELM, KELM, and the extreme learning machine (ELM) model.

Niu et al. (2020) provided short-term photovoltaic (PV) power generation forecasting based on random forest feature selection and CEEMD. A hybrid forecasting model was created combining random forest (RF), improved gray ideal value approximation (IGIVA), CEEMD, the particle swarm optimization algorithm based on dynamic inertia factor (DIFPSO), and backpropagation neural network (BPNN), called RF-CEEMD-DIFPSO-BPNN. Results revealed that RF-CEEMD-DIFPSO-BPNN was a promising approach in terms of PV power generation forecasting.

The GWL has considerably dropped in the Tashk–Bakhtegan basin and Maharlu lakes over recent years (Ashraf et al. 2021). The Aspas aquifer is one of the aquifers in the Basin that a high drop in GWL has experienced. The aquifer is located upstream of the basin. Due to the drop in the groundwater level, the effluent rivers have turned into influents rivers. Rivers in Aspas are of the head branches of the Kor River, and their drying or low water causes drying up of the Kor River and Bakhtegan Wetland.

Based on the review of the sources, two artificial intelligence models (ANN and SVR) have performed well. Thus, these models were used to investigate the GWL in the Aspas plain. To increase this models accuracy were used preprocessing tools. Two preprocessing tools WT and CEEMD have had a good performance in different studies, so in this study have been used. After the emergence of four combined models W-ANN, W-SVR, CEEMD-ANN, and CEEMD-SVR, the results were compared with the ANN and SVR models, and the best intelligent model was determined.

The main differences between our study and the latest studies (Bahmani et al. 2020; Bahmani and Quarda 2021) are

as follows: (i) in exception to this study, they used minimal local data, usually 3 to 8 observation wells. In this study, we used the aquifer unit hydrograph (extracted from the data of 40 observation wells) and the isohyetal, isothermal and isoevaporation maps that provide a global estimation of the aquifer accurately. (ii) The proposed hybrid approaches increased the accuracy of the intelligent models.

Materials and methods

Study area

The Tashk–Bakhtegan and Maharlu lakes basin, with an area of 31,451.8 km², is located in Fars Province in Iran. The basin is divided into two basins, namely Tashk–Bakhtegan and Maharlu, and 27 sub-basins. The Aspas study area with the code 4321 is located in the northwest of the basin. The region has a total area of 1590.5 km², including the plain (764.9 km²), highlands (816.5 km²), and lake (9.1 km²). The maximum height of the region is 3495 m (the peak of Bar Aftab mountain in the east), while the minimum height is 2061 m (Oujan river in the central plains). The most important city of the region is Sadeh. Figure 1 depicts the location of the study area and the Aspas aquifer in the Tashk–Bakhtegan and Maharlu lakes basin in Fars Province of Iran.

Data preparation

According to the data of the Aspas aquifer, 40 observation wells with a better reference period were selected. After introducing the GWL data of the observation wells in the HEC-4 software, the reference period of 2002–2020 was chosen to plot the unit hydrograph. The box plot method was used to assess the outlier data. In case of encountering any outlier data, they were removed and regenerated using the HEC-4 software. Ultimately, the unit hydrograph of the groundwater (water level variations) was plotted in a monthly scale for the Aspas aquifer, as shown in Fig. 2. Table 1 provides the location of the used meteorology stations inside and outside the Aspas aquifer. After evaluating and extending the data using the HEC-4 software and examining the outlier data, the monthly isohyetal maps were plotted for the Aspas study area, using 13 meteorology stations inside and outside it. Afterward, the monthly precipitation values were extracted for the aquifer. According to the maps, the average precipitation on the aquifer in the 19-year reference period (2002–2020) was 427 mm.

The data of eight meteorology stations inside and outside the Aspas sub-basin were employed to evaluate the temperature of the aquifer. After investigating the data and extending them using the method of differences, the outlier

Fig. 1 Location of the Aspas aquifer in Iran

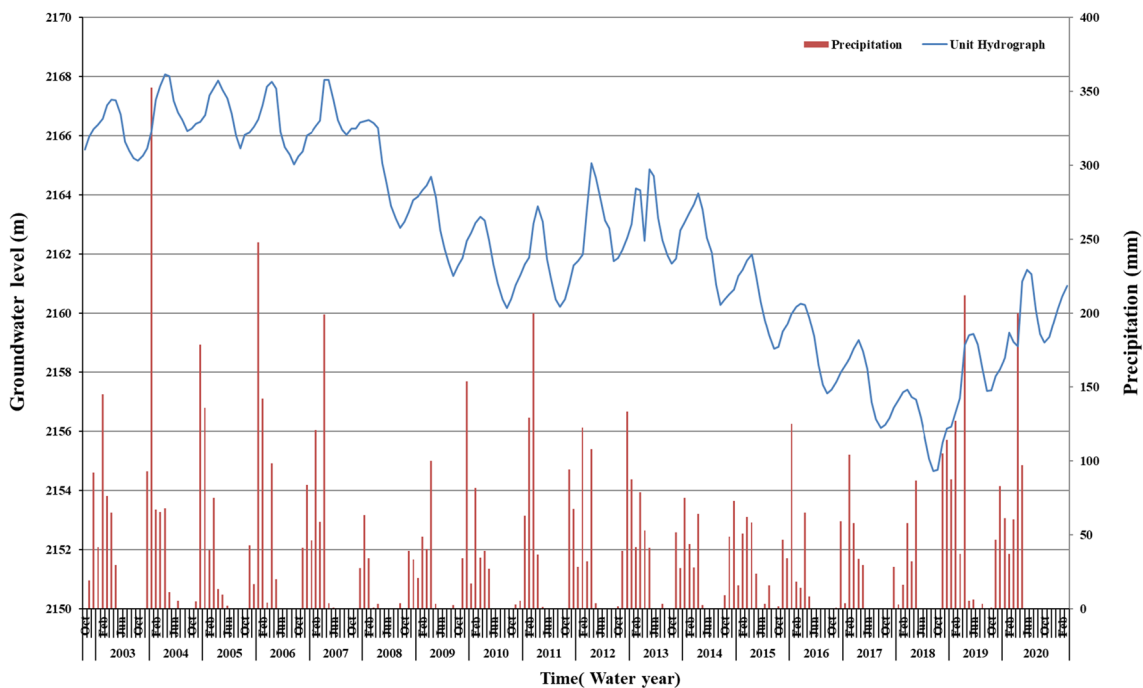
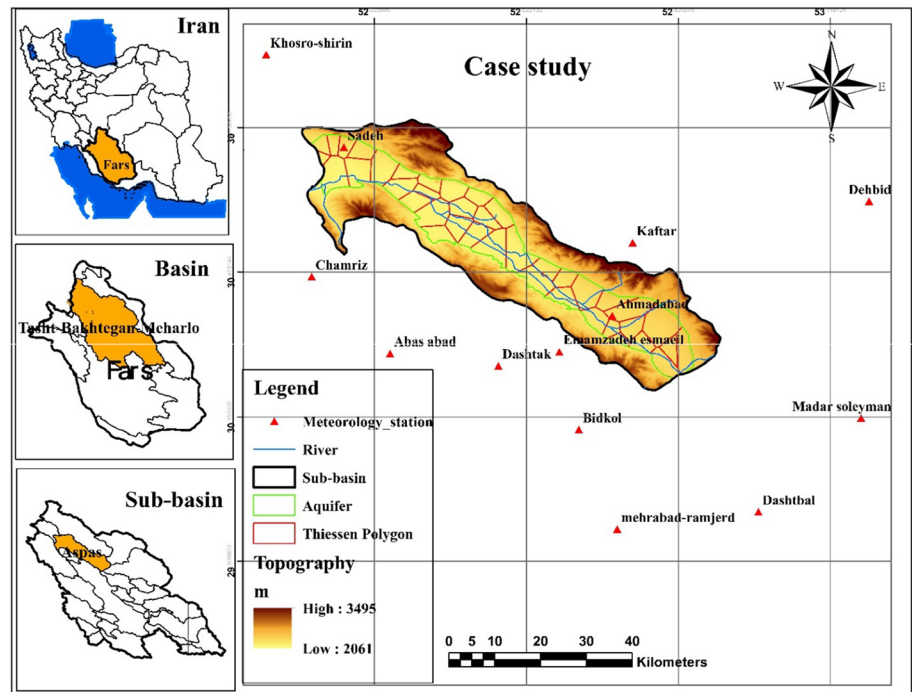


Fig. 2 Unit Hydrograph of groundwater in the Aspas aquifer

data were examined. The monthly isothermal maps were plotted for the study area. Then, the monthly temperature values were extracted for the Aspas aquifer. The average annual temperature of the Aspas aquifer was obtained 13 °C. The data of eight meteorology stations inside and outside the Aspas study area were employed to evaluate

the evaporation of the aquifer. After assessing the data and extending them based on the temperature-evaporation relationship for each station and examining their outlier data, the isoevaporation maps were plotted for the study area. Afterward, the monthly evaporation values were extracted

Table 1 Specifications of the stations used

Stations	Variable	Station code	UTMx	UTMy	Elevation (m)
Ahmad abad	Precipitation (mm)	43-068	662436	3363117	2233
Emamzadeh esmaeil		43-124	652573	3355239	1833
Bidkol	Temperature (°C), Evaporation (mm)	43-025	656499	3338446	1613
Chobkhaleh		43-061	585762	3379456	2053
Khosro shirin		43-055	596699	3419146	2340
Dezgard reisun		43-054	592058	3399077	2107
Dashtak		43-134	641114	3352077	2046
Dehbid		43-095	710365	3388840	2312
Dashtbal		43-035	690731	3321090	1673
Kemehr		43-202	584103	3368805	2377
Mehrabad ramjerd		43-008	664068	3316903	1616
Chamriz		Precipitation (mm), Temperature (°C), Evaporation (mm)	43-015	605703	3370902
Sadeh	43-201		611470	3399141	2192
Abas abad	43-082		620649	3354465	1690
Kaftar	43-029		666049	3379054	2350
Madar soleyman		43-105	709769	3341856	1861

for the aquifer. The average evaporation of the Aspas aquifer was obtained at 2273 mm.

