International Journal of Agriculture and Biosciences 2025 14(3): 424-435.
https://doi.org/10.47278/journal.ijab/2025.024

RESEARCH ARTICLE eISSN: 2306-3599; pISSN: 2305-6622

Groundwater Level Simulation using Hybrid Model

Shiva Khosravi and Amir Robati *

Department of Civil Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran
*Corresponding author: Gammirrr@gmail.com

ABSTRACT

Simulation of groundwater level fluctuations as the most important source of water supply plays a substantial role in gathering information for planning and managing water resources.
This study introduces a hybrid methodology combining discrete wavelet transforms (DWT) with artificial neural networks (ANN) to simulate groundwater level fluctuations. The optimal mother wavelet was considered, and the simulated values were compared to those generated by a robustly intelligent tool, such as weighted least squares support vector machine (WLS-SVM). To assess the robustness and efficiency of the DWT-ANN model, monthly groundwater level time-series data from three observation wells in the Bagheyn Plain, Iran, were used. Statistical indicators, including mean absolute error (MAE), root mean squared error (RMSE), and Nash-Sutcliffe efficiency (NSE), were calculated to evaluate the models' performance. Results showed that the DWT-ANN model achieved superior performance, yielding MAE, RMSE, and NSE values of 0.044, 0.064, and 0.9998, respectively, at the Saadi observation well. These findings underscore the DWT-ANN model's superiority over the WLS-SVM model in simulating groundwater levels for the selected wells. Furthermore, the DWT-ANN approach demonstrated enhanced accuracy, with simulated values closely aligning with observed data.

Keywords: Groundwater; Simulation; Wavelet transforms; Artificial neural network.

Article History
Article # 24-972
Received: 15-Nov-24
Revised: 07-Jan-25
Accepted: 09-Jan-25
Online First: 11-Feb-25

INTRODUCTION

Groundwater is an essential resource for potable water supply, playing a significant role in supporting environmental and economic activities. However, the over-extraction of aquifers has led to declining groundwater levels, which underscores the necessity for sustainable groundwater management (Adamowski, 2011). Accurate groundwater level simulations are crucial for developing effective strategies to manage these resources efficiently (Rahman et al., 2020). Traditional methods for measuring groundwater levels, such as advanced water level loggers and acoustic or optical sounding devices, are time-intensive and expensive. Therefore, artificial intelligence (AI) techniques offer a cost-effective alternative for simulating groundwater levels. Among these techniques, support vector machines (SVMs), artificial neural networks (ANNs), and wavelet transformations have proven highly effective for simulating nonlinear hydrological processes (Rajaee et al., 2019; Tao et al., 2022; Boo et al., 2024).

Support vector machines, first proposed by Vapnik (2013), have been successfully utilized for diverse modeling and simulation problems. Recent studies on groundwater level modeling using SVMs include works by Mallikarjuna et al. (2021), Anh et al. (2023), and Satapathy et al. (2023). Suykens et al. (2002) introduced the weighted least squares support vector machine (WLS-SVM), which has demonstrated superior accuracy compared to both SVM and the least squares version of SVM (LS-SVM). Several researchers have applied WLS-SVM in groundwater simulations, reporting minimal differences between observed and simulated values (Liu et al., 2009; Tang et al., 2019; Moravej et al., 2020). Additionally, Samani et al. (2023) found that the wavelet-LS-SVM model outperformed other standalone and hybrid wavelet-based AI methods for simulating groundwater levels.

Numerous studies have been conducted on the use of ANN models to simulate groundwater levels (Lohani et al., 2015; Lee et al., 2019; Derbela et al., 2020; Wunsch et al., 2021). ANN models, widely used for groundwater simulations, encounter challenges when processing non-stationary time-series data. These limitations necessitate data pre-processing (Zare et al., 2018; Zhang et al., 2019; Yadav et al., 2020; Samani et al., 2023).

Cite this Article as: Khosravi S and Robati A, 2025. Groundwater level simulation using hybrid model. International Journal of Agriculture and Biosciences 14(3): 424-435. https://doi.org/10.47278/journal.ijab/2025.024

A Publication of Unique Scientific Publishers

The wavelet transformation is a powerful tool for analyzing non-stationary hydrological signals (Hsu et al., 2010; Zhu et al., 2020; Gordu & Nachabe, 2021). Wavelet analysis decomposes time-series data in to approximation and details components. The new time-series data is utilized as input data for the ANN model. Therefore, the wavelet approach is combined with artificial intelligence methods such as ANN to create a wavelet–AI model.

The use of wavelet–ANN to simulate groundwater levels has been addressed in several research studies over the past years (Moosavi et al., 2013; Zhou et al., 2017; Ebrahimi et al., 2017). These studies highlight how wavelets have the capability to improve the performance of ANN for simulating groundwater levels. Nourani et al. (2014) provided an interesting review study on hybrid models (combining wavelet and artificial intelligence) for hydrological processes, such as groundwater level fluctuations. They recommended that the authors conduct future research in these fields. Recent studies, such as those by Vousoughi (2023) and Shahbazi et al. (2023), have also affirmed the enhanced performance of wavelet-ANN methods for simulating groundwater levels.

Rajaee et al. (2019) and Boo et al. (2024) demonstrated in their review studies that the largest number of studies on groundwater level simulation using AI methods were conducted in Iran. This point maybe attributed to the necessity of conducting more groundwater studies in arid and semi-arid regions such as Iran. Therefore, there is a need for an accurate groundwater level simulation model in these regions. Our research purpose is to develop the most accurate model for simulating groundwater level. A highly accurate groundwater level simulation model is a powerful tool for sustainable groundwater resource management. Accordingly, a hybrid approach combining discrete wavelet transforms (DWT) and artificial neural networks (ANN) is proposed, which selects the optimal wavelet transform for simulating groundwater fluctuations. Simulated values are compared with WLS-SVM results to assess the models' effectiveness. Statistical analysis confirms that the DWT-ANN method achieves superior accuracy compared to previously established models, such as those by Ebrahimi et al. (2017) and Wei et al. (2023).

MATERIALS & METHODS

Study Area and Data

The Bagheyn plain spans an area of 5420km^2 and is situated between latitudes 29°47′ to 30°31′ N and longitudes 56°18′ to 57°37′ E in Kerman Province, southeastern Iran (Fig. 1). (Fig. 1). Approximately 50% of the plain is mountainous, contributing significantly to aquifer recharge. The region's highest elevation, at 4,189 meters, is in the southwestern area, while the lowest elevation, at 1,626 meters, is in the northeast. The general slope of the plain trends from south to northwest. Within the plain, there are 65 observation wells, of which three were selected for this study (Fig. 1). The studied aquifer is unconfined, with an alluvial thickness ranging from 100 to 250 meters. The groundwater level time-series in the 10-year study period shows that the amount of groundwater recharge in the area was almost constant, with no significant fluctuations observed. The survey of water level fluctuations shows that groundwater reaches its highest level in November and its lowest in June. The depth of the groundwater level in the plain varies from 9.13 to 170.89 meters as of August 2020). The aquifer exhibits hydraulic conductivity values between 2 and 15 m/day and transmissivity coefficients of 200 to 3,000m^2/day. The storage coefficient also varies from at least 2% to a maximum of 20% in different sections of the plain. The average annual temperature for the Bagheyn plain ranges from 12.8°C to 15.6°C, with the average annual rainfall in this region ranging from 147 to 204.2mm at the Kerman station (during 2019). The annual precipitation occurs from December to April, with these five months accounting for 87% of the total annual precipitation.

Given the region's susceptibility to water shortages, groundwater serves as a critical resource for domestic, agricultural, and industrial purposes. Accurate simulation models are therefore essential for effective water resource management in the Bagheyn Plain.

The study area and the locations of the three selected observation wells in the Bagheyn Plain shown in Fig. 1. For the simulation of groundwater levels, monthly piezometric head data from April 2009 to March 2019 (120 months) were collected from three selected observation wells.
The data for the simulation was divided into two categories: training (75%) and testing (Fig. 2).

Table 1 provides the summary statistics of groundwater levels recorded at the three observation wells during the study period. These statistics include details about the wells' locations and descriptive data such as the range and mean of groundwater levels. This breakdown ensures a comprehensive understanding of the dataset used for model development and evaluation.

Table 1: Properties of observation wells' location and statistics of groundwater level in the Bagheyn plain during Apr 2009–Mar 2019.

Well ID	Wells name	Wells' location		Water table level (m)			
		Longitude	Latitude	Min	Max	Mean	SD
1	Saadi	473500	3341500	74.38	88.45	81.38	4.00
2	Dasht-Bagheyn	497200	3335750	56.41	63.46	60.31	2.12
3	Tolombe Badi	514650	3341250	54.93	60.37	57.78	1.70

Weighted Least Squares Support Vector Machines (WLS-SVM)

Support vector machines (SVMs) have been successfully applied to numerous machine learning problems, including groundwater level simulation (Suykens et al., 2002). The weighted least squares support vector machine (WLS-SVM) introduced by Suykens et al. in 2002, has proven to be more robust than the least squares version of SVM (LS-SVM). The interesting research by Suykens et al. (2002) provides greater motivation for applying the WLS-SVM method in groundwater level simulation.

