Larhyss Journal, ISSN 1112-3680, n°61, March 2025, pp. 343-369

COMPARATIVE STUDY OF MACHINE LEARNING MODELS IN PREDICTING WATER TABLE FLUCTUATIONS IN AZARSHAHR PLAIN, IRAN

FATEMI S.M., MOLAVI A.*

Department of Water Sciences and Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran

(*) ahad.molavi@gmail.com

Research Article – Available at http://larhyss.net/ojs/index.php/larhyss/index
Received July 13, 2024, Received in revised form February 18, 2025, Accepted February 21, 2025

ABSTRACT

Managing drought and water scarcity concerning underground water resources requires modeling methodologies with a simple yet effective framework. Due to temporal and financial constraints, machine learning models (MLMs) are crucial in this context. This study aims to predict daily underground water levels (UWL) in Azarshar Plain, Tabriz, Iran, using three MLMs (SVM, GEP, MLP). Covering 126 annual datasets for 34 wells from 2018 to 2021, various combinations were tested with different UWLs and lag times. Performance evaluation metrics including RMSE, MAE, R2, and DDR were employed. Results show satisfactory accuracy for all three MLMs, with SVM, GEP, and MLP being more accurate in 53%, 26%, and 20% of cases respectively among the 34 wells. The input configuration with a lag-time of two days (M2) emerged as the most optimal, yielding the most accurate simulations. Average values of RMSE, MAE, R2, and DDR for M2 during the testing period were calculated as 0.2457, 0.2077, 0.9482, and 31.53 respectively. In conclusion, these MLMs offer viable alternatives to numerical models for managing and predicting UWL, facilitating better water resource management in areas prone to drought and water scarcity.

Keywords: Simulation, Aquifer, Prediction, Performance Assessment, Groundwater.

INTRODUCTION

Groundwater sources are widely acknowledged as a paramount and invaluable reservoir of water resources in the realm of hydrogeological studies (Hountondji et al., 2020; Mehta et al., 2023; Srivastava et al., 2023; Deb, 2024). Groundwater resources are highly vulnerable to over-extraction, contamination especially by nitrate (Koussa and Berhail, 2021), and climate change, making their protection essential for ensuring long-term water

security and ecological balance (Kouassi et al., 2013; Ouhamdouch et al., 2016; Aroua, 2018; Zegait et al., 2021; Nakou et al, 2023). Unregulated withdrawal and pollution from industrial, agricultural, and domestic sources threaten groundwater quality, necessitating stringent conservation measures and sustainable management practices (Belhadj et al., 2017; Laghzal et al., 2019). Promoting natural groundwater recharge through watershed protection, afforestation, using flood waters, and wetland restoration helps maintain aquifer levels and prevents depletion in arid and semi-arid regions (Chibane and Ali-Rahmani, 2015; El Moukhayar et al., 2015; Benmmoussat et al., 2017; Remini, 2018; Qureshi et al., 2024). Artificial recharge techniques, such as managed aquifer recharge (MAR), infiltration basins, and other sources of supply, are critical solutions for replenishing groundwater reserves, mitigating drought impacts, and sustaining water availability for future generations (Gaaloul, 2015; Abaidia and Remini, 2020; Later and Labadi, 2024).

Climate fluctuations significantly affect both the quantity and quality of groundwater, posing challenges for sustainable water resource management (Zella and Smadhi, 2010; Ouis, 2012; El Fellah-Idrissi et al., 2017). Variations in precipitation patterns and prolonged droughts reduce groundwater recharge, leading to declining water tables and increased reliance on deep aquifers. Rising temperatures intensify evapotranspiration, further exacerbating groundwater depletion in arid and semi-arid regions. Additionally, fluctuations in rainfall influence contaminant transport, as extreme precipitation can lead to increased infiltration of pollutants, while reduced recharge can concentrate dissolved salts and pollutants, deteriorating water quality. The combined effects of drought, excessive pumping, and contamination risks highlight the urgent need for adaptive groundwater management strategies in response to climate variability (Bahir et al., 2015).

Aquifer systems must be thoroughly characterized with a high degree of precision to accurately determine the volume of available water resources. Indeed, understanding the volume of underground water resources is crucial for sustainable water management, ensuring the balance between extraction and natural recharge to prevent overexploitation and depletion (Baiche et al., 2015; Argaz, 2018). In addition, accurate assessment of aquifer reserves enables informed decision-making for water supply planning, agricultural irrigation, and industrial use, particularly in arid and semi-arid regions where groundwater is a primary resource. Moreover, precise knowledge of underground water volumes helps in mitigating the impacts of droughts and climate change, allowing for the development of adaptive strategies to secure long-term water availability. Furthermore, it is essential for controlling land subsidence risks, maintaining ecological stability, and ensuring the resilience of communities dependent on groundwater resources. Geological and hydrogeological investigations serve as essential tools for accurately evaluating aquifer resources (Meroni et al., 2021).

It is equally imperative to monitor groundwater quality to gain a comprehensive understanding of the intricate relationship between geochemical processes and water composition (Ngouala et al., 2016). This knowledge is essential for the implementation of effective and sustainable groundwater management strategies, ensuring the long-term preservation and responsible utilization of this critical resource.

The accurate identification and fundamental utilization of these resources, particularly in arid and semi-arid regions, exert a substantial influence on the sustainable advancement of numerous agricultural, societal, and economic endeavors (Umrigar et al., 2024; Alshehri and Rahman, 2023). The diminution in precipitation and the occurrence of drought in recent years have posed a formidable challenge to the replenishment cycle of subterranean aquifers, leading to a significant lowering of water levels while affecting the aquifers' recharge (Haouchine et al.; 2015; Nichane and Khelil, 2015). Concomitant with the escalating human expectations and the expansion of the agricultural sector, the decline in rainfall has been paralleled by a heightened extraction of subterranean water resources (Ahmed et al. 2023; Li et al., 2019). The aforementioned factors have ultimately culminated in a depletion of subterranean water resources and a reduction in the UWLs (Rajput et al., 2023). Numerous models have been employed for the prediction of UWLs, including ancestral methods (Kebizi et al., 2023). In contemporary hydrogeological practices, the development of a streamlined aquifer model and its subsequent simulation are deemed imperative. This is essential not only for the implementation of management scenarios but also for the selection of viable strategies to effectively administer groundwater resources, along with the determination of appropriate withdrawal rates from groundwater reservoirs (Dadhich et al., 2021).

The categorization of models employed for simulating groundwater levels predominantly falls into two principal groups: physical models and data-driven models (Bahmani and Quarda, 2021). For physical models, achieving a suitable synthesis of aquifer parameters is imperative to delineate the spatial variability of groundwater within the aquifer. Acquiring this information poses challenges, as it necessitates costly on-site surveys, consequently augmenting computational expenses. An alternative approach to physical modeling is offered by data-driven models, which facilitate accurate predictions without the need for an extensive and costly calibration process, and wherein the underlying physical mechanisms are not explicitly considered during model development (Sattari et al., 2018). Data-driven models, exemplified by soft computing and machine learning tools, present a viable means to surmount the constraints associated with physical models (Kantharia et al. 2024; Yoon et al., 2011).

The optimization of agricultural practices and the assessment of the potential for high-quality groundwater resources hold paramount significance. While physical and mathematical models serve as fundamental and precise instruments for ascertaining hydrogeological parameters and comprehending system processes, they are encumbered by cost and time constraints. Additionally, they necessitate comprehensive and diverse input data, resulting in a substantial computational workload and protracted model development and execution timelines. In recent years, in order to overcome the aforementioned limitations, MLMs have emerged and evolved as capable tools for simulating system behavior without the requirement for extensive domain expertise (Kaushik and Kumar, 2023; Leon et al., 2023; Abbaszadeh Shahri et al., 2022). A concise overview of the MLMs that have been implemented is provided in Table 1.