The modeling was performed after determining the monthly data of the GWL, precipitation, temperature, and evaporation. Before using the data, they were standardized (changed to values between zero and one). For this purpose, Eq. 1 was used as recommended by Solgi et al. (2017). The input parameters of the model included precipitation, temperature, evaporation, and the GWL in a given month, and the output parameter was the GWL in the next month. Of the whole data, 75% were used for training, and 25% were used for testing. The input parameters of the model included: precipitation (P_t), temperature (T_t), evaporation (E_t), and groundwater level (GWL_t) at time t and groundwater level at time $t + 1$ (GWL_{t+1}).

$$y = 0.5 + \left(0.5 \times \left(\frac{x - \bar{x}}{x_{\max} - x_{\min}} \right) \right) \tag{1}$$

In Eq. (1), x is the desired data, \bar{x} is the average data, x_{\max} is the maximum data, x_{\min} is the minimum data, and y is the standardized data.

Figure 3 shows the general flowchart of the work steps. As can be seen in this fig., first are received the data of observation wells and meteorological stations. Then, the data are completed and extended by Hec-4 software. The boxplot method is used to check outlier data. The unit hydrograph of groundwater is drawn. Maps of isohyetal, isothermperature and isoevaporation are drawn in the GIS software. Then, monthly values for the aquifer are extracted from these maps. These values are entered as input data to tools based on data analysis. The main signal is decomposed,

and the sub-signals are extracted. The sub-signals of each input parameter are entered as input to artificial intelligence models. After modeling by artificial intelligence models, the results of each model are compared by observed values. If the results are acceptable, the model results are extracted, otherwise, the models are run again to improve the results by changing the effective factors in the modeling. In the end, the results of all models are compared based on evaluation criteria and the best model is selected.

Artificial intelligence models

According to the evaluations of the intelligence models, two of them that had good performance in previous studies, i.e., ANN and SVR, were used. Afterward, two preprocessing tools, namely the WT and CEEMD, were employed to enhance the performance of the models.

ANN model

The following are effective in modeling artificial neural networks (ANNs) (Solgi 2014):

1. Sufficient training,
2. Number of layers in a network,
3. Number of neurons of the middle layers,
4. Training laws,
5. Simulation functions (transmission)

An important criterion in training a network is the number of iterations (epochs) or repetitions the network experiences while training. In training a network, determining the

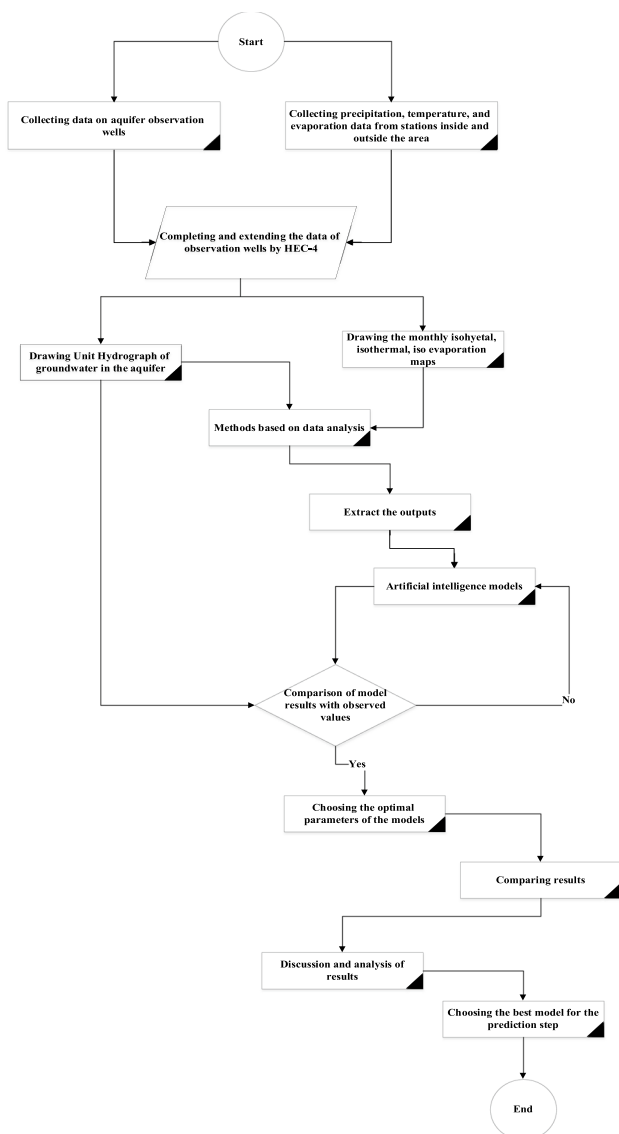


Fig. 3 Flowchart of work steps in this study

proper number of iterations is of great importance. In general, the more the number of iterations in training a network, the smaller the simulation (prediction) error. However, when the number of iterations exceeds a given value, the testing group error increases. The best number of iterations minimizes the errors in both training and testing groups.

The number of layers in a network is among the main criteria in designing ANNs. These networks are usually comprised of several layers. These layers include an input layer, middle layers, and an output layer. The number of the middle layers is determined by trial and error. In general, it is recommended to use ANNs with lower middle layers.

In ANNs, the number of neurons of the input and output layers is a function of the type of problem. However, there is no particular relationship for the number of neurons in

the middle layers, and these neurons are determined based on trial and error for each middle layer.

The internal mode of a neuron caused by the activation function is known as the activity level or action. Generally, each neuron sends its activity level to one or more of the other neurons in the form of a single signal. The activation functions of a layer's neurons are typically, but not essentially, the same. Moreover, these activation functions are nonlinear for the neurons of the hidden layer while being identity functions for the input layer's neurons. Using a reaction function, a neuron produces outputs for various inputs. Linear and sigmoid functions are two well-known functions (Karamuz and Araghinezhad 2010). An activation function is chosen based on the specific need of a problem that should be solved by the ANN. In practice, a limited number of activation functions is used, as listed in Fig. 4 (Alborzi 2001).

The network architecture in solving a problem refers to choosing the proper number of layers, the number of neurons in each layer, the way connections are made between neurons, and the proper activation functions for neurons, along with determining the training algorithm and the way the weights are adjusted.

Another major parameter of an ANN is the training function. Table 2 lists the training functions used in the current study.

SVR model

In this case, the kernel functions are used. In this study, four common kernels as the following were used:

$$k(x, y) = (x \cdot y + 1)^d \quad d = 2, 3, \dots$$

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right)$$

$$k(x, y) = \tanh(x \cdot y + \theta)$$

$$k(x, y) = (x \cdot y) \quad (2)$$

These kernels are the polynomial (Poly) kernel, RBF kernel, Sigmoid kernel, and linear (Lin) kernel, respectively (Raghavendra and Deka 2014). The RBF kernel has one parameter g (Gamma). In the sigmoidal kernel, just the default values, including zero and $1/k$ Gamma, are used. The linear kernel does not have any parameters. The polynomial kernel has two parameters d (the degree of the polynomial) and r (an accumulative fixed number). For more information on this topic, see references (Cortes and Vapni 1995; Raghavendra and Deka 2014). As listed in Table 3,

Fig. 4 A number of activation functions used in the ANN (Alborzi 2001)

Name	Input/Output Relation	Icon	MATLAB Function
Hard Limit	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$		hardlim
Symmetrical Hard Limit	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$		hardlims
Linear	$a = n$		purelin
Saturating Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$		satlin
Symmetric Saturating Linear	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = 1 \quad n > 1$		satlins
Log-Sigmoid	$a = \frac{1}{1 + e^{-n}}$		logsig
Hyperbolic Tangent Sigmoid	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
Positive Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$		poslin
Competitive	$a = 1 \quad \text{neuron with max } n$ $a = 0 \quad \text{all other neurons}$		compet

Table 2 Training functions used in this study

No	Training function	Function name in MATLAB software
1	Levenberg–Marquardt	Trainlm
2	BFGS Quasi-Newton	Trainbfg
3	Resilient backpropagation	Trainrp
4	Scaled conjugate gradient	Trainscg
5	Conjugate gradient with powell/Beale restarts	Traincgb
6	Fletcher–Powell conjugate gradient	Traincgf
7	Polak–Ribière conjugate gradient	Traincgp
8	One Step Secant	Trainoss
9	Variable learning rate gradient descent	Traingd
10	Bayesian regularization	Trainbr
11	Gradient descent with momentum	Traingdm
12	Gradient descent	Traingd

six combinations of architecture were considered to run the ANN and SVR models.