The WLS-SVM is characterized by an optimization function, which is defined in the primal weight space as follows (Suykens et al., 2002):

$$\min \mathrm{q}\,(\mathrm{w}, \mathrm{e}) = \frac{1}{2}\|\mathrm{w}\|^2 + \frac{1}{2}\gamma \sum_{\mathrm{k}=1}^{\mathrm{n}} \overline{\mathrm{v_k}}\, \mathrm{e_k^2} \quad (1)$$

Subject to

$$\mathrm{y_k} = \mathrm{w^T}\phi(\mathrm{x_k}) + \mathrm{b} + \mathrm{e_k} \quad \mathrm{k} = 1., \mathrm{n} \quad (2)$$

With $\{x_i, y_i\}_{k=1}^{n}$ being a set of n training samples, where $x_k \in R^n$ represents input vector data and $y_k \in R$ is output data. Additionally, $\phi(.): \mathrm{R^n} \to \mathrm{R^d}$ is a function that maps the

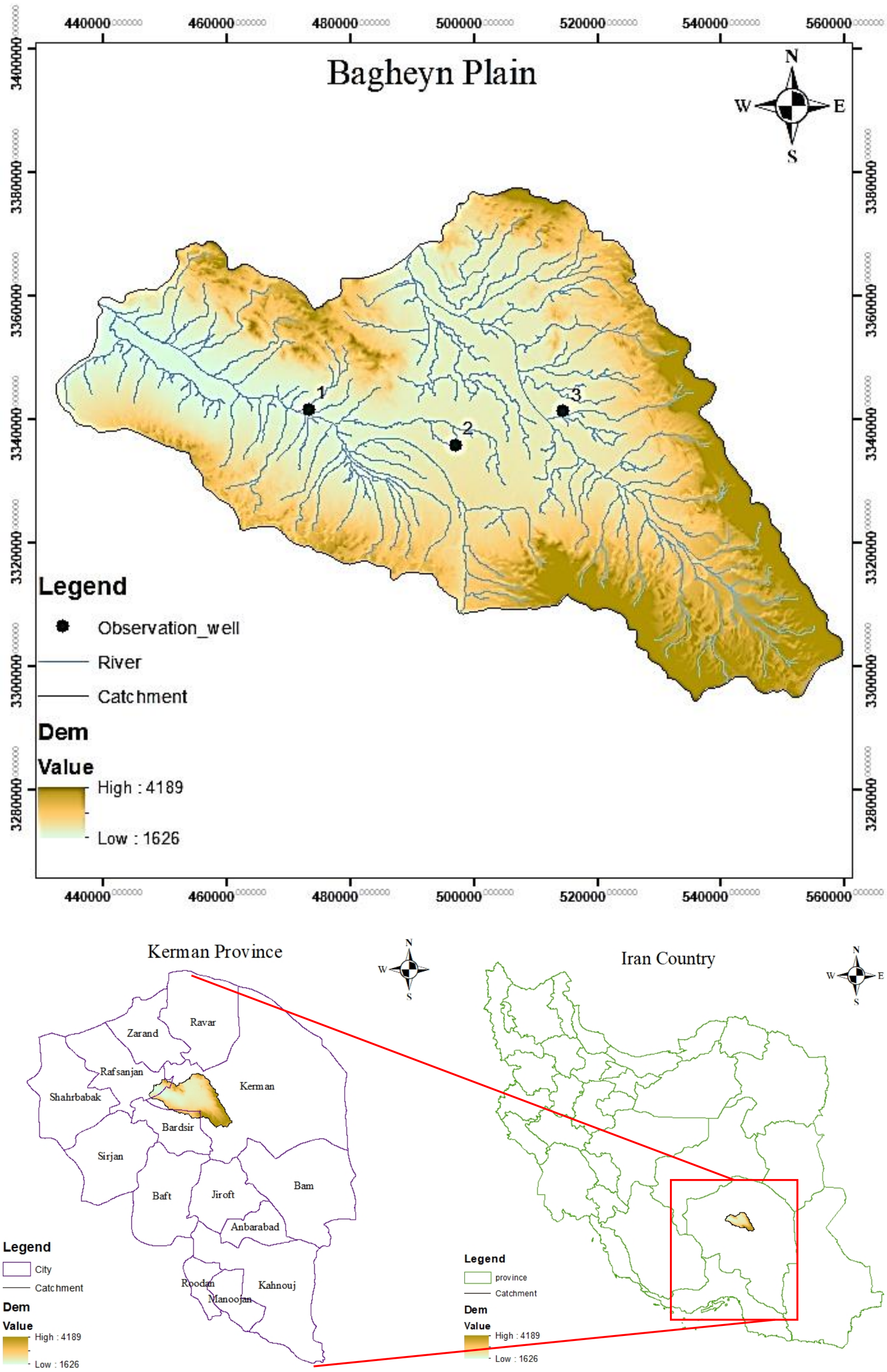


Fig. 1: The study area and the locations of observation wells.

input space into a higher dimensional feature space, $w \in R^d$ is the weight vector in primal weight space, $e_k \in R$ are error variables, b is the bias term, and γ is an adjustable constant. The weights $\overline{v}_k$ are determined based on the relationships proposed by David et al. (1998).

The model of WLS-SVM in the primal weight space is defined as follows:

$$y(x) = w^T\phi(x) + b \quad (3)$$

Due to the unknown structure of the function, it is not possible to calculate w from equation (1). Therefore, the model shown in equation (1) is calculated using the Lagrangian method as follows:

$$L(w, b, e; x) = q(w, e) - \sum_{k=1}^{n} \alpha_k \left(w^T\phi(x_k) + b + e_k - y_k\right) \quad (4)$$

The solution can be derived from the Karush-Kuhn-Tucker (KKT) conditions for optimality by eliminating variables w and e, resulting in the following system of linear equations:

$$\begin{bmatrix} \Omega + V_y & 1_n^T \\ 1_n & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix} \quad (5)$$

where

$$V_y = diag\{1/\gamma\overline{v}_1 . 1/\gamma\overline{v}_n\}$$
$$\Omega_{k,l} = \langle \phi(x_k), \phi(x_l) \rangle_H \; k, l = 1,2, . n$$
$$y = [y_1, . y_n]^T \quad (6)$$
$$1_n^T = [1 ., 1])$$
$$a = [a_1 ., a_n]$$

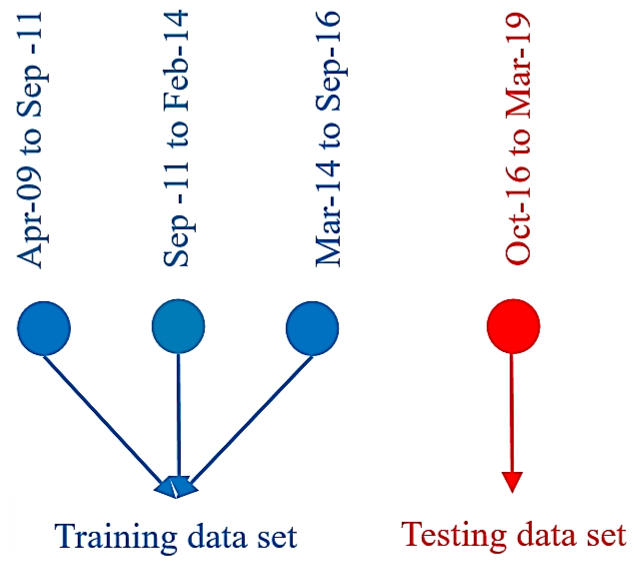


Fig. 2: Four-fold training and testing datasets.

A kernel $K(.,.)$is selected according to Mercer's condition, such that:

$$k(x_k, \overline{x_l}) = \langle \phi(x_k), \phi(x_l) \rangle_H \quad (7)$$

Consequently, the WLS-SVM model developed for function prediction is as follows:

$$y(x) = \sum_{k=1}^{n} \alpha_k k(x_k, x) + b \quad (8)$$

Where α_k and b are the solutions to the linear system.

In the present study, the radial basis function (RBF) with the parameter σ was used as the kernel function.

$$K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}) \quad (9)$$

Where σ is the kernel width.

Several studies have demonstrated that the RBF kernel provides strong performance in simulating hydrological variables. It has also been used in recent studies for groundwater level simulation (Tang et al., 2019).

Artificial Neural Networks (ANN)

An artificial neural network (ANN) consists of nodes known as artificial neurons, which model the neurons in a biological brain. In the brain's process, myelinated axons and neurons connect inputs at dendrites to outputs at axon terminals with signal flow. The artificial neural network receives (X_i) as inputs, recognizes the activation, and produces outputs (Y_i) depending on the input and activation. The multi-layer perceptron (MLP) employed in the research had three layers: input layer, hidden layer(s), and output layer. The chosen activation function for both the hidden and output layers was a tansig (Hyperbolic tangent sigmoid) function, which was determined to be the most effective through a process of trial and error. The Levenberg-Marquardt algorithm is used due to better performance and speed of training in the research. Each circular node symbolizes an artificial neuron, while an arrow indicates the connection from the output of one artificial neuron to the input of another (Fig. 4). The number of neurons in the hidden layer and delays were achieved using a trial-and-error approach and the data-series were categorized into three subsets: training, validation, and testing. Artificial neural networks have been proven to be beneficial tools for simulating future values based on past histories in the hydrological process (Nourani et al., 2014; Wunsch et al., 2021).