Table 1: Literature review of MLMs on UWL prediction

Reference	Model name	Results
Yao et al. (2024)	CEEMDAN–BiGRU–SVR–MWOA (CBSM) framework	The CBSM as a new and effective method is proposed for accurately simulating and predicting lake water levels.
Yi et al. (2024)	Random Forest (RF), ANN, SVM, Gradient Boosting, and Extreme Gradient Boosting (XGBoost)	The XGBoost outperformed other models.
Mohammed et al. (2023)	Genetic Algorithm-ANN (GA-ANN), Independent Component Analysis-ANN (ICA-ANN), extreme learning machine (ELM), outlier robust extreme learning machine (ORELM)	The results demonstrate that the closest point to the reference point is related to the ORELM method
	Long short-term memory (LSTM), gated recurrent unit (GRU), bidirectional LSTM (BiLSTM)	Deep learning models performed well for annual-type of ground water level
	SVM, Generalized Regression Neural Network (GRNN), Decision Tree (DT), Random Forest (RF), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Gated Recurrent (GRN) Network	The RF had the superior outcomes
	ANFIS, Improved Alpha-Guided Grey Wolf optimization (IA-GWO)	ANFIS-IA-GWO model had better results
Mozaffari et al. (2022)	PSO (Particle Swarm Optimization), SVR	The SVR-PSO hybrid model can be used as a superior tool for simulation.
Vadiati et al. (2022)	ANN, fuzzy logic (FL), ANFIS, SVM	the ANFIS model showed slightly better performance
Vu et al. (2021)	LSTM	The LSTM is reliable to predict missing water level data.
Seifi et al. (2020)	ANN, ANFIS, SVM, grasshopper optimization algorithm (GOA), cat swarm optimization (CSO), weed algorithm (WA), GA, krill algorithm (KA), PSO	ANFIS-GOA and SVM had, respectively, the best and worst performances among other models
Jeihouni et al. (2019)	ANN, ANFIS	the ANN was the best model.
Afzaal et al. (2019)	Multilayer perceptron (MLP), LSTM, convolutional neural network (CNN)	The MLP performed better
Das et sl. (2019)	Back Propagation Neural Network (BPNN), ANFIS	ANFIS had better performance than BPNN

Ebrahimi et al. (2017)	ANN, multi linear regression (MLR), SVR, wavelet-ANN (WNN), wavelet-MLR (WLR), wavelet-SVR (WSVR)	The WNN had accurate results.

Simulation of UWLs using MLMs reveals high accuracy within the study area, underscoring MLMs' potential. Notably, MLMs suit scenarios with limited data diversity, crucial when instrumentation lacks, necessitating accurate predictive modeling with minimal datasets. The Azarshahr Plain in East Azerbaijan Province, Iran, plays a crucial role socio-economically, particularly due to its groundwater resources serving as a primary water source for agriculture, industry, and urban needs. However, there are notable challenges in accurately simulating UWLs using MLMs in this area, possibly stemming from various factors. Therefore, intensified research efforts and contemporary methodologies are necessary for UWLs modeling in the Azarshahr Plain. The primary objective of using machine learning models (SVM, GEP, and MLP) for predicting underground water levels (UWL) in the Azarshahr Plain is to develop accurate daily UWL predictions. This is critical for managing water resources, especially in regions facing drought and water scarcity, by providing efficient alternatives to numerical models for better decision-making in groundwater management

The novel aspects of this research include applying multiple machine learning models to UWL prediction and evaluating their performance using comprehensive metrics across different lag-time configurations. The focus on daily predictions in a drought-prone area adds practical significance to the study

MATERIAL AND METHODS

Case study

Azarshahr plain, situated in the northwestern region of Iran, has experienced a conspicuous decline in UWL in recent years. Given the critical significance of this matter and the geographical proximity of Azarshahr plain to the Lake Urmia watershed, it becomes imperative to undertake a comprehensive investigation into the fluctuations of UWL within this locale. In the investigated region, 34 monitoring wells were utilized to assess the UWL, and their data spanning from September 23, 2018, to September 23, 2021, were employed for the purposes of this study. The dataset comprises 126 daily observations for each well over the three-year period. The ensuing table presents the geographic attributes, along with pertinent statistical features, pertaining to the water levels observed in each well throughout the specified temporal duration. Figure 1 illustrates the temporal variation of the measured UWL at the beginning and end of the study period. Within the figure, the positions of the observation wells are denoted by the "+" sign.

Table 2: A summary of wells statistics at the Azarshar plane, Tabriz

Well NO.	X(utm)	Y(utm)	Max	Min	STDEV	VAR	Mean	CV
1	583882	4183425	28.670	23.880	0.970	0.941	27.872	0.035
2	579493	4184475	22.030	12.540	2.119	4.491	17.504	0.121
3	579640	4185452	15.650	9.430	1.191	1.418	13.109	0.091
4	586366	4181777	29.220	25.570	0.821	0.674	28.012	0.029
5	577574	4192337	2.940	0.780	0.518	0.268	2.079	0.249
6	580250	4193150	16.450	7.100	2.116	4.476	12.506	0.169
7	576388	4183773	3.990	0.920	0.530	0.281	2.091	0.253
8	575019	4184860	4.230	0.850	0.670	0.449	3.133	0.214
9	577379	4183616	5.000	0.560	0.875	0.765	2.607	0.336
10	577363	4183608	11.600	6.800	0.879	0.773	9.760	0.090
11	586933	4182783	12.370	8.750	0.805	0.648	10.293	0.078
12	586218	4177195	28.000	14.670	2.562	6.565	25.241	0.102
13	569717	4191646	11.650	5.130	1.537	2.362	8.124	0.189
14	580299	4179724	3.250	0.970	0.391	0.153	2.141	0.183
15	574438	4186002	16.640	5.020	2.151	4.628	11.264	0.191
16	580516	4187705	4.020	1.720	0.517	0.267	2.927	0.177
17	583350	4179279	7.880	2.350	0.774	0.600	6.042	0.128
18	582118	4184209	7.790	2.710	1.195	1.428	5.841	0.205
19	581126	4183380	21.600	14.610	1.517	2.301	18.334	0.083
20	582112	4184211	11.500	6.380	1.267	1.605	8.768	0.144
21	580506	4186885	24.150	16.770	1.715	2.941	20.988	0.082
22	577202	4179979	7.480	4.200	0.632	0.400	6.046	0.105
23	578045	4180332	22.870	7.660	2.609	6.808	14.838	0.176
24	576729	4178811	7.350	2.850	0.974	0.948	6.010	0.162
25	576300	4180122	8.450	5.870	0.656	0.430	7.499	0.087
26	574844	4177649	5.860	3.250	0.486	0.236	5.183	0.094
27	579864	4173986	8.510	6.990	0.339	0.115	7.839	0.043
28	584480	4183477	29.780	1.340	10.520	110.669	14.984	0.702
29	576166	4176016	13.940	12.310	0.464	0.216	13.277	0.035
30	578841	4182026	25.000	13.470	2.523	6.365	20.038	0.126
31	582650	4185785	9.600	6.500	0.526	0.277	7.639	0.069
32	584079	4185578	8.150	7.360	0.190	0.036	7.755	0.024
33	585470	4187716	21.690	14.000	2.433	5.919	15.865	0.153
34	582980	4190811	8.000	5.380	0.645	0.416	6.613	0.098

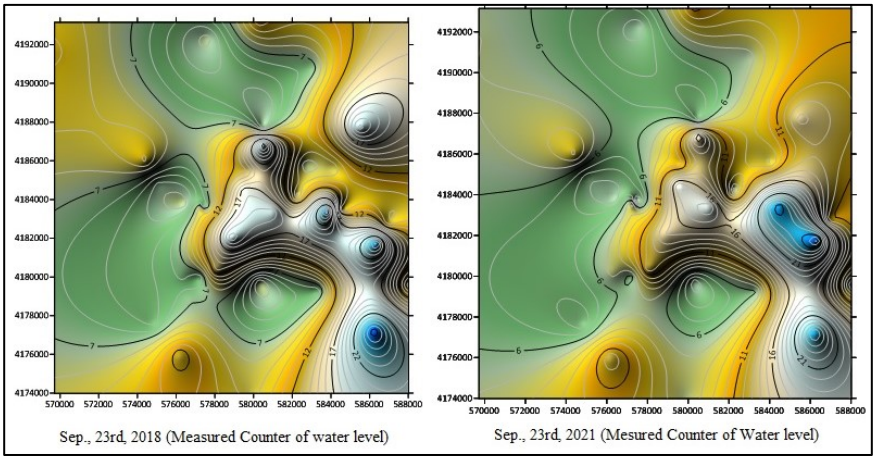

Figure 1: Measured water level counters at the beginning and end of the period

In this study, data preprocessing involved the following steps: (1) handling missing data: any missing data points were filled using interpolation methods to maintain the continuity of the time series; (2) outliers: outliers were detected through statistical methods and were either removed or replaced using mean/median values to prevent them from skewing the predictions.