Table 3 Details of different architectures of ANN and SVR models

Architecture	Input	Output
1	GWL_{t-1}, GWL_t	GWL_{t+1}
2	P_{t-1}, P_t	GWL_{t+1}
3	T_{t-1}, T_t	GWL_{t+1}
4	E_{t-1}, E_t	GWL_{t+1}
5	GWL_t, P_t, E_t, T_t	GWL_{t+1}
6	$GWL_{t-1}, GWL_t, P_{t-1}, P_t, E_{t-1}, E_t, T_{t-1}, T_t$	GWL_{t+1}

Preprocessing tools

In this study, used two tools includes the WT and CEEMD.

Wavelet transform (WT)

Wavelet transform (WT) is one of the efficient mathematical transformations in the field of signal processing. Wavelets are mathematical functions that provide a time-scale form of time series and their relationships for analyzing time series that include variables and non-constants. The wavelet function is one that has two important properties of oscillation and being short-term. The wavelet coefficients can be

calculated at any point of the signal (b) and for any value of the scale (a) in the following equation (Nourani et al. 2009).

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \psi\left(\frac{t-b}{a}\right) f(t) dt \quad (3)$$

In Eq. (3), ψ is the mother wavelet function, a is scales and b translates the function. For different values of a , b , the value of T is obtained.

Because the effect of time series data is differentiated before entering different models and the initial signal is decomposed into several sub-signals, it is possible to use an analysis that includes short-term and long-term effects, which in turn optimizes the model in future evaluations and estimates.

For more information on this topic, see references (Mallat 1998; Foufoula-Georgiou & Kumar 1994).

CEEMD method

The empirical mode decomposition (EMD) is a method for decomposing different signals in a process called screening. During this process, the main signal is decomposed into a number of components with different frequency contents. According to Eq. 4, the EMD method decomposes the main signal $x(n)$ to a number of intrinsic mode functions (IMFs) (Amirat et al. 2018).

$$x(n) = \sum_{i=1}^n c_i(x) + r_n(x) \quad (4)$$

In Eq. 4, $r_n(x)$ is the remaining component after n IMFs and $c_i(x)$. $C(x)$ is the wave-shape (harmonic) function extracted from the main signal, which does not have the conditions of the IMF. Data can have several IMFs at a time. These fluctuating modes are called intrinsic mode functions and have the two following conditions: 1. in all data, the number of extremum points equals the number of zero points or, at most, is different by a unit. 2. At each point, the average of the envelopes fitted to local maximum and minimum points should be zero.

Given the presence of alternation and noise in the signals, in some cases, the frequency-time domain distribution is interrupted, and the EMD performance is disturbed due to the difference in the modes. In order to solve this problem, Wu and Huang (2004) proposed a different method called ensemble empirical mode decomposition (EEMD). In the decomposition procedure of EEMD, a limited volume of white noise enters the main signal. By using the positive statistical aspects of the white noise, which is uniformly distributed in the frequency domain, the effect of the alternating noise is removed from the decomposition process. In the CEEMD method, a pair of positive and negative white noise

is added to the main data to create two series of IMF. Therefore, a combination of the main data and the additional noise is obtained in which the sum of IMFs equals the main signal.

Combining models with the WT method

This section discusses the formation of hybrid models obtained from the combination of WT with the ANN and SVR intelligence models. When an initial signal is decomposed using the WT method, and the resulting sub-signals are used as inputs to the ANN and SVR intelligence models, the hybrid models of W-ANN and W-SVR are obtained. As shown in Fig. 4, at first, the signals of the input parameters of precipitation, temperature, evaporation, and the GWL are decomposed using the WT. Then, the obtained sub-signals are introduced as inputs in the ANN and SVR models to create W-ANN and W-SVR hybrid models. Figure 5 depicts the $Pa(t)$, $Ta(t)$, $GWL_a(t)$ and $Ea(t)$ sub-signals of the overall scale (approximate) in the last level and other sub-signals of the detailed scale (detail) from level one to the last level.

Combining models with the CEEMD method

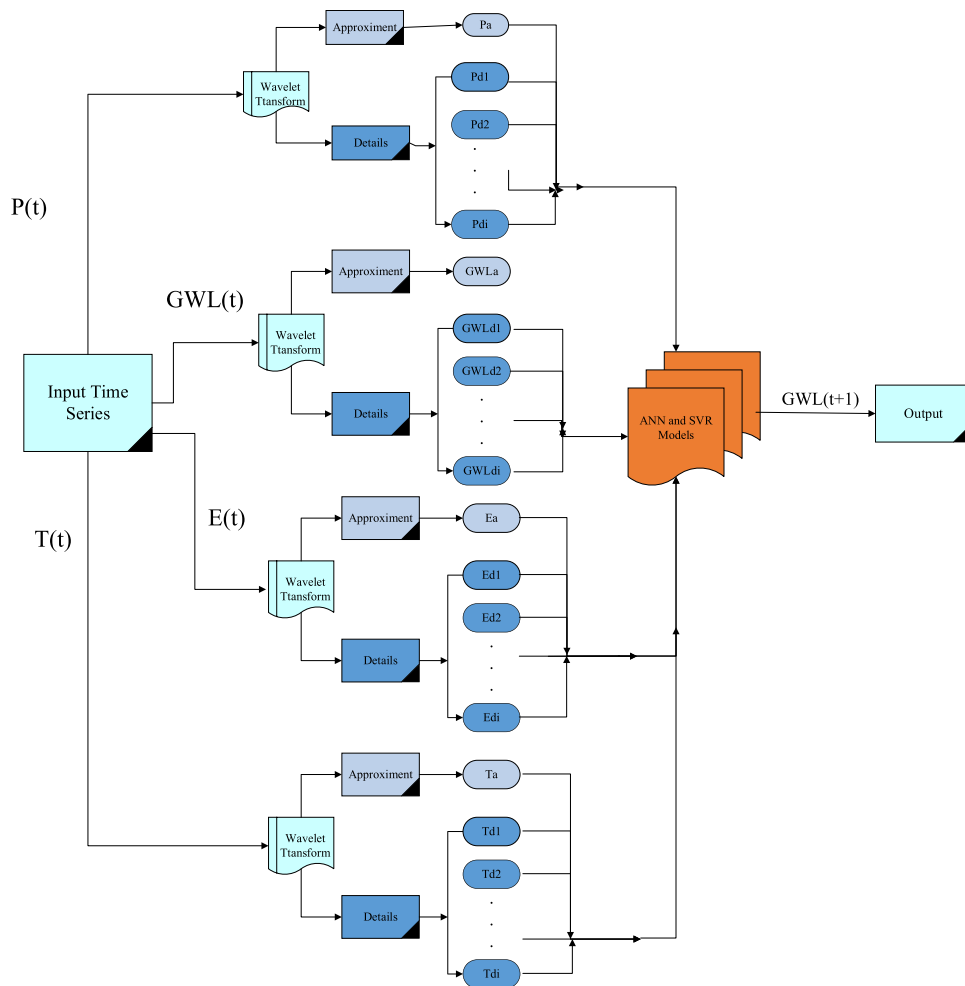
This section describes the creation of hybrid models obtained from the combination of CEEMD with the ANN and SVR intelligence models. When an initial signal is decomposed using the CEEMD method, and its sub-signals are used as inputs to the ANN and SVR intelligence models, the hybrid models of CEEMD-ANN and CEEMD-SVR are obtained. As shown in Fig. 5, at first, the signals of the input parameters of precipitation, temperature, evaporation, and the GWL are decomposed using the CEEMD. Then, the obtained sub-signals are introduced as inputs in the ANN and SVR models to create CEEMD-ANN and CEEMD-SVR hybrid models. Figure 6 depicts the $Pimf(t)$, $Timf(t)$, $GWL_{imf}(t)$ and $Eimf(t)$ sub-signals of the overall scale (IMF) in the last level and other sub-signals of the residuals from level one to the last level.

The CEEMD method has two parameters, i.e., the maximum number of IMFs and ℓ . It can be said that the number of IMFs is the level of decomposition in the WT method. In this study, the values of 0.1, 0.2, and 0.3 were used for ℓ . Some studies have used 0.2 for this parameter. Moreover, according to the structure of the study, the number of IMFs was one to six. Accordingly, the hybrid models were used in different modes by combining six IMFs and three ℓ values.