In groundwater level forecasting, X_i represents n-dimensional input vector of groundwater level values at different antecedent time lags and Y_i is n-dimensional output vector of the groundwater level for a subsequent period at a well (Nayak et al., 2006). The ANN model for groundwater level simulation was developed utilizing the MATLAB R2022a software program.

Wavelet Analysis

The concept of wavelets was first proposed by Morlet and Meyer developed the methods of wavelet analysis. Recently wavelet transforms have been investigated as a powerful tool for the analysis and de-noising of time series data, providing accurate results in hydrological processes. The groundwater level time series can be denoted as $\{X_t: t = 1,..,n\}$ where t represents the time index and n indicates the total number of groundwater level observations. If the time series properties do not change much over time, it is called a stationary signal. Groundwater level time series contain numerous nonstationary characteristics; thus, wavelets are ideal for studying nonstationary groundwater level data, where the mean and autocorrelation of the signal fluctuate over time.

A mother wavelet $\psi(t)$ is a mathematical function with a zero average (Mallat, 1999).

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \quad (10)$$

The wavelet transform breaks down signals into dilated and translated wavelets. A mother wavelet $\psi(t)$ is a waveform of effectively limited duration and can be expressed as follows (Hajizadeh et al., 2016):

$$\psi_{s,u}(t) = \frac{1}{\sqrt{|s|}} \psi(\frac{t-u}{s}) \quad (11)$$

Where s $(s > 0)$ is a scaling parameter or a scale that measures the degree of compression and u is the translation parameter that determines the time location of the wavelet.

The wavelet transforms are divided into discrete wavelet transform (DWT) and continuous wavelet transform (CWT).

The continuous wavelet transforms (CWT) of a function f(t) is described as follows:

$$f(u, s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^*(\frac{t-u}{s})dt \quad (12)$$

Where $\psi^*(t)$ represents the operation of the complex conjugate.

In the discrete wavelet transforms (DWT) concept, the scales s and u are considered as powers of 2 in the mother wavelet as follows:

$$s = 2^j; \quad u = 2^j n; \quad j, n \in Z \quad (13)$$

where Z is a set of integers.

By substituting the parameter values u and s in Eq. (11), the discrete wavelet transform (DWT) is obtained as follows:

$$\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi(\frac{t-2^j n}{2^j}) \quad (14)$$

Where j and n are the dilation and the translation parameters, respectively.

The original signal S is processed through both low-pass and high-pass filters and transforms the time series into approximation (A) and details (D) signals, respectively (Fig. 3).

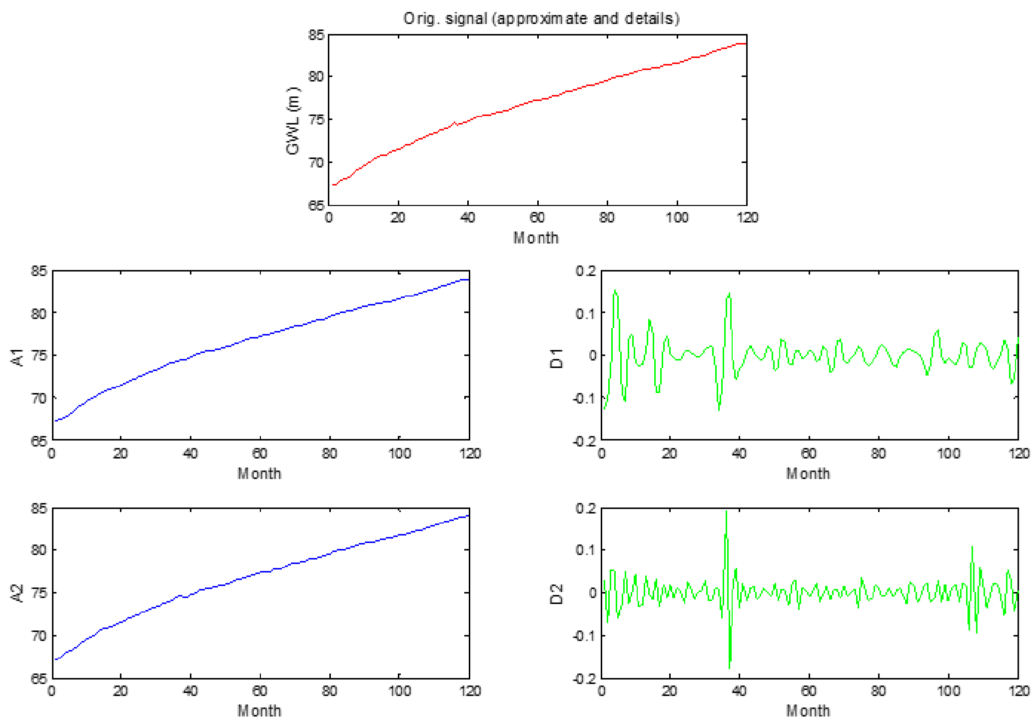


Fig. 3: Wavelet decomposition of groundwater level time series using Db4 wavelet at level two.

The selection of an appropriate efficient mother wavelet and decomposition level are two important issues in wavelet analysis. In the present study, the decomposition of the groundwater level time series was employed with the Daubechies mother wavelets (Db4), which was proven in past studies (Moosavi et al., 2013; Rajaee et al., 2019, Saroughi et al., 2024) to have a robust performance in simulating the hydrological process (Fig. 4).

The decomposition level of the wavelet depends on the historical data length (Karthika et al., 2015; Nury et al., 2017; Shahbazi et al., 2023):

$$L = int[log\, N_S] \quad (15)$$

Where

L = decomposition level,

N_S = length of the time-series data

The Proposed Hybrid Model of DWT–ANN

Although ANN indicate significant performance in simulation, there are still drawbacks present in ANN. A problem with artificial neural networks is the presence of limitations when dealing with non-stationary data. Hydrological phenomena such as fluctuations in groundwater levels, are highly non-stationary. ANN models may not be able to accurately simulate hydrological time-series (Adamowski et al., 2010; Cui et al., 2022).

To enhance the accuracy of ANN, using wavelet transformation as a preprocessing method for time-series data has been suggested by Bahmani et al. (2020) and Shahbazi et al. (2023). Wavelet transforms decompose the original time series, and they can be very beneficial for non-stationary data. In this study, a hybrid model combining of DWT with ANN, referred to as DWT-ANN, proposed to resolve the aforementioned drawback. The Flow diagram of the DWT-ANN model is shown in Fig. 4.

The implementation of the DWT-ANN model includes the following steps:

Step 1: Select the original time series data (monthly groundwater level) and identify the length of historical data.
Step 2: Calculate the number of decomposition levels using Equation15.
Step 3: Utilize the mother wavelet (db4) to decompose the original time series data into approximation and details components (Fig. 4).
Step 4: Enter the time-lagged series of approximation and details into the ANN as input data.
Step 5: Use a trial-and-error approach to identify neurons in the hidden layer.
Step 6: Divide the data into training and testing data (Fig. 2).
Step 7: Run ANN code to predict the test series.

This hybrid approach leverages the strengths of wavelet decomposition to preprocess the data, effectively removing noise and enhancing the ANN model's ability to handle non-stationary signals. By combining DWT and ANN, this methodology addresses the limitations of standalone ANN models and improves the simulation of groundwater level fluctuations.

DWT–ANN Model Performance Evaluation

Statistical indicators were calculated to determine the adequacy of the DWT-ANN method compared to the WLS-SVM model, as one of the appropriately selected criteria. The mathematical expressions for these error estimates are presented in Equations (16) to (18).

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|S_i - O_i| \text{ Mean Absolute Error} \quad (16)$$

$$\mathrm{RMSE} = \left[\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2\right]^{\frac{1}{2}} \text{ Root Mean Squared Error} \quad (17)$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n}(S_i-O_i)^2}{\sum_{i=1}^{n}(\bar{S}_i-O_i)^2} \text{ Nash–Sutcliffe Efficiency} \quad (18)$$

Where O_i and S_i are observed and simulated value, respectively; and n represents the length of data.

The best efficiency is achieved when MAE and RMSE are close to zero, and NSE is near to 1. Model efficiency (NSE), as proposed by Nash and Sutcliffe (Duc et al., 2023), ranges from 0 to 1, with an acceptable level of performance.