An overview on SVM

The inception of the SVM algorithm can be attributed to Vapnik (1995), representing a seminal contribution to the field of machine learning. Concurrently, Cortes and Vapnik (1995) introduced a significant refinement in the form of the prevailing standard rendition of SVM, known as the "soft margin." This modification has since played a pivotal role in enhancing the algorithm's practical utility. The SVM constitutes a category of supervised learning models endowed with dedicated learning algorithms. These algorithms are adept at scrutinizing data, discerning underlying patterns, and are amenable for employment in tasks pertaining to classification and regression analysis. It is pertinent to note that the inherent challenge within the SVM arises when attempting to delineate sets that exhibit a lack of linear separability within the confines of a finite-dimensional space, despite the problem's original formulation within such a space. To address the issue of non-linear separability, a proposed strategy entails the transformation of the initial finite-dimensional space into a considerably higher-dimensional space. This maneuver is hypothesized to facilitate a more tractable separation of sets within the newly constructed space. In order to maintain a manageable computational burden, the SVM approaches incorporate specific mappings designed to facilitate the straightforward computation of dot products in terms of original space parameters. These mappings are expressed in relation to a kernel function denoted as K(x,y), chosen to align with the problem at hand. In the context of a higher-dimensional space, hyper-planes are defined as the locus of points characterized by a constant dot product with a vector existing within that space.

The vectors defining these hyper-planes can be selected as linear combinations involving parameters αi, which correspond to transformed feature vectors present in the dataset. With this choice of hyper-plane formulation, the points within the feature space, subsequently mapped onto the hyper-plane, adhere to the following mathematical relationship:

$$\sum_{i=1} \alpha_i k(x_i,x)=\text{constant} \quad (1)$$

It is imperative to note that when the kernel function K(x, y) diminishes as the variable y departs from the point x, each term within the summation serves as a metric for assessing the proximity of the test point x to its corresponding data point xi. This formulation allows for the collective sum of kernels mentioned above to serve as a quantification of the relative proximity of each test point to the data points x originating from either of the sets that require discrimination. Various kernel types are elaborated upon in Table 3 for reference and application (Baudhanwala et al., 2024; Aderemi et al., 2023; Fuladipanah and Majediasl, 2021; Fuladipanah et al., 2021).

Table 3: Different Kernel functions

Function	Expression
Linear	$K(x_i,x_j)=(x_i,x_j)$
Polynomial	$K(x_i,x_j)=[(x_i,x_j)+1]^d$
Radial basis function	$K(x_i,x_j)=\exp[-\frac{\|x_i-x_j\|^2}{2\sigma^2}]$
Sigmoid	$K(x_i,x_j)=\tanh[-\alpha(x_i,x_j)+c]$

An overview on the GEP

The GEP is a computational technique used in evolutionary algorithms for symbolic regression, function optimization, and machine learning. The GEP focuses on evolving and discovering mathematical expressions and symbolic models, making it particularly useful for problems where the functional form of the solution is unknown or complex. Chromosome structure, function sets and terminal sets, encoding, evolutionary process, fitness function can be counted as the key elements of the GEP structure. In the GEP, solutions are represented as linear strings of symbols called "chromosomes." These chromosomes consist of genes, each of which represents a component of the solution. Genes can be of fixed length or variable length, allowing for flexibility in modeling complex functions. The GEP employs a set of functions and a set of terminals. Functions represent mathematical operations (e.g. addition, subtraction, multiplication), while terminals represent constants or variables from the problem domain. The combination of functions and terminals defines the set of possible mathematical expressions. Genes in a GEP chromosome encode a mathematical expression or a symbolic model. The combination of genes in a chromosome forms an expression tree, where functions act as internal nodes and terminals act as leaf nodes. The GEP utilizes genetic operators such as mutation, recombination (crossover), and selection to evolve populations of

chromosomes over generations. During evolution, new expressions are generated by modifying or recombining existing genes or sub-expressions. The fitness function evaluates how well a chromosome's encoded expression approximates the desired behavior or solves the problem at hand. The goal of GEP is to evolve chromosomes with high fitness values (Ferreira, 2001; Azamathulla et al., 2018; Birbal et al., 2021; Majedi-Asl et al., 2022; Balahang and Ghodsian, 2023; Leon et al., 2023; Fuladipanah et al., 2023). Figure 2 presents a general step of the GEP model.

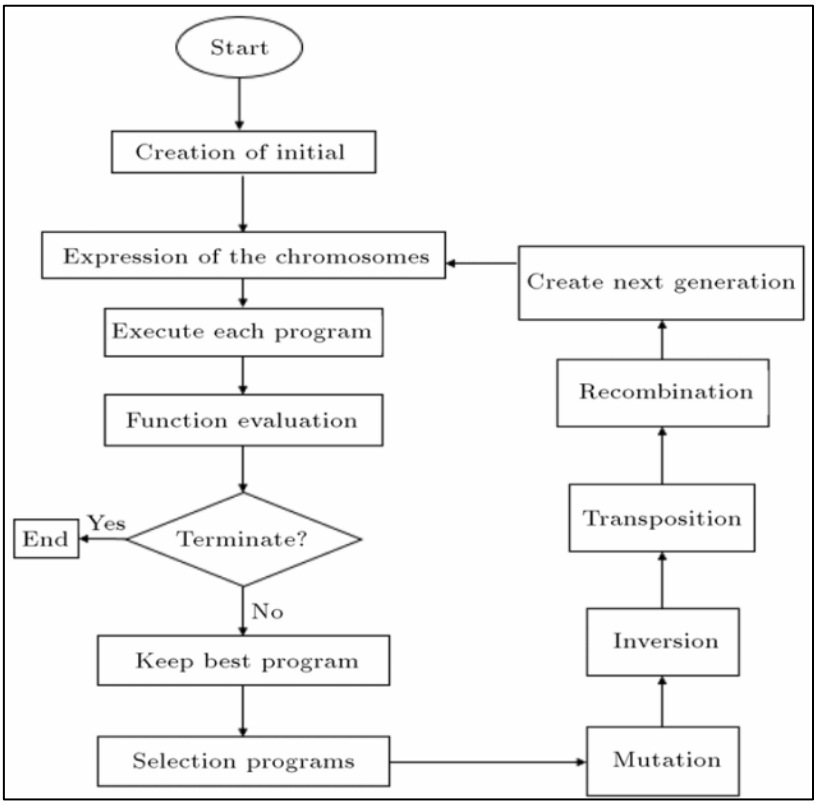

Figure 2: Flowchart of the GEP model

Overview on the MLP

The MLP is a class of artificial neural networks that have proven to be versatile and effective in various machine learning applications. It, a type of feed-forward neural network, has gained prominence in solving complex problems across diverse domains such as image recognition, natural language processing, and financial forecasting. The MLPs are characterized by their layered architecture, consisting of an input layer, one or more hidden layers, and an output layer. The architecture of an MLP is defined by the arrangement and connectivity of its layers. Each layer contains a set of nodes (neurons)

that process information and pass it to the next layer. The input layer receives the raw input data, while the output layer produces the final prediction or classification. Hidden layers, with their non-linear activation functions, enable the network to learn complex relationships in the data. The training of MLPs involves adjusting the weights and biases to minimize the difference between the predicted output and the true output. This process, known as back-propagation, utilizes optimization algorithms such as gradient descent to iteratively update the network parameters. The choice of activation functions, learning rates, and regularization techniques plays a crucial role in training the network effectively and preventing over-fitting. Activation functions introduce non-linearity into the network, enabling it to learn complex patterns. Common activation functions include sigmoid, hyperbolic tangent (tanh), and rectified linear unit (ReLU). The choice of activation function influences the network's capacity to capture and represent intricate relationships within the data (Azamathulla and Ahmad, 2013).