Evaluation criteria

In the performance assessment of models, different qualitative and quantitative parameters should be evaluated to clearly observe the effectiveness of each input parameter on the results. Accordingly, the following parameters were used

Fig. 5 Schematic diagram of the W-ANN and W-SVR models



to examine the efficiency of the methods (Solgi et al. 2017; Bahmani et al. 2020).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (G_{i_{obs}} - G_{i_{pre}})^2}{n}} \tag{5}$$

$$R^2 = \frac{\left(\sum_{i=1}^n (G_{i_{obs}} - \overline{G_{i_{obs}}})(G_{i_{pre}} - \overline{G_{i_{pre}}}) \right)^2}{\sum_{i=1}^n (G_{i_{obs}} - \overline{G_{i_{obs}}})^2 \sum_{i=1}^n (G_{i_{pre}} - \overline{G_{i_{pre}}})^2} \tag{6}$$

$$AIC = m \times \ln(RMSE) + 2(N_{par}) \tag{7}$$

In equations above, the parameters include:

n : number of data, i counter variable, $G_{i_{obs}}$: observational data, $\overline{G_{i_{obs}}}$: average of the observational data, $G_{i_{pre}}$: computational data, $\overline{G_{i_{pre}}}$: average predicted data, m : number of the parameters of the model and N_{par} : number of the trained data.

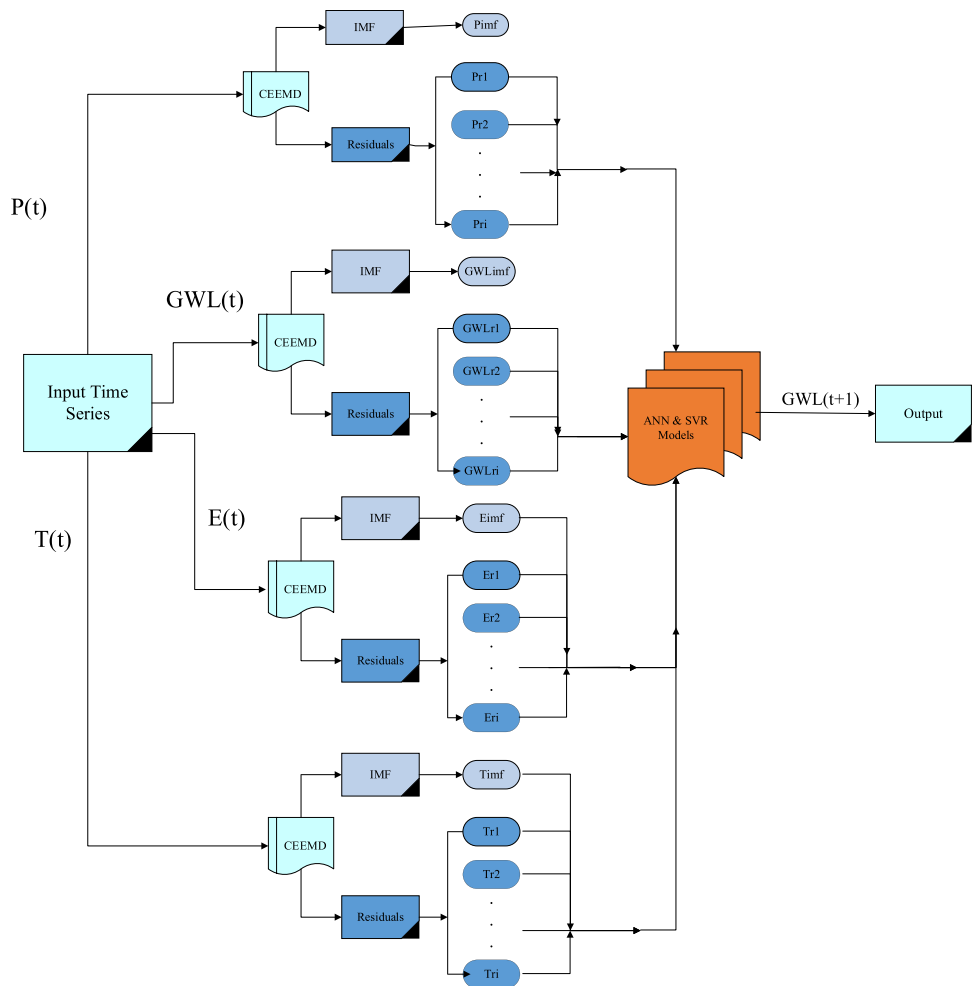
The coefficient of determination (R^2) determines the agreement between the data created by the model and the

real data. The values closer to one indicate better agreement and lower error. Therefore, this parameter was used to evaluate the effectiveness of each parameter, such as the type of the wavelet function, number of middle neurons, and decomposition degree of wavelet in the performance of all models. Moreover, the parameter of RMSE is the root mean square error of the computational and observational data. It is clear that lower values of this parameter indicate the better training and simulation of the data. Regarding the Akaike information criterion (AIC), it can be said that lower Akaike coefficients reveal the better performance of the models. The low value of the Akaike coefficient is caused by two factors, i.e., the error of the model and the number of parameters. Therefore, it is a good criterion to evaluate models.

Results and discussion

This section provides the results of the intelligence models executed for the GWL estimation using the data obtained from the plotted isohyetal, isothermal, and isoevaporation maps and the unit hydrograph of the groundwater.

Fig. 6 Schematic diagram of the CEEMD-ANN and CEEMD-SVR models



Afterward, the preprocessing methods were used, and the obtained hybrid models and their results were provided. Eventually, the results of the models were compared based on the statistical criteria.

Artificial intelligence models

Results of the ANN model

The ANN was modeled by coding in the MATLAB software and using all effective parameters previously mentioned.

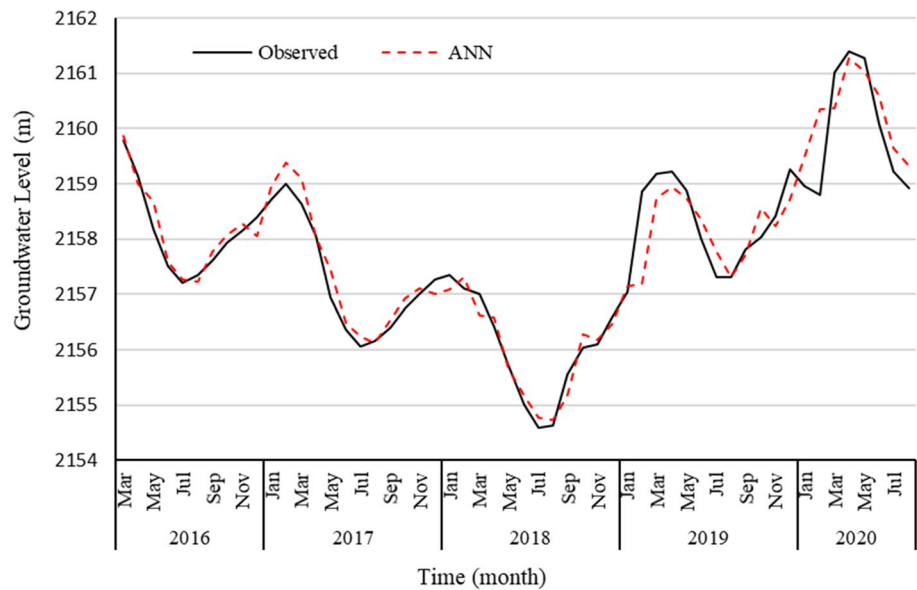
For this purpose, different structures were evaluated in each combination. Table 4 provides the results of the best structure in each combination. As can be seen in the table, combination No. 6 with eight input parameters and three neurons in the middle layer had the best performance. Its coefficient of determination was 0.927, and its error was 0.0158. In this prior combination, the training law of trainlm and stimulation (transmission) function of tansig had the best performance compared to other training laws and transmission functions. The performance of this model can be compared with the observational values in Fig. 7. As can be

Table 4 The best performance in the ANN model combinations at the monthly scale

Combination	Model parameters				R^2		RMSE	
	Structure	Epoch	Learning rules	Transfer function	Train	Test	Train	Test
1	2-6-1	1000	Trainlm	Tansig	0.955	0.888	0.0198	0.0215
2	2-5-1	111	Trainbfg	Tansig	0.184	0.165	0.0839	0.2302
3	2-4-1	1000	Trainlm	Tansig	0.192	0.144	0.0834	0.2263
4	2-4-1	1000	Trainlm	Tansig	0.329	0.224	0.0760	0.2126
5	4-3-1	225	Trainbr	Tansig	0.969	0.912	0.0163	0.0179
6	8-3-1	1000	Trainlm	Tansig	0.985	0.927	0.0113	0.0158

Bold values represent the best performance

Fig. 7 Comparison of the ANN model with observed data



seen, the ANN model had better performance in minimum points compared to the peak points.