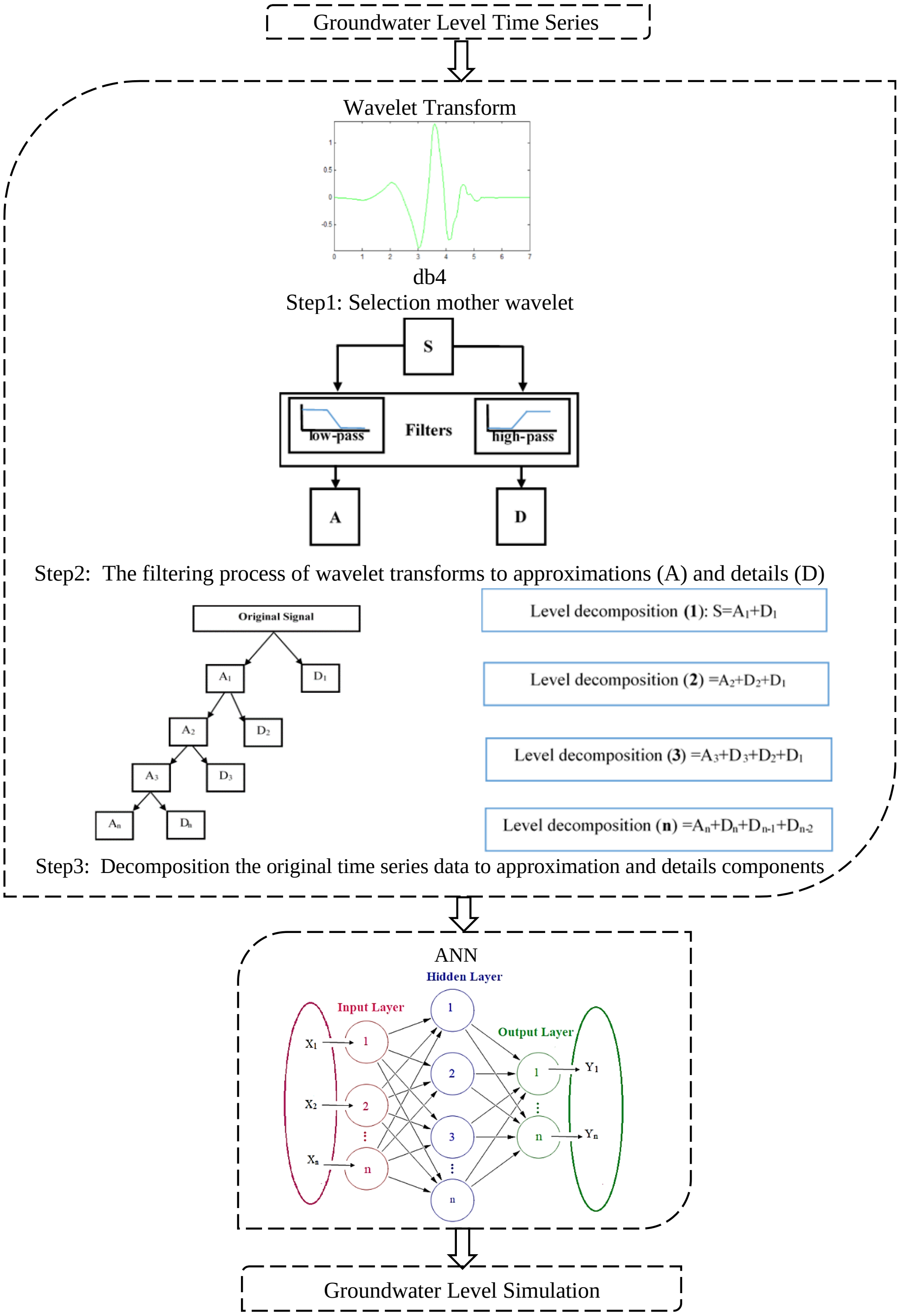


Fig. 4: DWT-ANN model

Additionally, the determination coefficient (R^2) is the degree of collinearity between observed and simulated data. R^2 also ranges from 0 to 1, with values greater than 0.5 considered acceptable (Hlaing et al., 2024).

The DWT-ANN method is used to evaluate monthly groundwater levels by training a novel method. To train the Artificial Intelligence model (AI), observed groundwater level data is broken down into two sub-series: approximations (A) and details (D). approximations A1, A2 and details D1, D2 are employed as inputs for the hybrid model. The inputs from 75% and 25% of the data are then utilized to train and test the performance of the DWT-ANN method.

RESULTS

The simulated groundwater level values for the DWT-ANN and WLS-SVM models were obtained at all three observation wells. The performance of both DWT-ANN and WLS-SVM models for the wells in terms of the three evaluation criteria provided in Table 2. Three criteria for performance evaluation, including MAE, RMSE and NSE were used to verify the robustness of the models. As shown in Table 2, both the DWT-ANN and WL-SSVM models had smaller values of MAE and RMSE at all three observation wells during the training phase compared to the testing phase. It seems logical that simulation models perform better in the training phase than the testing phase.

Across all wells, the minimum MAE and RMSE values achieved for DWT-ANN model at the Saadi well. The MAE and RMSE values were 0.044, 0.064 obtained for training phase respectively. The testing phase yielded the best results for the DWT-ANN model at the Tolombe-Badi well, with MAE and $RMSE$ values of 0.133, and 0.168, respectively. These values demonstrate a nearly perfect fit between observed and simulated results. The results in Table 2 indicate that both models achieved good simulation accuracy, with $RMSE$ values ranging from 0.168 for DWT-ANN to 0.257 for WLS-SVM in the testing phase, indicating that the DWT-ANN model has superior simulation capability.

A model can be considered accurate if the calculated NSE value is greater than 0.8 (Shu et al., 2008). It can be seen from Table 2 that the values of NSE for all cases are close to or above 0.8. Therefore, both the DWT-ANN and WLS-SVM models have good performance for simulating groundwater levels. NSE values also demonstrate satisfactory simulation accuracy, with DWT-ANN achieving the highest efficiency at 0.9998 (Fig. 5). At all three observation wells, higher NSE values and lower MAE and RMSE were observed for the DWT-ANN model compared to the WLS-SVM model. This outcome highlights the superiority of the DWT-ANN model over the WLS-SVM model, making DWT-ANN the most effective in simulating groundwater levels.

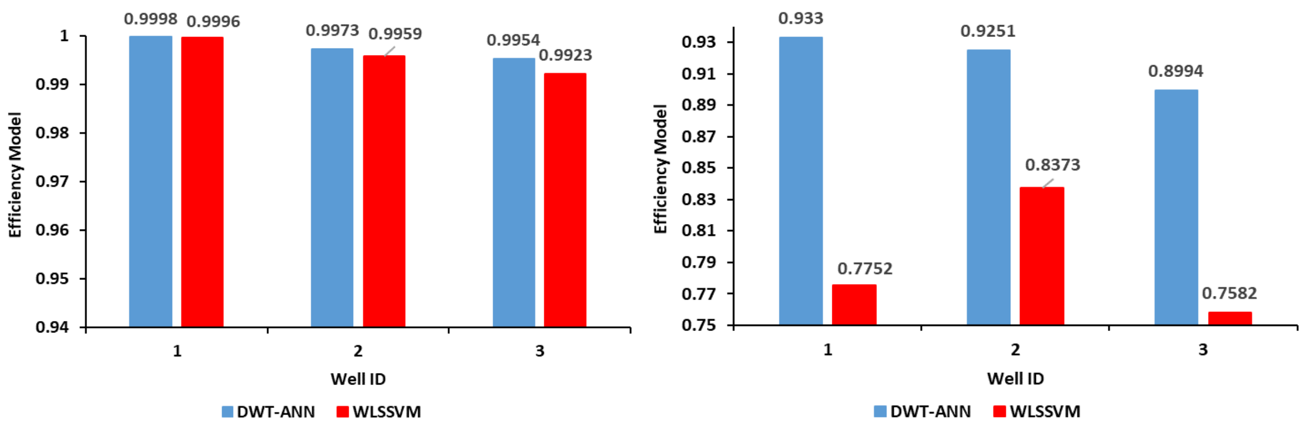
The comparison of simulated and observed values using DWT-ANN and WLS-SVM for three selected observation wells were illustrated in Fig. 6. At the three selected observation wells, the simulated values for the DWT-ANN were closer to the observed values compared to WLS-SVM model. Additionally, Fig. 7 and Fig. 8 display the scatter plots of observed groundwater level value versus simulated groundwater levels at three selected observation wells for DWT-ANN and WLS-SVM models in training and testing phases, respectively. R^2 describes the degree of collinearity between observed and simulated data, ranging from 0.9363 to 0.9976 for the DWT-ANN model and 0.8442 to 0.9944 for the WLS-SVM model in training and testing phases. The R^2 values for the DWT-ANN model at three selected observation wells are significantly higher than those for the WLS-SVM model, with the value of 0.9976 achieved at the Saadi well in training phase. Furthermore, the R^2 values for both the DWT-ANN and WLS-SVM models in the training phase are higher than those in the testing phase at all three selected observation wells. Therefore, the DWT-ANN model is considered a more accurate model for simulating groundwater levels. The results of the DWT-ANN model showed a closer collinearity between the simulated and observed groundwater level values compared to the WLS-SVM model.