Performance criteria

The validation process for each developed model was conducted with the utilization of predetermined criteria, specifically the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). These criteria were defined by the following mathematical equations:

$$R^2=\frac{\sum_{i=1}^{N}(O_i-\bar{O})(P_i-\bar{P})}{\sqrt{\sum_{i=1}^{N}(O_i-\bar{O})^2}\sqrt{\sqrt{\sum_{i=1}^{N}(P_i-\bar{P})^2}}} \quad (2)$$

$$RMSE=\sqrt{\frac{\sum_{i=1}^{N}(O_i-P_i)^2}{N}} \quad (3)$$

$$MAE=\frac{\left|\sum_{i=1}^{N}(O_i-P_i)\right|}{N} \quad (4)$$

Where x_0 and x_P are observed and predicted values, respectively; $\bar{x}_0$ and $\bar{x}_P$ are the mean of the observed and the predicted values, respectively, and N is the total number of dataset. The indices mentioned earlier represent average error values of the implemented models. In order to address this limitation, a novel statistical metric, termed the developed discrepancy ratio (DDR) has been introduced by Noori et al. (2010):

$$DDR=\frac{\text{Predicted Value}}{\text{Observed Value}}-1 \quad (5)$$

For enhanced evaluation and visualization, it is advantageous to compute the Gaussian function of DDR values and depict them in a standard normal distribution format. To achieve this, the initial step involves standardizing the DDR values, followed by the application of the Gaussian function to obtain the normalized DDR values. The standard normal distribution for each model was graphically presented. In the resulting figures, it was observed that a stronger inclination of the error distribution graph toward the center

line, accompanied by larger values of the maximum when standardizing the DDR values, corresponds to a higher degree of accuracy.

RESULTS AND DISCUSSION

To anticipate the UWL in individual wells, the data presented in Table 4 were utilized for each of the machine learning models introduced in the preceding section. Within this table, five input variables encompass the UWL observed one to five days prior, denoted by Lag-1 to Lag-5 symbols, respectively. The different input configurations with lag times (M1-M5) were used to test the influence of previous UWL observations on current predictions. By incorporating lag times from one to five days to capture the temporal dependencies and enhance the accuracy of the models by testing various combinations of input variables.

Table 4: The input combination for the MLMs to predict the UWL

Model number	Input	Explanation
M1	WT(t-1)	Lag-1
M2	WT(t-1), WT(t-2)	Lag-2
M3	WT(t-1), WT(t-2), WT(t-3)	Lag-3
M4	WT(t-1), WT(t-2), WT(t-3), WT(t-4)	Lag-4
M5	WT(t-1), WT(t-2), WT(t-3), WT(t-4), WT(t-5)	Lag-5

Within this research work, a set of inputs, as delineated by the five models detailed in Table 4, for each of the MLMs, was employed for each of the 34 wells. A cumulative total of 510 simulations were conducted in the course of this research. A concise summary of the outcomes derived from these simulations is presented in Table 5. The initial column in Table 5 enumerates the well identifiers, while the second column delineates the optimal input combination for the respective **machine learning model.** The third column indicates the superior MLM. Columns 4 to 7 and 8 to 11, respectively, present the values of performance evaluation indicators during the training and testing phases for the two steps. The statistical data presented in this table reveals that, in the majority of wells, the input configuration denoted as M1 and M2 emerges as the most optimal combination for achieving accurate predictions of water levels. Notably, the M5 model did not emerge as the preferred combination of input parameters in any of the conducted simulations. Upon scrutinizing the numerical values of the performance evaluation indicators, it becomes evident that all three MLMs exhibit a notable potential for simulating the UWL. However, in comparative analysis, the SVM model consistently emerges with the highest frequency of superior models. In order to facilitate a more precise comparison between the observed and computed data, contour lines representing the water table during both the training and testing phases on the initial and final days are depicted in Figure 3. It is important to acknowledge that the choice of these specific days was dictated by the constraints associated with the presentation of numerous figures within the article. As discerned from

the performance evaluation indicators in Table 5 and Figure 3, a notably strong agreement between the measured and computed data is evident. The regulatory parameters of the GEP model are delineated in Table 6.

Table 5: The superior MLMs for wells involved with appropriate inputs

Well No.	Type of input	Model number	Training phase				Testing phases			
			RMSE	MAE	R^2	DDR_{max}	RMSE	MAE	R^2	DDR_{max}
1	M2	SVM	0.26989	0.02317	0.94821	41.94	0.1685	0.1382	0.9250	67.22
2	M2	SVM	0.42962	0.03790	0.97091	16.97	0.3580	0.3066	0.9752	21.69
3	M3	SVM	0.42962	0.03790	0.91579	13.19	0.3580	0.3066	0.8864	16.78
4	M3	GEP	0.31904	0.26342	0.87919	37.66	0.3155	0.2682	0.8929	36.49
5	M4	GEP	0.41035	0.34320	0.75301	2.41	0.2660	0.1943	0.7744	3.48
6	M2	MLP	0.25584	0.22012	0.98788	20.34	0.1937	0.1604	0.9951	29.99
7	M1	SVM	0.25336	0.02231	0.86339	2.94	0.2723	0.2304	0.8215	3.37
8	M1	MLP	0.23224	0.19893	0.88549	4.71	0.1675	0.1465	0.8762	8.07
9	M1	GEP	0.41035	0.34320	0.71252	2.63	0.1560	0.1354	0.9692	8.04
10	M1	SVM	0.15824	0.01383	0.97395	23.72	0.1929	0.1767	0.9511	25.56
11	M2	SVM	0.35588	0.03006	0.80562	12.40	0.2134	0.1851	0.8309	17.56
12	M2	SVM	0.23742	0.02001	0.99337	42.09	0.3032	0.2679	0.9580	37.37
13	M2	SVM	0.19414	0.01666	0.98167	16.72	0.3307	0.2591	0.9848	12.83
14	M2	MLP	0.23740	0.19493	0.84509	4.64	0.1214	0.1036	0.9452	7.32
15	M1	SVM	0.29423	0.02465	0.98917	15.59	0.1825	0.1520	0.9842	27.11
16	M3	MLP	0.23512	0.20128	0.83722	4.68	0.1758	0.1518	0.8528	7.80
17	M3	GEP	0.24472	0.20467	0.90197	10.23	0.2033	0.1665	0.9291	9.14
18	M3	SVM	0.29263	0.02640	0.94324	7.32	0.2653	0.2366	0.9400	9.41
19	M2	GEP	0.25295	0.22232	0.97633	28.70	0.2261	0.1925	0.9694	32.39
20	M1	GEP	0.28269	0.24073	0.90212	11.43	0.3250	0.2845	0.8312	13.75
21	M1	MLP	0.35263	0.30353	0.95946	22.86	0.2562	0.2201	0.9826	33.22
22	M1	SVM	0.22545	0.01926	0.91509	12.36	0.1831	0.1516	0.8919	12.59
23	M2	SVM	0.26064	0.02244	0.98884	21.17	0.2203	0.1859	0.9935	28.41
24	M4	SVM	0.26087	0.02297	0.94456	9.19	0.2503	0.2197	0.9034	9.07
25	M3	SVM	0.29410	0.02476	0.90429	13.67	0.2257	0.1888	0.9102	14.16
26	M1	MLP	0.24102	0.20151	0.84815	8.77	0.2440	0.2218	0.7249	8.89
27	M1	GEP	0.42856	0.37122	0.33765	7.55	0.3066	0.2542	0.3708	10.51
28	M2	GEP	0.28242	0.24037	0.99942	5.86	0.2588	0.2184	0.9981	30.25
29	M2	MLP	0.35531	0.29557	0.69195	17.32	0.2596	0.2161	0.4980	20.30
30	M2	GEP	0.23601	0.20440	0.99260	34.76	0.1685	0.1358	0.9427	19.01
31	M2	SVM	0.24083	0.02120	0.20387	12.68	0.1916	0.1704	0.2969	16.36
32	M2	SVM	0.21831	0.01904	0.38105	27.40	0.1796	0.1482	0.9914	50.85
33	M3	SVM	0.17079	0.01519	0.93049	15.52	0.1298	0.1142	0.8302	18.85
34	M1	SVM	0.23542	0.02050	0.92445	11.06	0.2053	0.1696	0.9466	11.24