Results of the SVR model

The SVR was modeled by coding in the MATLAB software and using all effective parameters previously mentioned. For this purpose, different structures were evaluated in each combination. Table 5 provides the results of the best structure in each combination. As can be seen in the table, combinations 5 and 6 had the same coefficient of determination, but given its lower error, combination No. 6 was known as the superior combination. The coefficient of determination and error of this combination were 0.918 and 0.0168, respectively. In this superior combination, the line kernel had the best performance among all kernels. The performance of this model can be compared with the observational values in Fig. 8. As can be seen in the figure, the estimations of the SVR model in the maximum and minimum points were different from the observational values.

Results of the W-ANN model

According to Nourani et al. (2009), Eq. 8 can be used as an initial estimation in determining the decomposition level in the WT method on a monthly time scale. In this equation, L denotes the decomposition level, and N is the number of data. Given the number of data in this study, which was 216, L was calculated at two. In order to improve the accuracy, the decomposition levels were considered from one to four. For this purpose, the WT method was coded in MATLAB software.

$$L = \text{Int}[\log(N)] \tag{8}$$

In order to execute the W-ANN model, the sub-signals obtained from the execution of the WT in different decomposition levels were introduced as inputs to the ANN model. To this end, the W-ANN model was studied in different decomposition levels with various wavelet functions in different structures, whose results are listed in Table 6. According to the table, wavelet function

Table 5 The best performance in the SVR model combinations at the monthly scale

Combination	Model parameters					R^2		RMSE	
	Kernel	Gamma	Sin2	B	Polynomials degree	Train	Test	Train	Test
1	Lin	5.157	-	-0.0003	-	0.954	0.886	0.0198	0.0216
2	Poly	0.0001	419.84	-0.0001	3	0.122	0.214	0.0870	0.2313
3	Lin	0.0555	-	0.0006	-	0.102	0.153	0.0880	0.2294
4	Poly	1.0749	0.1919	0.2020	3	0.228	0.230	0.0816	0.2139
5	Lin	4.0064	-	-0.0002	-	0.967	0.919	0.0167	0.0194
6	Lin	1.4603	-	-0.0002	-	0.972	0.918	0.0156	0.0168

Bold values represent the best performance

Fig. 8 Comparison of the SVR model with observed data

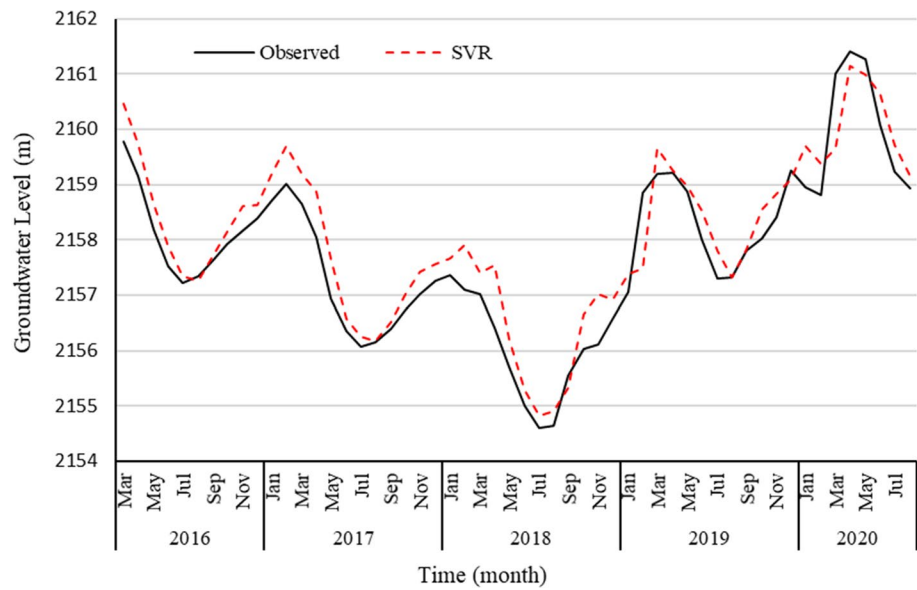


Table 6 Result of W-ANN model at the monthly scale with different mother wavelets and decomposition levels

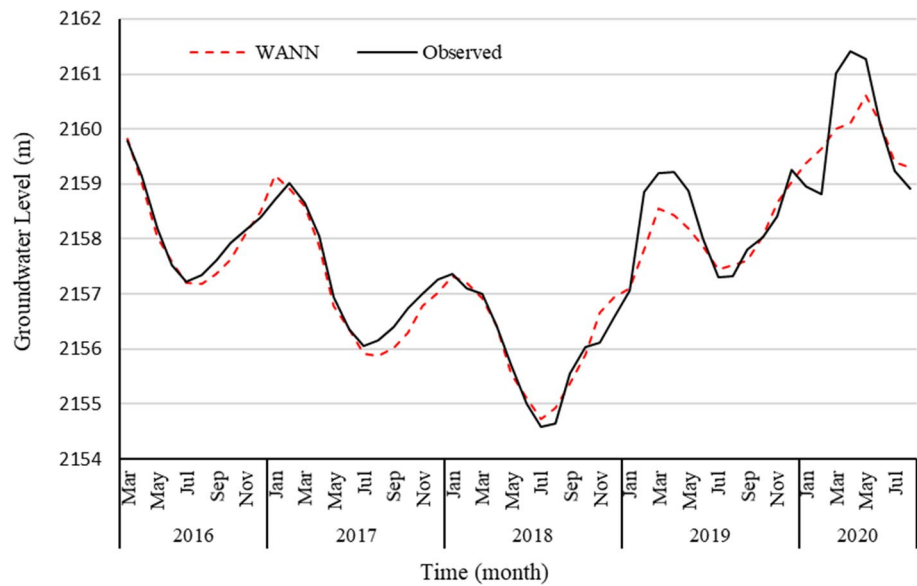
Wavelet Function	Level	Model parameters				R^2		RMSE	
		Structure	Epoch	Learning rules	Transfer function	Train	Test	Train	Test
Coif1	1	8-3-1	133	Trainbr	Tansig	0.991	0.920	0.0087	0.0301
	2	12-7-1	1000	Trainbr	Tansig	0.987	0.879	0.0107	0.0279
	3	16-4-1	429	Trainbr	Tansig	0.995	0.842	0.0068	0.0254
	4	20-4-1	1000	Trainbr	Purelin	0.987	0.885	0.0106	0.0264
Db2	1	8-3-4-1	238	Trainbr	Tansig	0.995	0.871	0.0067	0.0291
	2	12-4-1	230	Trainbr	Tansig	0.997	0.769	0.0050	0.0337
	3	16-3-1	386	Trainbr	Tansig	0.991	0.834	0.0089	0.0290
	4	20-4-1	1000	Trainbr	Purelin	0.987	0.927	0.0107	0.0166
Db4	1	8-4-1	339	Trainlm	Purelin	0.983	0.836	0.0119	0.0429
	2	12-6-1	123	Trainlm	Purelin	0.995	0.645	0.0065	0.0504
	3	16-3-3-1	1000	Trainbr	satlin	0.983	0.929	0.0120	0.0177
	4	20-5-1	455	Trainbr	Tansig	0.999	0.767	0.0029	0.0620
Sym3	1	8-4-1	1000	Trainbr	Purelin	0.984	0.938	0.0116	0.0149
	2	12-4-4-1	1000	Trainbr	Tansig	0.984	0.937	0.0116	0.0163
	3	16-4-1	159	Trainlm	Purelin	0.984	0.903	0.0116	0.0252
	4	20-4-1	436	Trainbr	Tansig	0.997	0.658	0.0052	0.0445
Haar	1	8-4-1	1000	Trainbr	Tansig	0.986	0.921	0.0108	0.0196
	2	12-3-3-1	320	Trainbr	Tansig	0.985	0.923	0.0112	0.0193
	3	16-6-1	1000	Trainbr	Purelin	0.985	0.921	0.0113	0.0190
	4	20-5-5-1	179	Trainbr	Tansig	0.985	0.924	0.0112	0.0191

Bold values represent the best performance

Sym3 in decomposition level 1 had the best performance. This superior combination had eight input neurons with four hidden layers with the training law of trainbr and the stimulation function of purelin. In this combination, RMSE = 0.0149 and $R^2 = 0.938$, which indicated its better performance than other combinations. The performance of

this model can be compared with the observational values in Fig. 9. As can be seen in the figure, the W-ANN model did not have proper performance in the peak points.