DISCUSSION

The growth of urban populations in cities has resulted in a notable increase in water demand, especially for drinking purposes. However, the excessive exploitation and utilization of groundwater resources have led to an annual decline in groundwater levels. This problem of water scarcity is particularly prominent in arid and semi-arid regions (Zafar et al., 2023; Zafar et al., 2024). The Bagheyn Plain, located in Iran, is situated in the mid-latitude belt of semi-arid and arid regions of the Earth. Therefore, it has been selected for groundwater simulation in this study. Over the last two decades, numerous studies using various Artificial Intelligence (AI) methods such as Artificial Neural Networks (ANN) have been conducted to simulate groundwater levels (Bahmani & Ouarda, 2021). AI-models are capable of hydro-geophysical and topographical data. Due to this

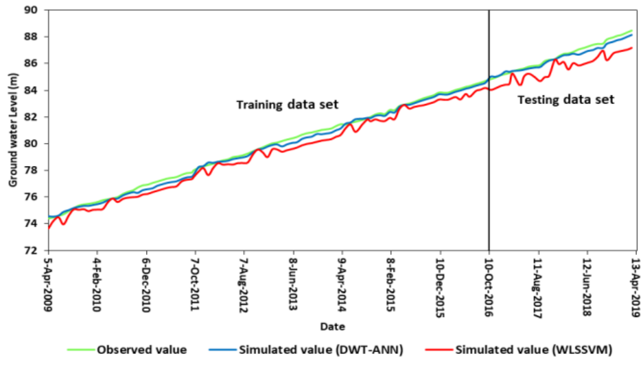
Table 1: The statistical parameters of both models.

Well ID	Well Name	Models	Training			Testing		
			MAE(m)	RMSE(m)	NSE	MAE(m)	RMSE(m)	NSE
1	Saadi	(DWT-ANN)	0.044	0.064	0.9998	0.241	0.257	0.9330
		WLS-SVM	0.061	0.086	0.9996	0.440	0.470	0.7752
2	Dasht-Bagheyn	(DWT-ANN)	0.059	0.010	0.9973	0.163	0.174	0.9251
		WLS-SVM	0.070	0.124	0.9959	0.212	0.258	0.8373
3	Tolombe- Badi	(DWT-ANN)	0.054	0.075	0.9954	0.133	0.168	0.8994
		WLS-SVM	0.069	0.097	0.9923	0.219	0.260	0.7582

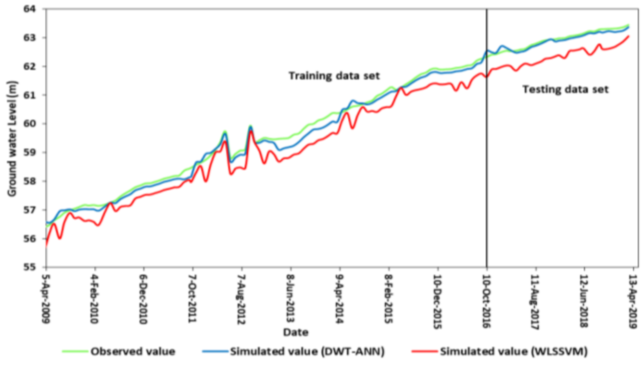


(a) Training phase (b) Testing phase

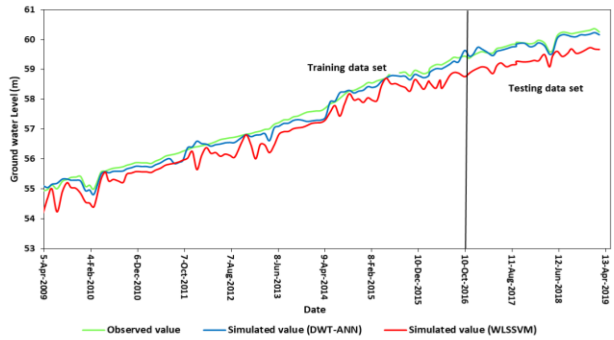
Fig. 5: Efficiency model (NSE) on the resulting groundwater level at the observation wells in DWT-ANN and WLS-SVM models; (a) Training phase, (b) Testing phase.



(a) Saadi well



(b) Dasht-Bagheyn well



(c) Tolombe-Badi wells

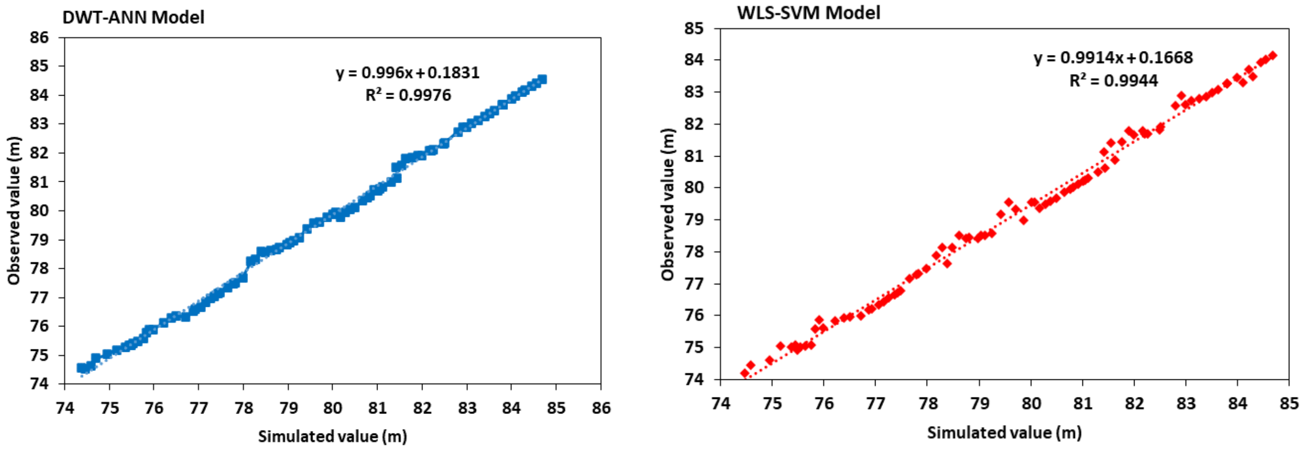
Fig. 6: Comparison of the observed groundwater level and the DWT-ANN and WLS-SVM simulated groundwater level at (a) Saadi, (b) Dasht-Bagheyn and (c) Tolombe-Badi wells.

ability, AI methods are more appealing compared to physically based and numerical methods. Since groundwater level fluctuations are a nonlinear phenomenon in the natural environment, it is necessary to consider robust models (Najafabadipour et al., 2022).

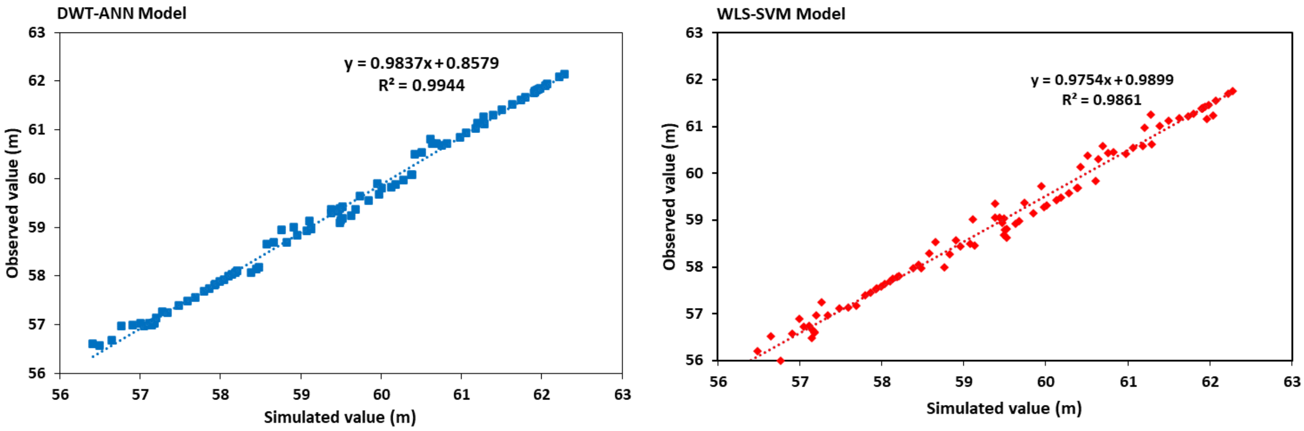
However, ANN models have shown limitations with complex non-linearity and non-stationary time-series data such as groundwater level fluctuations as demonstrated by prior research (Kayhomayoon et al., 2021; Osman et al., 2021; Cui et al., 2022); Therefore, there is a need for data transformation to improve hydrological non-stationary signals. Meanwhile, several studies have proven that wavelet data transformation is an effective method for removing noise from time-series data. They have also introduced the decomposed signal as an input variable to the AI models (Wu et al. 2021; Zhang et al., 2023; Wang et al., 2023). In the present study, the monthly groundwater level time-series of three observation wells in the Bagheyn plain during the time period of April 2009–March 2019 (120 months) were selected as input data for the model. According to the previous study (Shahbazi et al., 2023), 75% of the data were used for training and 25% were used for testing. Wavelet transform has two main forms: continuous wavelet transform (CWT) and discrete wavelet transform (DWT). However, in their 2019 review study, Rajaee et al. (2019) recommended using DWT for decomposing hydrological time-series data. In the present study, DWT was chosen for use in the hybrid model based on previous findings. According to previous studies on Wavelet-ANN models, the choice of the mother wavelet may impact the accuracy of hybrid models (Bahmani et al. 2020; Freire & Santos, 2020). Rajaee et al. (2019) recommended using the db2 and db4 mother wavelets for analyzing groundwater level time-series data due to their shapes. Furthermore, Wei et al. (2023) indicated that the Wavelet-ANN model with the db mother wavelet performed better than other mother wavelets for forecasting groundwater level. Moreover, Saroughi et al. (2024) emphasized that the Daubechies mother wavelets were the most appropriate wavelet for estimating groundwater level. In this study, based on previous findings, the db4 mother wavelet was chosen for use in the hybrid model.