Table 6: Values of tuned parameters for optimized MLMs

Well No.	Tuned parameters of SVM	Well No.	Tuned parameters of GEP	Well No.	Tuned parameters of MLP
1	(C=150, ε=0.10, γ=10)	4	Population size: 150 Number of genes: 3 Gene head length: 9 Gene tail length: 11	6	MLP: 2-5-1 Activation function: Tanh
2	(C=90, ε=0.10, γ=9)	5	Population size: 180 Number of genes: 3 Gene head length: 10 Gene tail length: 15	8	MLP: 1-3-1 Activation function: Tanh
3	(C=85, ε=0.10, γ=10)	9	Population size: 170 Number of genes: 3 Gene head length: 9 Gene tail length:19	14	MLP: 2-4-1 Activation function: Sigmoid
7	(C=200, ε=0.2, γ=10)	17	Population size: 320 Number of genes: 3 Gene head length: 15 Gene tail length:20	16	MLP: 3-7-1 Activation function: Tanh
10	(C=120, ε=0.10, γ=15)	19	Population size: 220 Number of genes:4 Gene head length: 15 Gene tail length: 18	21	MLP: 1-4-1] Activation function: Sigmoid
11	(C=120, ε=0.10, γ=23)	20	Population size:200 Number of genes: 4 Gene head length: 20 Gene tail length: 15	26	MLP: 1-5-1 Activation function: Sigmoid
12	(C=150, ε=0.50, γ=6)	27	Population size: 230 Number of genes: 3 Gene head length: 15 Gene tail length:24	29	MLP: 2-6-1 Activation function: Tanh
13	(C=200, ε=0.10, γ=10)	28	Population size: 240 Number of genes: 4 Gene head length: 16 Gene tail length: 35		
15	(C=150, ε=0.10, γ=2)	30	Population size: 350 Number of genes: 3 Gene head length: 21 Gene tail length: 25		
18	(C=200, ε=0.10, γ=10)	**Note:**			
22	(C=120, ε=0.20, γ=20)	The following tunes parameters are set as fixed values for all models			
23	(C=95, ε=0.10, γ=2)				
24	(C=80, ε=0.10, γ=25)	• Mutation rate: 0.044			
25	(C=150, ε=0.10, γ=9)	• Inversion rate: 0.1			
31	(C=100, ε=0.10, γ=18)	• Gene transposition rate: 0.1			
32	(C=100, ε=0.10, γ=19)	• One point recombination rate: 0.3			
33	(C=120, ε=0.10, γ=20)	• Two-point recombination rate: 0.3			
34	(C=110, ε=0.20, γ=26)	• Gene recombination rate: 0.1			
		• Fitness function: RMSE			

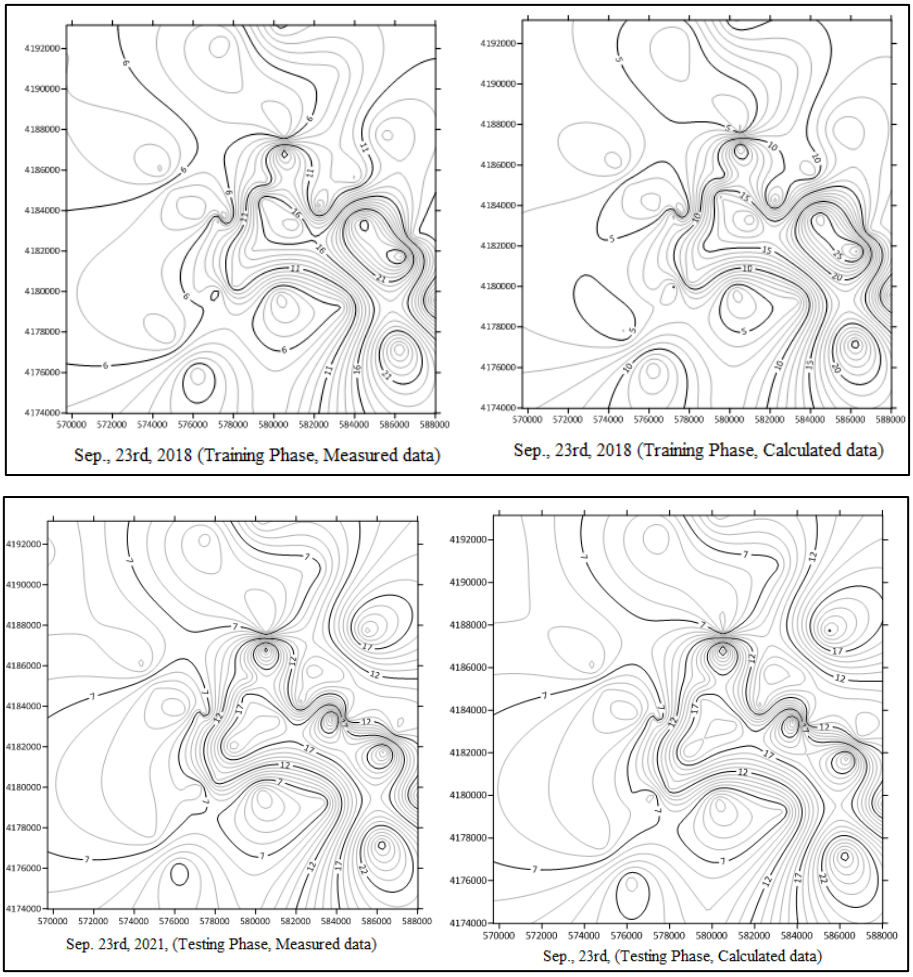

Figure 3: Water level counter line during training and testing phases on the first and last day of simulation

To elaborate further on the simulation process and the selection of the optimal model for each observation well, detailed descriptions of three chosen models are provided below, based on their performance. In the context of well number one, the SVM model has been selected as the optimal simulator, employing the M2 combination as the designated input parameter. The computed values for RMSE, MAE, R^2, and DDR_{max} indices during both the training and testing processes were determined as follows: Training Phase (RMSE=0.26989, MAE=0.02317, R^2=0.94821, DDR_{max}=46.64); Testing Phase (RMSE=0.1685, MAE=0.1382, R^2=0.9250, DDR_{max}=67.22). As established, the outcomes of the statistical indicators pertaining to performance evaluation during the test period surpass those of the training phase, indicative of the model's accurate training

performance. Notably, the tuned parameters for the SVM model in this simulation were set to C=150 and γ=2.5. In Fig. 4, the depiction illustrates the distribution of observational and computational data in relation to a line with a slope of 1:1 throughout both the training and test stages. Notably, the output of the SVM model exhibits a proximity to the 1:1 line in contrast to the other three models. The performance evaluation comparison of the three models, utilizing the DDR index, is depicted graphically in Fig. 5. As illustrated, the SVM model attains the highest value on the vertical axis during both the training and test stages, signifying its superiority. Additionally, Fig. 6 presents the degree of concordance between observational and calculated data, accompanied by the distribution of residual errors for each of the training and test stages.