Fig. 9 Comparison of the W-ANN model with observed data



Results of the W-SVR model

In order to execute the W-SVR model, the sub-signals obtained from the execution of the WT in different decomposition levels were introduced as inputs to the ANN model. To this end, the W-SVR model was studied in

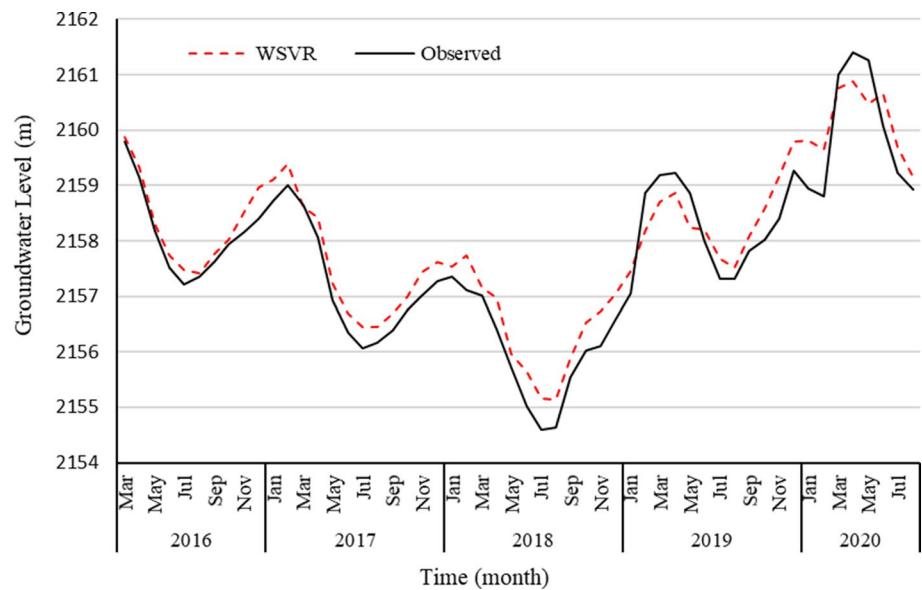
different decomposition levels with various wavelet functions in different structures, whose results are listed in Table 7. According to the table, wavelet function Coif1 in decomposition level 2 had the best performance. This superior combination had an RBF kernel. In this combination, RMSE = 0.0164 and $R^2 = 0.949$, which indicated its better performance than other combinations. The performance of this model can be compared with the

Table 7 Result of W-SVR model at the monthly scale with different mother wavelets and decomposition levels

Wavelet function	Level	Model parameters					R^2		RMSE	
		Kernel	Gamma	Sin2	B	P- Degree	Train	Test	Train	Test
Coif1	1	RBF	1277.39	1231.07	6.6989	–	0.992	0.942	0.0084	0.0149
	2	RBF	2292.08	1365.20	7.4714	–	0.992	0.949	0.0086	0.0164
	3	RBF	883.19	1182.99	0.0001	–	0.991	0.930	0.0087	0.0174
	4	RBF	804.06	118,838	0.0004	–	0.991	0.860	0.0088	0.0279
Db2	1	RBF	1520.50	2495.24	0.0003	–	0.987	0.011	0.0108	0.0692
	2	RBF	6487.95	1728.11	0.0001	–	0.986	0.872	0.0108	0.0390
	3	Lin	0.49	–	–0.0002	–	0.986	0.898	0.0109	0.0405
	4	Lin	0.51	–	0.0002	–	0.986	0.901	0.0109	0.0380
Db4	1	Lin	1.4804	–	–0.0001	–	0.983	0.901	0.0121	0.0395
	2	Lin	0.4913	–	–0.0001	–	0.983	0.835	0.0122	0.0432
	3	Lin	0.44	–	–0.0003	–	0.983	0.898	0.0122	0.0205
	4	RBF	7422.24	2246.52	0.0001	–	0.997	0.761	0.0047	0.0290
Sym3	1	Lin	2.03	–	–0.0001	–	0.984	0.937	0.0117	0.0147
	2	RBF	3132.17	1422.53	0.0006	–	0.993	0.830	0.0077	0.0039
	3	Lin	1.6352	–	–0.0001	–	0.984	0.810	0.0117	0.0477
	4	Lin	1.5527	–	–0.0001	–	0.984	0.703	0.0117	0.0694
Haar	1	RBF	1469.69	3458.91	–0.0534	–	0.991	0.938	0.0086	0.0195
	2	RBF	1521.31	2628.48	–0.0003	–	0.991	0.932	0.0089	0.0173
	3	RBF	4681.02	1126.31	–0.4914	–	0.992	0.936	0.0084	0.0172
	4	RBF	1012.11	6522.21	–0.0081	–	0.991	0.933	0.0088	0.0169

Bold values represent the best performance

Fig. 10 Comparison of the W-SVR model with observed data



observational values in Fig. 10. As can be seen in the figure, the W-SVR model did not have proper performance in the minimum and maximum points.

Results of the CEEMD-ANN model

The R software was employed to execute the CEEMD method. At first, the "hht" package was installed on the R software. Then, sub-signals were extracted using the CEEMD method through coding in the R software environment. The program was executed for the maximum

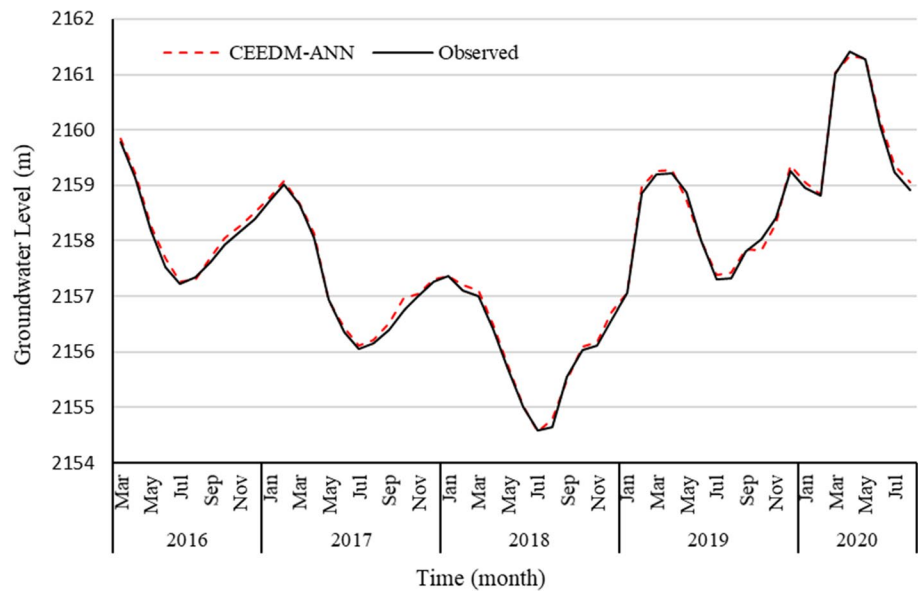
IMF number of 10. The results showed that the maximum number of sub-signals could be six. Thus, the code of the program was executed for the IMF numbers of one to six. Afterward, the obtained sub-signals were introduced as inputs to the ANN model. The results of the evaluation are given in Table 8. According to the table, the best performance in the IMF equaled one, and the value of ϵ was obtained 0.2. In this superior structure, the training law of trainbr and transmission function of tansig were the best. This structure had an R^2 of 0.98 and an RMSE of 0.0035. The performance of this model can be compared with the

Table 8 Result of CEEMD-ANN model at the monthly scale

Number of IMF	ϵ	ANN model parameters				R^2		RMSE	
		Structure	Epoch	Learning rules	Transfer function	Train	Test	Train	Test
1	0.1	8-5-1	1000	Trainbr	Purelin	0.976	0.929	0.0143	0.0180
	0.2	8-4-4-1	1000	Trainbr	Tansig	0.997	0.997	0.0052	0.0035
	0.3	8-4-1	163	Trainlm	Purelin	0.971	0.918	0.0157	0.0194
2	0.1	1-5-1	165	Trainlm	Purelin	0.980	0.947	0.0131	0.0144
	0.2	1-4-1	1000	Trainbr	Purelin	0.976	0.920	0.0143	0.0187
	0.3	1-4-1	234	Trainlm	Purelin	0.973	0.932	0.0154	0.0181
3	0.1	1-4-1	256	Trainlm	Purelin	0.984	0.941	0.0116	0.0139
	0.2	12-4-1	1000	Trainbr	Purelin	0.981	0.931	0.0129	0.0150
	0.3	12-3-3-1	144	Trainbfg	Purelin	0.980	0.927	0.0132	0.0170
4	0.1	20-4-1	306	Trainbfg	Purelin	0.984	0.936	0.0119	0.0146
	0.2	20-3-3-1	247	Trainlm	Purelin	0.982	0.919	0.0124	0.0165
	0.3	20-5-1	1000	Trainbr	Purelin	0.981	0.925	0.0127	0.0163
5	0.1	24-4-1	1000	Trainbr	Purelin	0.985	0.927	0.0113	0.0155
	0.2	24-4-1	1000	Trainbr	Purelin	0.985	0.921	0.0113	0.0186
	0.3	24-4-1	220	Trainbfg	Purelin	0.980	0.918	0.0131	0.0198
6	0.1	24-4-1	1000	Trainbr	Purelin	0.984	0.927	0.0118	0.0172
	0.2	24-6-6-1	1000	Trainbr	Purelin	0.981	0.925	0.0128	0.0183
	0.3	28-4-1	29	Trainlm	Purelin	0.981	0.923	0.0126	0.0187

Bold values represent the best performance

Fig. 11 Comparison of the CEEMD-ANN model with observed data



observational values in Fig. 11. As can be seen in the figure, the CEEMD-ANN model had a suitable performance in almost all points, and its estimates were close to the observational values.