Additionally, the accuracy of hybrid models can be influenced by selecting the appropriate decomposition level. A high level of decomposition may not always be beneficial in improving the accuracy of the model; it is important to identify the optimal level of decomposition. In the present study, two decomposition levels are utilized based on the length of historical groundwater level data (Shiri et al., 2021).

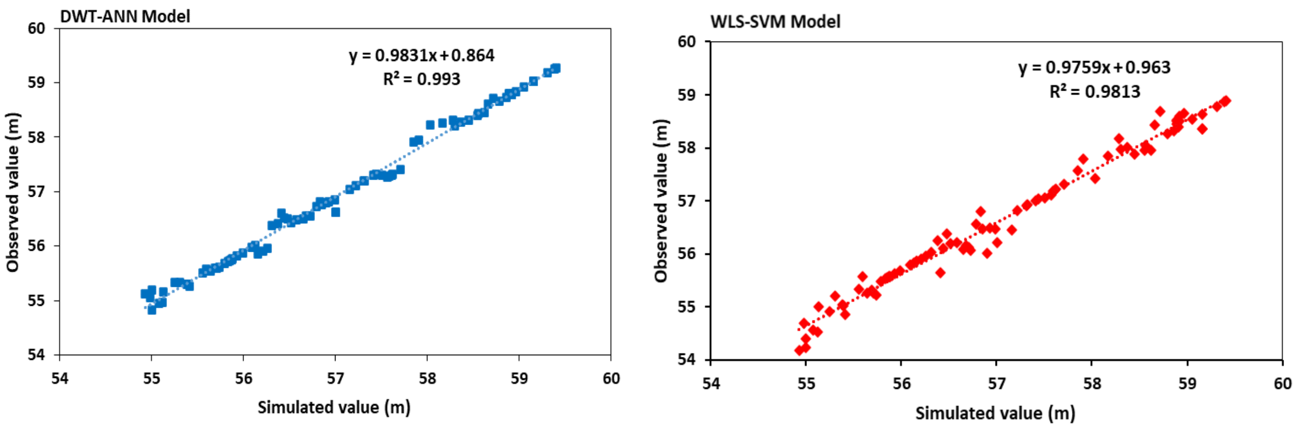
The use of wavelet transform for preprocessing groundwater level time-series data can improve the simulation accuracy of the ANN model. Considering all selected wells, the results of the present study indicate that the DWT-ANN performed better than the WLS-SVM model. The results of this study are agreed with those of several related studies. Graf et al. (2019) indicated the superior performance of the hybrid DWT-ANN model. Similarly, Freire & Santos (2020) demonstrated that combining ANN model with preprocessing tools like wavelet improved their performances. A study by Shahbazi et al. (2023) also supports these findings, emphasizing that hybrid models



(a) Saadi well



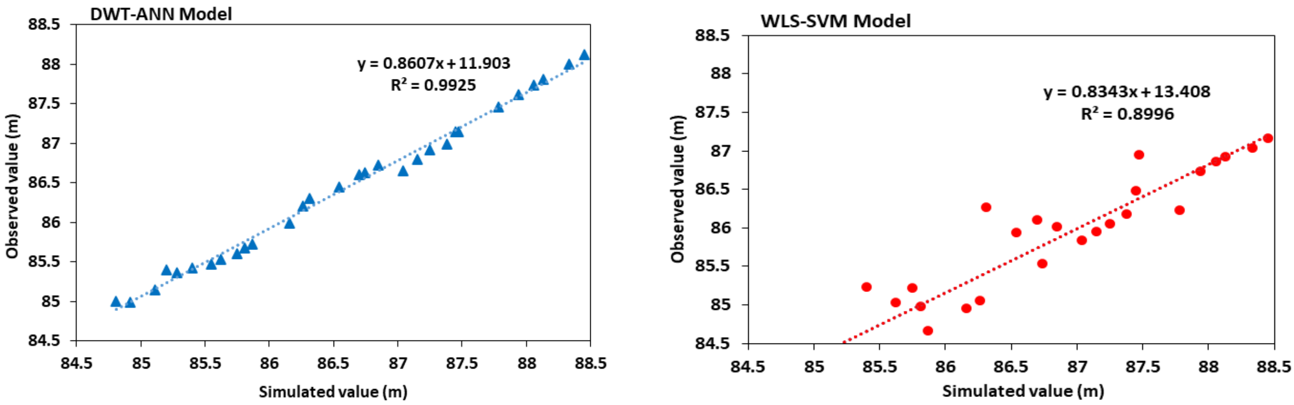
(b) Dasht-Bagheyn well



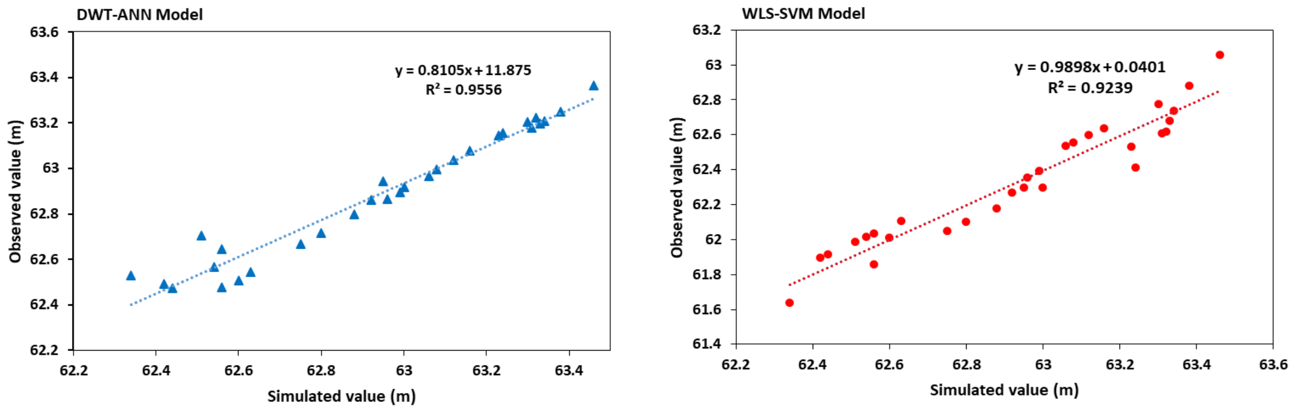
(c) Tolombe-Badi wells

Fig. 7: Scatter plot of observed groundwater level vs. the DWT-ANN and WLS-SVM simulated groundwater level in training phase at (a) Saadi, (b) Dasht-Bagheyn and (c) Tolombe-Badi wells.

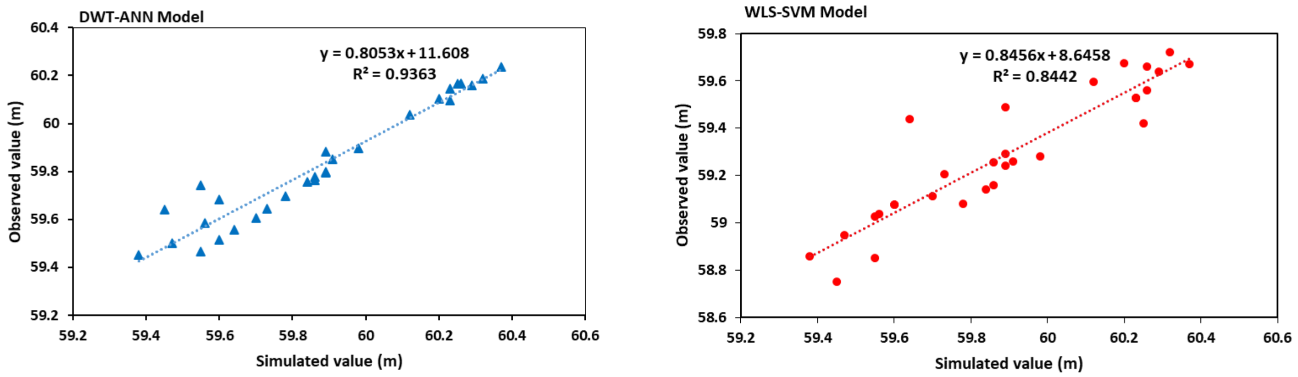
combining Wavelet with ANN significantly improve predictive performance for hydrological time series. Additionally, another relevant study by Samani et al. (2023) indicated that hybrid wavelet-machine learning conjunction models were superior to other stand-alone model to predict groundwater level. Similar to previous studies (Wei et al., 2023), statistical indicators such as mean absolute error (MAE), root mean squared error (RMSE), and Nash-Sutcliffe efficiency (NSE) were calculated to evaluate the performance of the hybrid model. After considering the three statistical indicators (MAE, RMSE and NSE) and conducting visual analysis, it became clear that the DWT-ANN model demonstrated improved performance. The RMSE values for the DWT-ANN model varied from 0.010 m to 0.075 m in the training phase and from 0.168 m to 0.257 m in the testing phase. The MAE results also followed the RMSE indicator, and NSE values ranging from 0.9954 to 0.9998 in the training phase and from 0.8994 to 0.9330 in the testing phase. These results establish the Wavelet-ANN combination as having the highest simulation accuracy. These findings are consistent with those reported by Wei et al. and Shahbazi et al. (2023), which revealed that the novel hybrid model combining wavelet with ANN is more accurate compared to other stand-alone models.