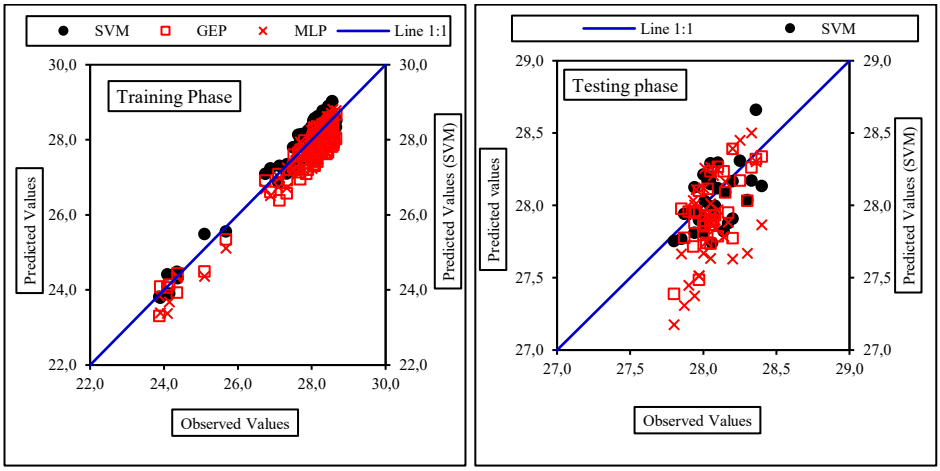

Figure 4: Scatter plot of SVM outcome for the Well #1

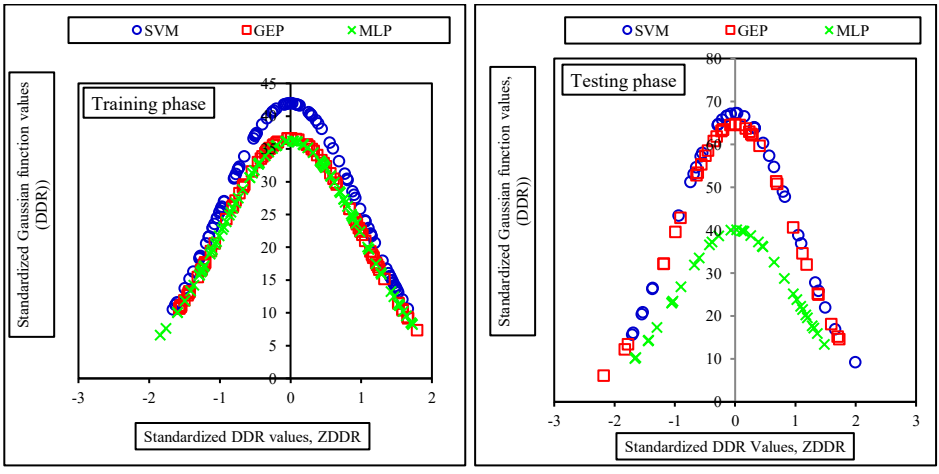

Figure 5: Distribution of DDR index for Well #1

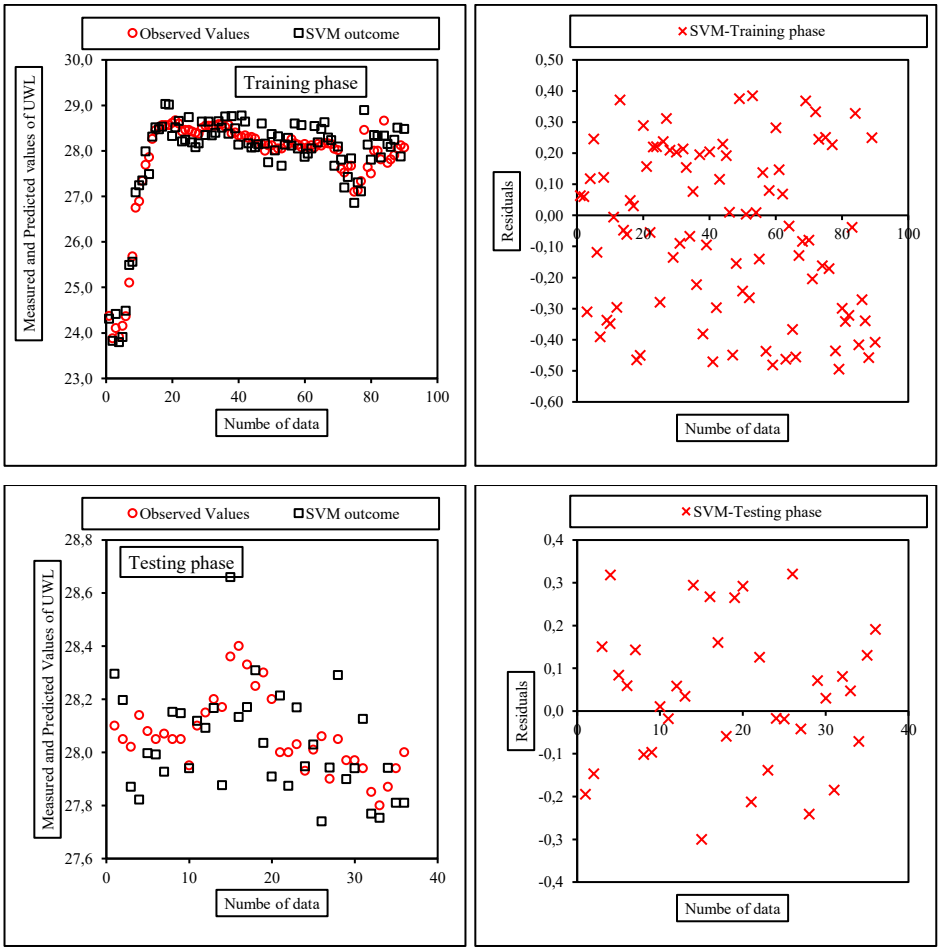

Figure 6: Performance of the SVM for well #1

For well #30, wherein the GEP model with the input combination M2 is identified as the superior and optimal model, the computed values for RMSE, MAE, R^2, and DDR_{max} indices during the training and validation stages are reported as follows: (0.23601, 0.2044, 0.9926, 34.76) and (0.1685, 0.1358, 0.9427, 19.01), respectively.

To visually assess the performance of the GEP model on input data, the interaction between observational and computational data is depicted in Figure 7. Notably, the curve's characteristics underscore the GEP model's outputs being in close proximity, exhibiting a shorter distance compared to the other two models. Figure 8 provides a graphical representation of the DDR index distribution for the three models. Evidently, the GEP model attains the highest point on the vertical axis during both the training and test periods, affirming its superiority for the well. Additionally, Figure 9 illustrates the

concordance between observational and computational data, coupled with the distribution of residual errors specifically for the GEP model.