Results of the CEEMD-SVR model

The program was executed for the IMF numbers of one to six. Then, the resulting sub-signals were introduced as inputs to the SVR model to obtain the CEEMD-SVR hybrid model.

Table 9 lists the results of using this hybrid model. According to the table, the best performance in the IMF equaled two, and the value of ϵ was obtained 0.1. Kernel Lin had the best performance. The values of R^2 and RMSE were equal to 0.95 and 0.0155, respectively. The performance of this model can be compared with the observational values in Fig. 12. As can be seen in the figure, the CEEMD-SVR hybrid model did not have proper performance in the peak points.

Table 9 Result of CEEMD-SVR model at the monthly scale

Number of IMF	ϵ	SVR Model Parameters					R^2		RMSE	
		Kernel	Gamma	Sin2	B	P- Degree	Train	Test	Train	Test
1	0.1	RBF	4594.20	1266.71	0.0007	–	0.980	0.935	0.0133	0.0182
	0.2	RBF	146.33	205.76	0.0001	–	0.977	0.930	0.0142	0.0154
	0.3	RBF	805.59	767.47	0.0005	–	0.974	0.919	0.0149	0.0215
2	0.1	Lin	2.42	–	–0.0003	–	0.980	0.948	0.0133	0.0155
	0.2	Lin	4.30	–	–0.0002	–	0.976	0.920	0.0143	0.0187
	0.3	Lin	3.66	–	–0.0001	–	0.973	0.932	0.0154	0.0182
3	0.1	Lin	1.37	–	0.0003	–	0.984	0.940	0.0116	0.0140
	0.2	RBF	1216.76	1132.01	–0.0003	–	0.984	0.940	0.0118	0.0142
	0.3	Lin	1.09	–	–0.0002	–	0.980	0.924	0.0132	0.0179
4	0.1	Lin	1.23	–	–0.0004	–	0.984	0.937	0.0119	0.0149
	0.2	RBF	459.93	419.69	–0.0001	–	0.988	0.923	0.0101	0.0163
	0.3	Lin	1.61	–	–0.0002	–	0.988	0.908	0.0103	0.0190
5	0.1	Lin	4.95	–	–0.0003	–	0.985	0.928	0.0113	0.0155
	0.2	Lin	1.29	–	–0.0002	–	0.985	0.919	0.0113	0.0188
	0.3	Lin	1.33	–	–0.0001	–	0.980	0.910	0.0132	0.0217
6	0.1	Lin	5.43	–	0.0001	–	0.983	0.744	0.0118	0.0408
	0.2	RBF	4259.04	1146.82	–0.0002	–	0.990	0.920	0.0094	0.0164
	0.3	Lin	2.86	–	–0.0008	–	0.981	0.922	0.0127	0.0274

Bold values represent the best performance

Fig. 12 Comparison of the CEEMD-SVR model with observed data

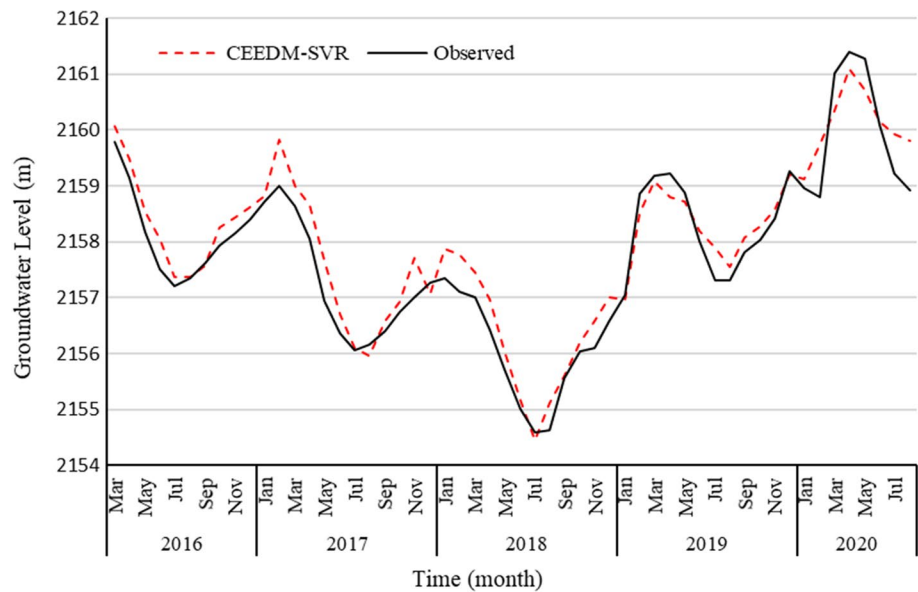


Table 10 Comparison of the models used in this study

Model Type	R^2		RMSE		AIC	
	Train	Test	Train	Test	Train	Test
SVR	0.972	0.918	0.0156	0.0168	305.35	91.66
ANN	0.985	0.927	0.0113	0.0158	304.07	91.41
W-ANN	0.984	0.938	0.0116	0.0149	304.18	91.18
W-SVR	0.992	0.949	0.0086	0.0164	302.96	91.57
CEEMD-SVR	0.980	0.948	0.0133	0.0154	304.72	91.32
CEEMD-ANN	0.997	0.998	0.0052	0.0035	300.99	85.35

Bold values represent the best performance

Comparison of the results of the models

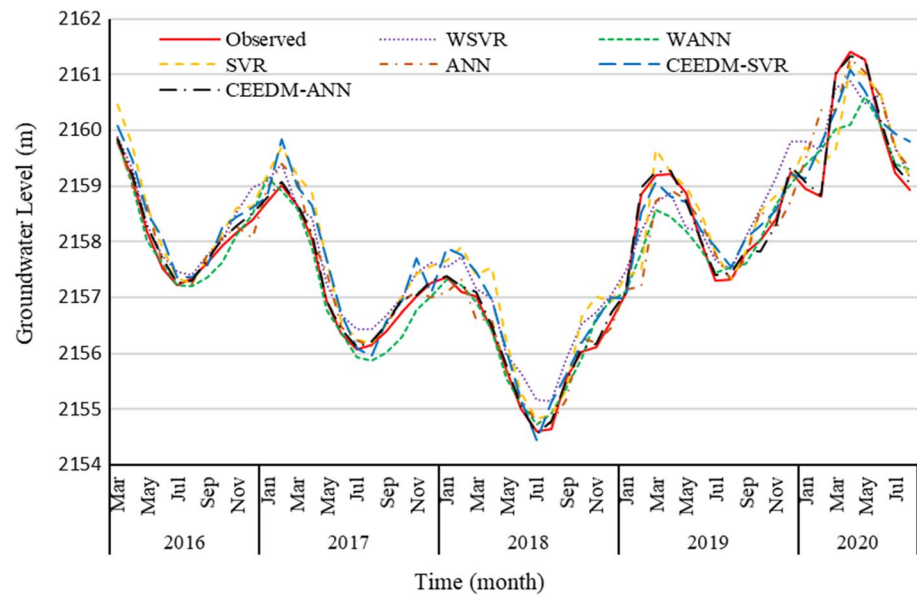
This section compares the results obtained from the execution of different models in different modes. According to the reference period used for the data of the testing step, the different methods were compared in the years 2016 to 2020. As can be seen in Table 10, the use of preprocessing tools improved the performance of the ANN and SVR models, such that using the WT resulted in an improvement of 1.10% in the ANN model and 3.07% in the SVR model. Furthermore, the use of the CEEMD model resulted in improvements by 7.08% and 2.89% in the ANN and SVR models, respectively. A comparison of the results reveals that the hybrid CEEMD-ANN model with a coefficient of determination of 0.99 and an error of 0.0035 had the best performance. This model had an Akaike coefficient of 85.5, which was lower than other models. Figure 13 compares the different models in the testing period based on the unit hydrographs of the aquifer. As can be seen, the hybrid models had a better performance, among which the hybrid CEEMD-ANN model

was closer to the observational values. Therefore, this model can be used in forecasting the GWL of the Aspas aquifer.

The results of this study are consistent with the results of Bahmani et al. 2020; Bahmani and Ouarda 2021, based on the better performance of models combined with WT and CEEMD. It is also consistent with the results of Adamowski and Chan (2011) and Solgi et al (2017) who stated that the model combined with the artificial neural network had the best performance.

Due to the fact that preprocessing tools divide initial signals into several sub-signals and these sub-signals are included in the models, this preprocessing of the data causes the execution time of the models to be reduced and the performance to be increased. So, in the artificial neural network model, it usually reduces the number of layers and neurons, and in fewer repetitions, it achieves better results, and in the SVR model, it also shortens the execution time of the program. On the other hand, the use of the standardization formula is another effective parameter in improving the result, which was used in all simple and combined models in this study.