(a) Saadi well



(b) Dasht-Bagheyn well



(c) Tolombe-Badi wells

Fig. 8: Scatter plot of observed groundwater level vs. the DWT-ANN and WLS-SVM simulated groundwater level in testing phase at (a) Saadi, (b) Dasht-Bagheyn and (c) Tolombe-Badi wells.

According to the authors' literature review, the findings of this study were more accurate compared to those of previous studies. For instance, Wei et al. (2023) reported that the wavelet-ANN model for groundwater forecasting using the best Daubechies mother wavelet function, achieved an optimal RMSE of 0.562m and NSE of 0.990. In another related study, Shahbazi et al. (2023) reported a coefficient of determination (R^2) of 0.984 and 0.938 in the training and testing phases for groundwater modeling using wavelet-ANN, respectively. In the present study, the R^2 value for the DWT-ANN model ranged from 0.993 to 0.9976 in the training phase and from 0.9363 to 0.9925 in the testing phase. It is clear that the maximum R^2 values (0.9976 in training and 0.9925 in testing phases) in the present study are higher than the indicator reported in the study by Shahbazi et al. (2023). Therefore, it can be concluded that the present hybrid model of DWT-ANN is a promising and robust tool for simulating groundwater levels.

Conclusion

Comparing the statistical indicators of both the DWT-ANN and WLS-SVM models in the training and testing phases indicated that the DWT-ANN with Db4 mother wavelet was a more powerful tool for simulating groundwater levels than the WLS-SVM model. The performance evaluation results of the DWT-ANN and WLS-SVM models indicate that the DWT-ANN model performs more efficiently than the WLS-SVM method. The discrete wavelet transform (DWT) can remove noise, making the hybrid model of DWT and ANN more accurate than the WLS-SVM method. The results indicate that the DWT-ANN method could be an accurate and reliable simulation tool for groundwater levels. It can also handle the planning of water resources management strategies in arid area. Since fieldwork for groundwater monitoring projects is costly and time consuming, the hybrid simulation approach to groundwater levels proposed in this research can help reduce costs and provide accurate groundwater level simulations. It is noteworthy that the proposed hybrid method can simulate groundwater level with high accuracy. Therefore, the hybrid simulation approach for groundwater levels proposed in the research can enhance groundwater monitoring projects.

Author's Contribution: A. Robati conceived and performed the analysis. Sh. Khosravi collected the data. Sh. Khosravi and A. Robati wrote and revised the manuscript. Both authors approved final version of the manuscript.

REFERENCES

Adamowski, J., & Chan, H.F. (2011). A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology*, 407(1-4), 28-40. https://doi.org/10.1016/j.jhydrol.2011.06.013

Adamowski, J., & Sun, K. (2010). Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *Journal of Hydrology*, 390(1-2), 85-91. https://doi.org/10.1016/j.jhydrol.2010.06.033

Anh, D.T., Pandey, M., Mishra, V.N., Singh, K.K., Ahmadi, K., Janizadeh, S., Tran, T.T., Linh, N.T.T., & Dang, N.M. (2023). Assessment of groundwater potential modeling using support vector machine optimization based on Bayesian multi-objective hyperparameter algorithm. *Applied Soft Computing*, 132, 109848. https://doi.org/10.1016/j.asoc.2022.109848

Bahmani, R., Solgi, A., & Ouarda, T.B. (2020). Groundwater level simulation using gene expression programming and M5 model tree combined with wavelet transform. *Hydrological Sciences Journal*, 65(8), 1430-1442. https://doi.org/10.1080/02626667.2020.1749762

Bahmani, R., & Ouarda, T.B. (2021). Groundwater level modeling with hybrid artificial intelligence techniques. *Journal of Hydrology*, *595*, 125659. https://doi.org/10.1016/j.jhydrol.2020.125659

Boo, K.B.W., El-Shafie, A., Othman, F., Khan, M.M.H., Birima, A.H., & Ahmed, A.N. (2024). Groundwater level forecasting with machine learning models: A review. *Water Research*, 121249. https://doi.org/10.1016/j. watres.2024.121249

Cui, F., Al-Sudani, Z.A., Hassan, G.S., Afan, H.A., Ahammed, S.J., & Yaseen, Z.M. (2022). Boosted artificial intelligence model using improved alpha-guided grey wolf optimizer for groundwater level prediction: Comparative study and insight for federated learning technology. *Journal of Hydrology*, 606, 127384. https://doi.org/10.1016/j.jhydrol. 2021.127384

David, H.A. (1998). Early sample measures of variability. *Statistical Science*, 12, 368-377.

Deka, P.C. (2014). Support vector machine applications in the field of hydrology: a review. *Applied Soft Computing*, 19, 372-386. https://doi.org/10.1016/j.asoc.2014.02.002

Derbela, M., & Nouiri, I. (2020). Intelligent approach to predict future groundwater level based on artificial neural networks (ANN). Euro-Mediterranean *Journal for Environmental Integration*, 5, 1-11. https://doi.org/10.1007/s41207-020-00185-9

Duc, L., & Sawada, Y. (2023). A signal-processing-based interpretation of the Nash–Sutcliffe efficiency. *Hydrology and Earth System Sciences*, 27(9), 1827-1839. https://doi.org/10.5194/hess-27-1827-2023

Ebrahimi, H., & Rajaee, T. (2017). Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine. *Global and Planetary Change*, 148, 181-191. https://doi.org/10.1016/j.gloplacha.2016.11.014

Freire, P.K.D.M.M., & Santos, C.A.G. (2020). Optimal level of wavelet decomposition for daily inflow forecasting. *Earth Science Informatics*, 13(4), 1163-1173. https://doi.org/10.1007/s12145-020-00496-z

Graf, R., Zhu, S., & Sivakumar, B. (2019). Forecasting river water temperature time series using a wavelet–neural network hybrid modelling approach. *Journal of Hydrology*, 578, 124115. https://doi.org/10.1016/j.jhydrol. 2019.124115

Gordu, F., & Nachabe, M.H. (2021). A physically constrained wavelet-aided statistical model for multi-decadal groundwater dynamics predictions. *Hydrological Processes*, 35(8), e14308. https://doi.org/10.1002/hyp. 14308

Hajizadeh, A.R., Salajegheh, J., & Salajegheh, E. (2016). Performance evaluation of wavelet and curvelet transforms based-damage detection of defect types in plate structures. *Structural Engineering and Mechanics*, 60(4), 667-691. https://doi.org/10.12989/sem.2016.60.4.667

Hlaing, P.T., Humphries, U.W., & Waqas, M. (2024). Hydrological Model Parameter Regionalization: Runoff Estimation Using Machine Learning Techniques in the Tha Chin River Basin, Thailand. *MethodsX*, 102792. https://doi.org/10.1016/j.mex.2024.102792

Hsu, K.C., & Li, S.T. (2010). Clustering spatial–temporal precipitation data using wavelet transform and self-organizing map neural network. *Advances in Water Resources*, 33(2), 190-200. https://doi.org/10.1016/j. advwatres.2009.11.005

Karthika, B.S., & Deka, P.C. (2015). Prediction of air temperature by hybridized model (Wavelet-ANFIS) using wavelet decomposed data. *Aquatic Procedia*, 4, 1155-1161. https://doi.org/10.1016/j.aqpro.2015.02.147

Kayhomayoon, Z., Ghordoyee, M.S., Arya, A.N., & Kardan, M.H. (2021). A new approach for regional groundwater level simulation: clustering, simulation, and optimization. *Natural Resources Research*, *30*, 4165-4185. https://doi.org/10.1007/s11053-021-09913-6

Lee, S., Lee, K.K., & Yoon, H. (2019). Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors. *Hydrogeology Journal*, 27(2). https://doi.org/10.1007/s10040-018-1866-3

Liu, J., Chang, J.X., & Zhang, W.G. (2009). Groundwater level dynamic prediction based on chaos optimization and support vector machine. In 2009 Third International Conference on genetic and evolutionary computing (pp. 39-43). IEEE.

Lohani, A.K., & Krishan, G. (2015). Groundwater level simulation using artificial neural network in southeast Punjab, India. *Journal of Geology and Geosciences*, 4(3), 206. http://dx.doi.org/10.4172/2329-6755.1000206

Mallat, S. (1999). A wavelet tour of signal processing. 3rd Ed., Academic Press.