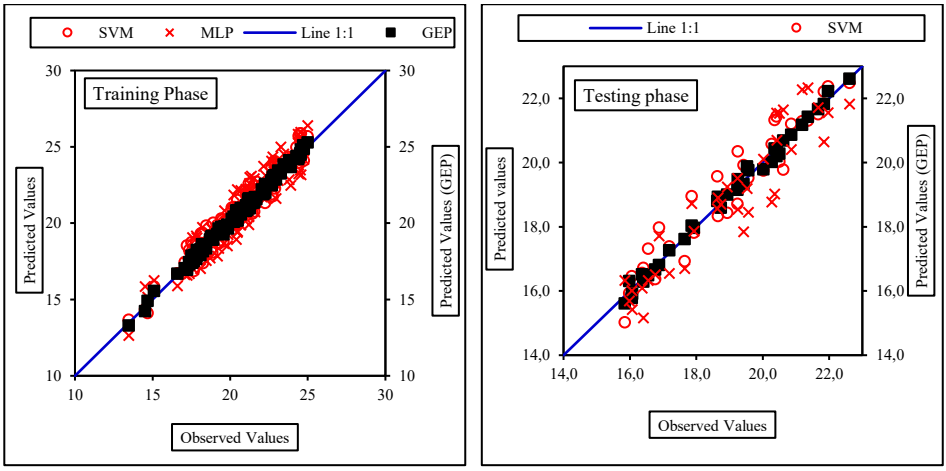

Figure 7: Scatter plot of the GEP outcome for the Well #30

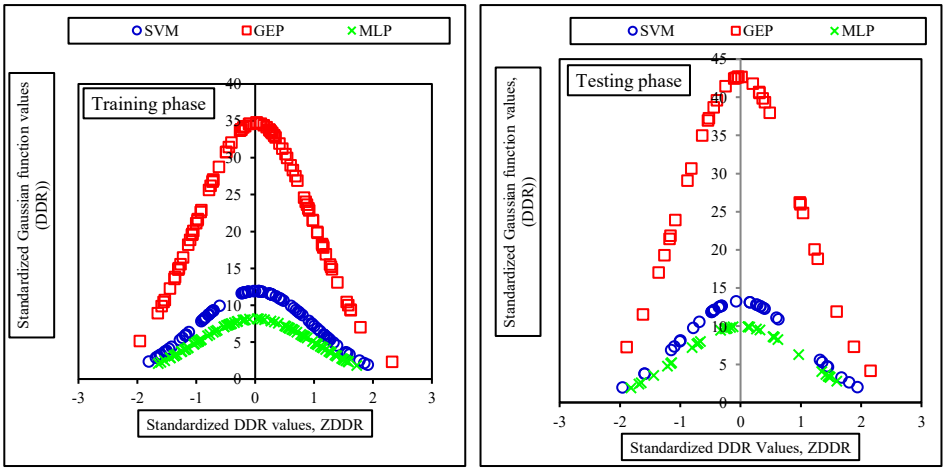

Figure 8: Distribution of DDR index for Well #30

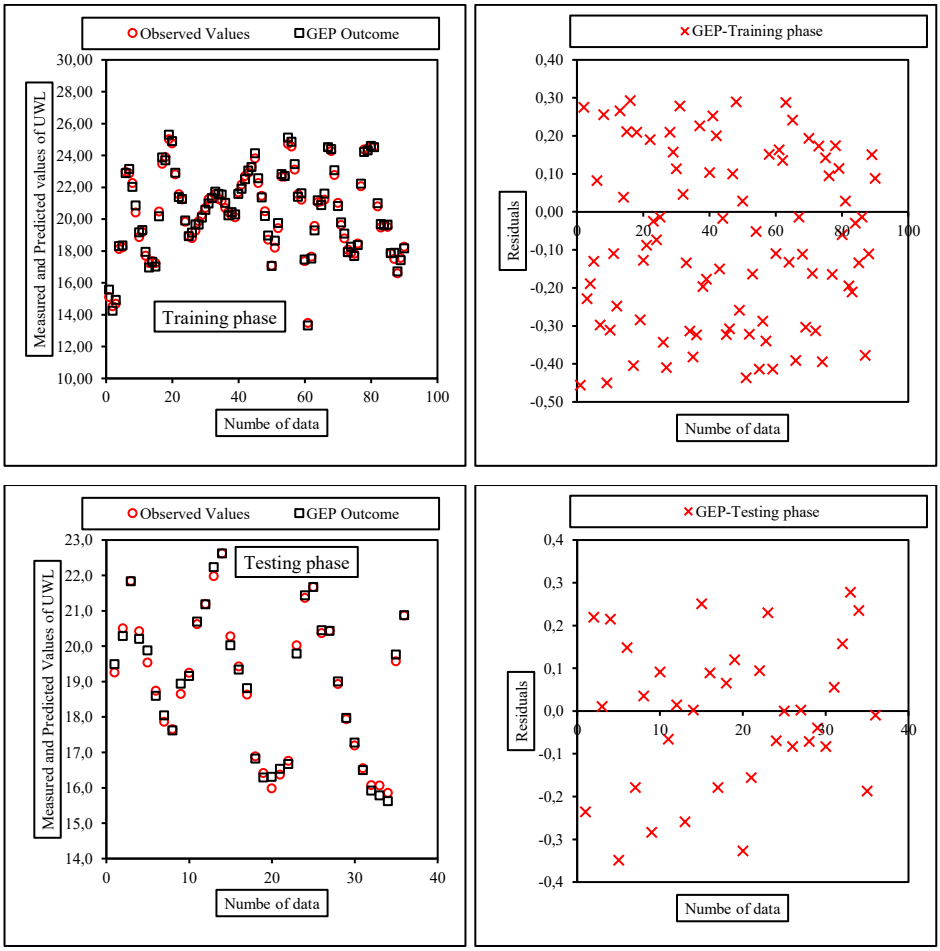

Figure 9: Performance of the SVM for well #30

The performance of the MLP model, identified as the superior model for well #21 based on the optimal combination of M1, is illustrated in Figs. 10 to 12. While the values of RMSE=0.3526, MAE=0.3035, R^2=0.9595, and DDR_{max}=22.86 shows the precise prediction of the MLP model for the training phase, those of the testing stage are RMSE=0.2562, MAE=0.2201, R^2=0.9826, DDR_{max}=33.22. In Figure 10, the evident proximity of points associated with the MLP model to the 1:1 line, in comparison to the two GEP and SVM models, is apparent. This characteristic is discernible in both the training and test phases. Figure 11 showcases the discrepancy in the maximum value of the DDR index on the axis, clearly emphasizing the superiority of the MLP model compared to the other two models. This distinction is noticeable in both training and test scenarios. Furthermore, Figure 12 presents the observed data values corresponding to each measured data point, alongside the distribution of residual errors for the MLP model.