Fig. 13 Comparison of the models used-step of test



Selecting a suitable decomposition level affects the accuracy of hybrid models. A high level of decomposition is not always helpful to increase the accuracy of the model and an optimal decomposition level must also be identified.

The results of CEEMD-based models illustrate the importance of ϵ for modeling and its impact on the accuracy of the hybrid models. No specified rule for ϵ determination has been presented and selecting ϵ depends on the features of a time series such as mean and extreme values. Therefore, it is recommended to find an optimal value for ϵ , for a given specific hydrological time series.

Conclusions

The aim of this paper was to identify the most reliable modeling technique in prediction of GWL monthly drop through comparison of models outcomes with observed GWL in existing wells. Two intelligent modeling techniques, ANN and SVR, were combined with two preprocessing tools WT and CEEMD resulting in four hybrid models. These hybrid methods were utilized to predict the GWL variations in the Aspas alluvial aquifer located in Tashk--Bakhtegan and Maharlu lakes basin. Model input data included precipitation, temperature, evaporation, and GWL of past months which were standardized before being used in the models. After model training process and during model testing phase, all models were implemented for the reference period 2016–2020 for which real GWL was available from existing observation wells. This allows to assess which model provides the closest fit to the real observed data. The performance of each modeling technique was evaluated by three main criteria including R^2 , RMSE, and AIC. The

results indicated that the hybrid models outperformed the intelligence models of ANN and SVR due to decomposed and de-noised input data. Among hybrid models, CEEMD-ANN was found to be the most accurate one with 7.08% performance improvements as compared to ANN model. Therefore, it is recommended to use this hybrid model (CEEMD-ANN) in forecasting the GWL variations in the Aspas aquifer according to findings of this paper.

There are a number of limitations in this research which can be pursued and further developed in future studies. Firstly, the proposed models used existing observation wells data as input data which is not usually available for recent months and might revert back up to a year ago. This undermined the need and importance of up-to-date data for purposes of water resource planning and management. To remediate this issue, the authors recommend utilizing GRACE satellite data and precipitation, temperature and evaporation satellite data as input data for intelligent models which are usually available for recent months and even days. Secondly, additional intelligent models such as Random Forest can be implemented in future studies to assess whether they provide more accuracy in GWL prediction.

Acknowledgements The authors are grateful to the Research Council of the Shahid Chamran University of Ahvaz for financial support. Also great thanks to the Regional Water Company of Fars and Iran Water Resources Management Company for sharing the required data.

Funding The authors received funding from Shahid Chamran University of Ahvaz (GN: SCU.WH99.589).

Declarations

Conflict of interest The authors declare that they have no known competing personal relationships that could have appeared to influence the work reported in this paper

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adamowski J, Chan FH (2011) A wavelet neural network conjunction model for groundwater level forecasting. *J Hydrol* 407(1–4):28–40. <https://doi.org/10.1016/j.jhydrol.2011.06.013>
- Alborzi M (2001) Introduction to neural networks. Scientific Publishing Institute of Industrial University Sharif, Tehran ((Persian))
- Amirat Y, Benbouzid M, Wang T, Bacha K, Feld G (2018) EEMD-based notch filter for induction machine bearing faults detection. *Appl Acoust* 133:202–209. <https://doi.org/10.1016/j.apacoust.2017.12.030>
- Ashraf S, Nazemi A, AghaKouchak A (2021) Anthropogenic drought dominates groundwater depletion in Iran. *Sci Rep* 11(9135):1–10. <https://doi.org/10.1038/s41598-021-88522-y>
- Awajan AM (2019) A review on empirical mode decomposition in forecasting time series. *Italian J Pure Appl Math* 42:301–323
- Bahmani R, Ouarda TBMJ (2021) Groundwater level modeling with hybrid artificial intelligence techniques. *J Hydrol* 595:1–12. <https://doi.org/10.1016/j.jhydrol.2020.125659>
- Bahmani R, Solgi A, Ouarda TBMJ (2020) Groundwater level simulation using gene expression programming and M5 model tree combined with wavelet transform. *Hydrol Sci J* 65(8):1430–1442. <https://doi.org/10.1080/02626667.2020.1749762>
- Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20:273–295
- Eskandari A, Solgi A, Zarei H (2018) Simulating fluctuations of groundwater level using a combination of support vector machine and wavelet transform. *J Irri Sci Eng* 41(1):165–180 ((Persian))
- Eskandari A, Faramarzyan yasuj F, Solgi A, Zarei H (2019) Evaluation of combined ANFIS with wavelet transform to modeling and forecasting groundwater level. *J Water Manag Res* 9(18):56–69 ((Persian))
- Foufoula-Georgiou E, Kumar P (1994) Wavelet in geophysics: an introduction. Academic Press, San Diego New. <https://doi.org/10.1016/B978-0-08-052087-2.50007-4>
- Huang Y, Schmitt FG, Lu Z, Liu Y (2009) Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis. *J Hydrol* 373(1–2):103–111. <https://doi.org/10.1016/j.jhydrol.2009.04.015>
- Karamooz M, Araghi Nejad SH (2010) Advanced hydrology, 2nd edn. Amirkabir University of Technology Press, Tehran, p 464 ((Persian))
- Mallat S (1998) A wavelet tour of signal processing. Academic Press is an imprint of Elsevier, San Diego
- Mirarabi A, Nassery HR, Nakhaei M, Adamowski J, Akbarzadeh AH, Alijani F (2019) Evaluation of data-driven models (SVR and ANN) for groundwater-level prediction in confined and unconfined systems. *Environ Earth Sci* 78(15):478–489. <https://doi.org/10.1007/s12665-019-8474-y>
- Moosavi V, Vafakhah M, Shirmohammadi B, Ranjbar M (2014) Optimization of wavelet-ANFIS and wavelet-ANN hybrid models by taguchi method for groundwater level forecasting. *Arabian J Sci Eng* 39(3):1785–1796. <https://doi.org/10.1007/s13369-013-0762-3>
- Niu D, Wanga K, Sun L, Wua J, Xu X (2020) Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: a case study. *Appl Soft Comput* 93:1–14. <https://doi.org/10.1016/j.asoc.2020.106389>
- Nourani V, Komasi M, Mano A (2009) A multivariate ANN-wavelet approach for rainfall–runoff modeling. *Water Res Manag* 23:2877–2894. <https://doi.org/10.1007/s11269-009-9414-5>
- Raghavendra NS, Deka PC (2014) Support vector machine applications in the field of hydrology: a review. *Appl Soft Comput* 19:372–386. <https://doi.org/10.1016/j.asoc.2014.02.002>
- Rajaei T, Ebrahimi H, Nourani V (2019) A review of the artificial intelligence methods in groundwater level modeling. *J Hydrol* 572:336–351. <https://doi.org/10.1016/j.jhydrol.2018.12.037>
- Salehi SM, Radmanesh F, Zarei H, Mansouri B, Solgi A (2019) A combined time series-wavelet model for prediction of groundwater level (case study: firuzabad plain). *J Irri Sci Eng* 41(4):1–16. <https://doi.org/10.22055/JISE.2018.14057>. ((Persian))
- Sang YF, Wang Z, Liu C (2012) Period identification in hydrologic time series using empirical mode decomposition and maximum entropy spectral analysis. *J Hydrol* 424–425:154–164. <https://doi.org/10.1016/j.jhydrol.2011.12.044>
- Sattari MT, Mirabbasi R, Shamsi Sushab R, Abraham J (2017) Prediction of groundwater level in ardebil plain using support vector regression and M5 tree model. *Nat'l Groundwater Assoc* 56(4):636–646. <https://doi.org/10.1111/gwat.12620>
- Solgi A, Zarei H, Nourani V, Bahmani R (2017) A new approach to flow simulation using hybrid models. *Appl Water Sci* 7:3691–3706. <https://doi.org/10.1007/s13201-016-0515-z>
- Solgi A (2014) Stream flow forecasting using combined neural network wavelet model and comparison with adaptive neuro fuzzy inference system and artificial neural network methods (case study: Gamasyab river, Nahavand). M.Sc. Thesis, Department of Hydrology and Water Resource, Shahid Chamran University of Ahvaz
- Suryanarayana CH, Sudheer CH, Mahammood V, Panigrahi BK (2014) An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam. *India Neurocomputing* 145:324–335. <https://doi.org/10.1016/j.neucom.2014.05.026>
- Wang K, Niu D, Sun L, Zhen H, Liu J, De G, Xu X (2019) Wind power short-term forecasting hybrid model based on CEEMD-SE method. *Processes* 7(11):843. <https://doi.org/10.3390/pr7110843>
- Wu Z, Huang NF (2004) A study of the characteristics of white noise using the empirical mode decomposition method. *Proc R.S Lond* 460A:1597–1611. <https://doi.org/10.1098/rspa.2003.1221>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.