Mallikarjuna, B., Sathish, K., Venkata, K.P., & Viswanathan, R. (2021). The effective SVM-based binary prediction of ground water table. *Evolutionary Intelligence*, 14(2), 779-787. https://doi.org/10.1007/ s12065-020-00447-z

Moosavi, V., Vafakhah, M., Shirmohammadi, B., & Behnia, N. (2013). A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. *Water Resources Management*, 27, 1301-1321. https://doi.org/10.1007/s11269-012-0239-2

Moravej, M., Amani, P., & Hosseini-Moghari, S.M. (2020). Groundwater level simulation and forecasting using interior search algorithm-least square support vector regression (ISA-LSSVR). *Groundwater for Sustainable Development*, 11, 100447. https://doi.org/10.1016/j.gsd.2020.100447

Nayak, P.C., Rao, Y.S., & Sudheer, K.P. (2006). Groundwater level forecasting in a shallow aquifer using artificial neural network approach. *Water Resources Management*, 20, 77-90. https://doi.org/10.1007/s11269-006-4007-z

Najafabadipour, A., Kamali, G., & Nezamabadi-Pour, H. (2022). Application of artificial intelligence techniques for the determination of groundwater level using spatio–temporal parameters. *ACS Omega*, *7*(12), 10751-10764. https://doi.org/10.1021/acsomega.2c00536

Nourani, V., Baghanam, A.H., Adamowski, J., & Kisi, O. (2014). Applications of hybrid wavelet–artificial intelligence models in hydrology: a review. *Journal of Hydrology*, 514, 358-377. http://dx.doi.org/10.1016/j.jhydrol. 2014.03.057

Nury, A.H., Hasan, K., & Alam, M.J.B. (2017). Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *Journal of King Saud University-Science*, 29(1), 47-61. https://doi.org/10.1016/j.jksus.2015.12.002

Osman, A.I.A., Ahmed, A.N., Chow, M.F., Huang, Y.F., & El-Shafie, A. (2021). Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engineering Journal*, *12*(2), 1545-1556. https://doi.org/10.1016/j.asej.2020.11.011

Rahman, A.S., Hosono, T., Quilty, J.M., Das, J., & Basak, A. (2020). Multiscale groundwater level forecasting: Coupling new machine learning approaches with wavelet transforms. *Advances in Water Resources*, 141, 103595. https://doi.org/10.1016/j.advwatres.2020.103595

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

Samani, S., Vadiati, M., Nejatijahromi, Z., Etebari, B., & Kisi, O. (2023). Groundwater level response identification by hybrid wavelet–machine learning conjunction models using meteorological data. *Environmental Science and Pollution Research*, 30(9), 22863-22884. https://doi.org/10. 1007/s11356-022-23686-2

Saroughi, M., Mirzania, E., Achite, M., Katipoğlu, O.M., Al-Ansari, N., Vishwakarma, D.K., & Yadav, K.K. (2024). Evaluate effect of 126 pre-processing methods on various artificial intelligence models accuracy versus normal mode to predict groundwater level (case study: Hamedan-Bahar Plain, Iran). *Heliyon*, 10(7). https://doi.org/10.1016/j.heliyon.2024.e29006

Satapathy, D.P., & Sahoo, S.K. (2023). Prediction of Ground Water Level using SVM-WOA Approach: A Case Study. *Journal of Scientific & Industrial Research*, 82(02), 269-277. https://doi.org/10.56042/jsir.v82i2.70212

Shahbazi, M., Zarei, H., & Solgi, A. (2023). De-noising groundwater level modeling using data decomposition techniques in combination with artificial intelligence (case study Aspas aquifer). *Applied Water Science*, 13(4), 88. https://doi.org/10.1007/s13201-023-01885-7

Shiri, N., Shiri, J., Yaseen, Z.M., Kim, S., Chung, I.M., Nourani, V., & Zounemat-Kermani, M. (2021). Development of artificial intelligence models for well groundwater quality simulation: Different modeling scenarios. *Plos one*, *16*(5), e0251510. https://doi.org/10.1371/journal.pone.0251510

Shu, C., & Ouarda, T.B. (2008). Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. *Journal of Hydrology*, 349(1-2), 31-43. https://doi.org/10.1016/j.jhydrol.2007.10. 050

Suykens, J.A., De Brabanter, J., Lukas, L., & Vandewalle, J. (2002). Weighted least squares support vector machines: robustness and sparse approximation. *Neurocomputing*, 48(1-4), 85-105. https://doi.org/10.1016/S0925-2312(01)00644-0

Tang, Y., Zang, C., Wei, Y., & Jiang, M. (2019). Data-driven modeling of groundwater level with least-square support vector machine and spatial–temporal analysis. *Geotechnical and Geological Engineering*, 37(3), 1661-1670. https://doi.org/10.1007/s10706-018-0713-6

Tao, H., Hameed, M.M., Marhoon, H.A., Zounemat-Kermani, M., Heddam, S., Kim, S., & Yaseen, Z.M. (2022). Groundwater level prediction using machine learning models: A comprehensive review. *Neurocomputing*, 489, 271-308. https://doi.org/10.1016/j.neucom.2022.03.014

Vapnik, V. (2013). The nature of statistical learning theory. Springer science & business media.

Vousoughi, F.D. (2023). Wavelet-based de-noising in groundwater quality and quantity prediction by an artificial neural network. *Water Supply*, 23(3), 1333-1348. doi: 10.2166/ws.2023.021

Wang, D., Li, P., He, X., & He, S. (2023). Exploring the response of shallow groundwater to precipitation in the northern piedmont of the Qinling Mountains, China. *Urban Climate*, 47, 101379. https://doi.org/10.1016/j .uclim.2022.101379

Wei, A., Li, X., Yan, L., Wang, Z., & Yu, X. (2023). Machine learning models combined with wavelet transform and phase space reconstruction for groundwater level forecasting. *Computers & Geosciences*, 177, 105386. https://doi.org/10.1016/j.cageo.2023.105386

Wu, C., Zhang, X., Wang, W., Lu, C., Zhang, Y., Qin, W., Tick, R.G., Liu, B., & Shu, L. (2021). Groundwater level modeling framework by combining the wavelet transform with a long short-term memory data-driven model. *Science of the Total Environment*, 783, 146948. https://doi.org/10.1016/j .scitotenv.2021.146948

Wunsch, A., Liesch, T., & Broda, S. (2021). Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). *Hydrology and Earth System Sciences*, 25(3), 1671-1687. https://doi.org/10.5194/hess-25-1671-2021

Yadav, B., Gupta, P.K., Patidar, N., & Himanshu, S.K. (2020). Ensemble modelling framework for groundwater level prediction in urban areas of India. *Science of the Total Environment*, 712, 135539. https://doi.org/10.1016/j.scitotenv.2019.135539

Yoon, H., Hyun, Y., Ha, K., Lee, K.K., & Kim, G.B. (2016). A method to improve the stability and accuracy of ANN-and SVM-based time series models for long-term groundwater level predictions. *Computers & Geosciences*, 90, 144-155. https://doi.org/10.1016/j.cageo.2016.03.002

Zafar, M.M., Chattha, W.S., Khan, A.I., Zafar, S., Subhan, M., Saleem, H., Ali, A., Ijaz, A., Anwar, Z., Qiao, F., Shakeel, A., Seleiman, M.F., Wasonga, D.O., Parvaiz, A., Razzaq, A., & Xuefei, J. (2023). Drought and heat stress on cotton genotypes suggested agro-physiological and biochemical features for climate resilience. *Frontiers in Plant Science*, *14*, 1265700. https://doi.org/10.3389/fpls.2023.1265700-

Zafar, M.M., Razzaq, A., Chattha, W.S., Ali, A., Parvaiz, A., Amin, J., Saleem, H., Shoukat, A., Elhindi, K.M., Shakeel, A., Ercisli, S., Qiao, F., & Jiang, X. (2024). Investigation of salt tolerance in cotton germplasm by analyzing agro-physiological traits and ERF genes expression. *Scientific Reports*, *14*(1), 11809. https://doi.org/10.1038/s41598-024-60778-0

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. https://doi.org/10.1016/j.jher. 2017.11.004

Zhang, J., Zhang, X., Niu, J., Hu, BX., Soltanian, MR., Qiu, H., & Yang, L. (2019). Prediction of groundwater level in seashore reclaimed land using wavelet and artificial neural network-based hybrid model. *Journal of Hydrology*, 577, 123948. https://doi.org/10.1016/j.jhydrol.2019.123948

Zhang, Q., Li, P., Ren, X., Ning, J., Li, J., Liu, C., Wang, Y., & Wang, G. (2023). A new real-time groundwater level forecasting strategy: Coupling hybrid data-driven models with remote sensing data. *Journal of Hydrology*, 625, 129962. https://doi.org/10.1016/j.jhydrol.2023.129962

Zhou, T., Wang, F., & Yang, Z. (2017). Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction. *Water*, 9(10), 781. https://doi.org/10.3390/w9100781

Zhu, S., Ptak, M., Yaseen, Z.M., Dai, J., & Sivakumar, B. (2020). Forecasting surface water temperature in lakes: A comparison of approaches. *Journal of Hydrology*, *585*, 124809. https://doi.org/10.1016/j.jhydrol. 2020.124809