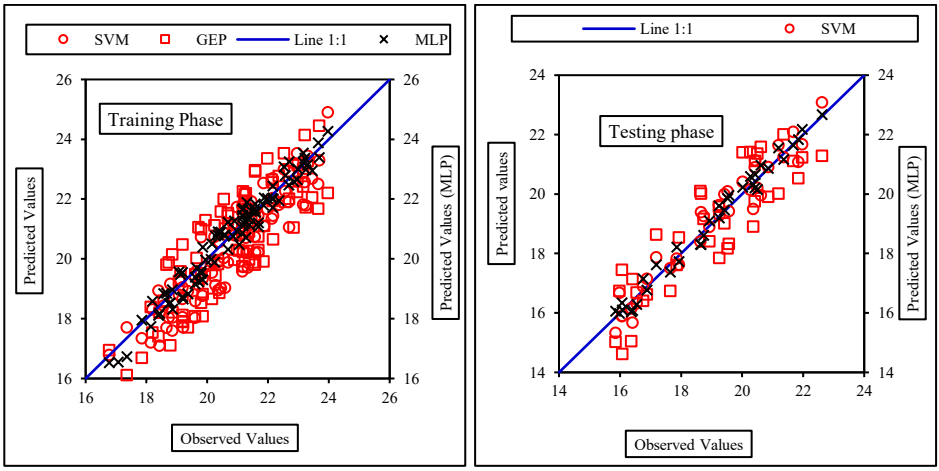

Figure 10: Scatter plot of the GEP outcome for the Well #21

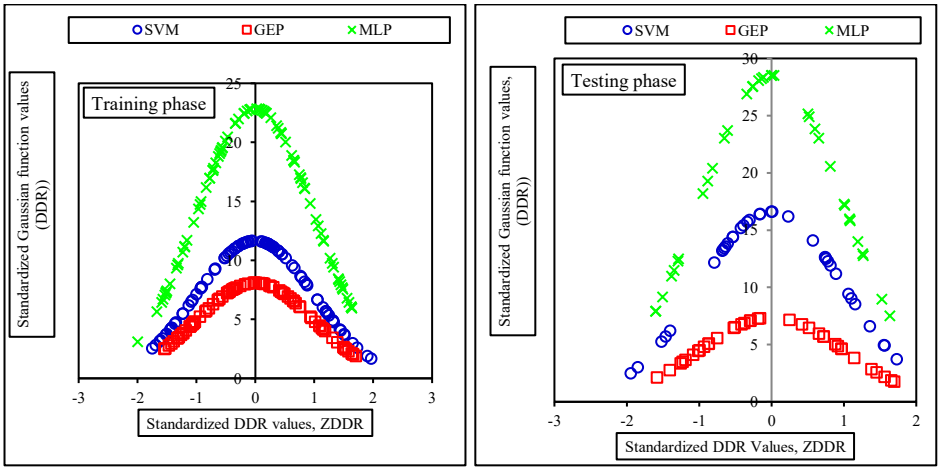

Figure 11: Distribution of DDR index for Well #21

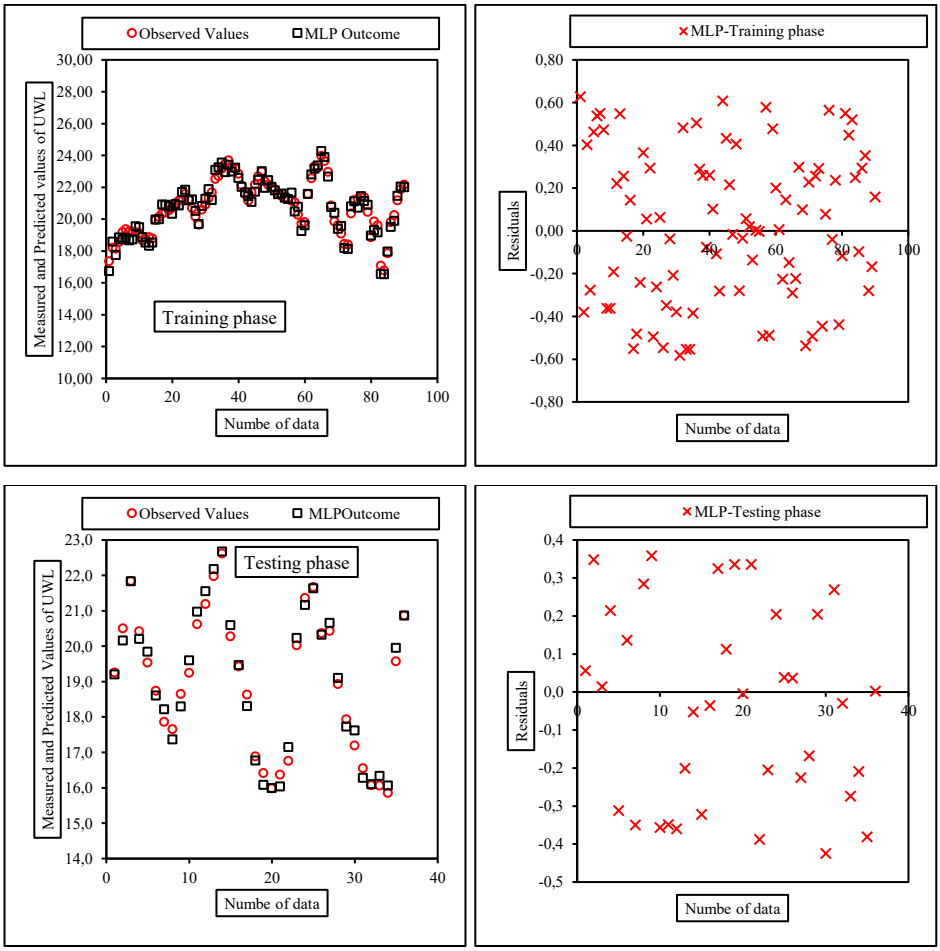

Figure 12: Performance of the SVM for #21

CONCLUSION

Given the paramount significance of subterranean water resources and the strategic imperative of their utilization across diverse sectors, precise anticipation of fluctuations in UWLs is of utmost importance. In this study, the efficacy of three distinct machine learning models namely SVM, GEP, and MLP was harnessed to prognosticate the UWLs within the Azarshahr plain of Tabriz, Iran. The analysis leveraged data spanning a three-year period derived from 34 observation wells. The assessment of model performance was executed through the deployment of RMSE, MAE, and DDR indices. The outcomes of the simulations revealed that all three models demonstrated commendable predictive capabilities regarding UWLs. The simulation results show the superiority of SVM, GEP

and MLP models in 53, 26 and 20 percent of the number of wells, respectively. The results indicated the M2 input combination (inputs with 2-lag) lead to the highest precise in the outcomes. The values of 0.2457, 0.2077, 0.9482, and 31.53 can be expressed as the average values of RMSE, MAE, R^2 and DDR_{max} for the SVM model during testing phase. One reason could be the effectiveness of the SVM in handling complex, nonlinear data patterns due to its ability to map inputs to higher-dimensional spaces using kernel functions, making it more adaptable to variations in groundwater levels.Increasing the lag time affected predictive accuracy by potentially adding unnecessary complexity to the models. While some lag time helps capture temporal dependencies, too much can introduce irrelevant past data, leading to overfitting and decreased accuracy. Each model exhibited the capacity to simulate outcomes with the desired precision through the amalgamation of input parameters. However, among the triad of models, the SVM emerged as the most adept in predicting UWLs, thereby earning distinction as the preeminent model across a substantial number of wells. The difference in accuracy across the models could be attributed to the inherent strengths of each algorithm in handling different types of nonlinear relationships, the choice of hyperparameters, and the ability of each model to generalize the patterns in the dataset. The complexity of the groundwater system and the chosen input configurations also play a role.

The findings imply that MLMs can serve as effective tools for groundwater management in arid and semi-arid regions, offering an alternative to traditional numerical models. Accurate UWL predictions allow for better resource planning and management in water-scarce regions.

Future studies could explore hybrid models combining the strengths of SVM, GEP, and MLP, or optimize the hyperparameters further. Additionally, incorporating more advanced machine learning techniques like deep learning or long short-term memory (LSTM) networks could enhance prediction accuracy for UWL.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

ABAIDIA S., REMINI B. (2020). Ghrib and Boukourdane (Algeria): when the water discharged from the dams feed the alluvial aquifer, Larhyss Journal, No 44, pp. 133-159.

ABBASZADEH SHAHRI A., SHAN C., LARSSON S. (2022). A novel approach to uncertainty quantification in groundwater table modeling by automated predictive deep learning, Natural Resources Research, Vol. 31, Issue 3, pp. 1351-1373.

ADEREMI B.A., OLWAL T.O., NDAMBUKI J.M., RWANGA S.S. (2023). Groundwater levels forecasting using machine learning models: A case study of the groundwater region 10 at Karst Belt, South Africa, Systems and Soft Computing, Vol. 5, Paper ID 200049.

AFZAAL H., FAROOQUE A.A., ABBAS F., ACHARYA B., ESAU T. (2019). Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning, Water, Vol. 12, Issue 5, pp 1-18.

AHMED A.K.A., EL-RAWY M., IBRAHEEM A.M., AL-ARIFI N., ABD-ELLAH M. K. (2023). Forecasting of Groundwater Quality by Using Deep Learning Time Series Techniques in an Arid Region, Sustainability, Vol. 15, Issue 8, pp. 1-16.

ALSHEHRI F., RAHMAN A. (2023). Coupling Machine and Deep Learning with Explainable Artificial Intelligence for Improving Prediction of Groundwater Quality and Decision-Making in Arid Region, Saudi Arabia, Water, Vol. 15, Issue 12, pp. 1-28.

ARGAZ A. (2018). 1d model application for integrated water resources planning and evaluation: case study of Souss river basin, Morocco, Larhyss Journal, No 36, pp. 217-229.

AROUA N. (2018). Water resources in SNAT 2030. between economic needs and ecological requirements, Larhyss Journal, No 35, pp. 153-168. (In French)

AZAMATHULLA H.M., AHMAD Z. (2013). Estimation of critical velocity for slurry transport through pipeline using adaptive neuro-fuzzy interference system and gene-expression programming, Journal of Pipeline Systems Engineering and Practice, Vol. 4, Issue 2, pp. 131-137.

AZAMATHULLA H.M., RATHNAYAKE U., SHATNAWI A. (2018). Gene expression programming and artificial neural network to estimate atmospheric temperature in Tabuk, Saudi Arabia, Applied Water Science, Vol. 8, pp. 1-7.

BAHIR M., EL MOUKHAYAR R., CARREIRA P. SOUHEL A. (2015). Isotopic tools for groundwater management in semi-arid area: case of the wadi Ouazzi basin (Morocco), Larhyss Journal, No 23, pp. 23-39.

BAHMANI R., OUARDA T.B. (2021). Groundwater level modeling with hybrid artificial intelligence techniques, Journal of Hydrology, Vol. 595, Paper ID 125659.

BAICHE A., SIDI MOHAMED H., ABLAOUI H. (2015). Overexploitation of water resources of the Mostaganem plateau aquifer, Larhyss Journal, No 22, pp. 153-165. (In French)

BALAHANG S., GHODSIAN M. (2023). Estimation of rectangular and triangular side weir discharge, ISH Journal of Hydraulic Engineering, Vol. 29, Issue 1, pp. 12-23.

BAUDHANWALA D., MEHTA D., KUMAR V. (2024). Machine learning approaches for improving precipitation forecasting in the Ambica River basin of Navsari District, Gujarat, Water Practice & Technology, Vol., 19, Issue 4, pp. 1315-1329.

BELHADJ M.Z., BOUDOUKHA A., AMROUNE A., GAAGAI A., ZIANI D. (2017). Statistical characterization of groundwater quality of the northern area of the basin of Hodna, M'sila, southeastern Algeria, Larhyss Journal, No 31, pp. 177-194. (In French)

BEMMOUSSAT A., ADJIM M., BENSAOULA F (2017). Use of the ZYGOS model for the estimation of groundwater recharge in Sikkak watershed (Northen west of Algeria), Larhyss Journal, No 30, pp. 105-119. (In French)

BIRBAL P., AZAMATHULLA H., LEON L., KUMAR V., HOSEIN J. (2021). Predictive modelling of the stage–discharge relationship using gene-expression programming, Water supply, Vol. 21, Issue 7, pp. 3503-3514.

CHIBANE B., ALI-RAHMANI S.E. (2015). Hydrological based model to estimate groundwater recharge, real- evapotranspiration and runoff in semi-arid area, Larhyss Journal, No 23, pp. 231-242.

CORTES C., VAPNIK V. (1995). Support-vector networks, Machine learning, Vol. 20, pp. 273-297

CORTES C., VAPNIK V.N. (1995). Support-vector networks, Machine learning, Vol. 20, Issue 3, pp. 273-297.

DADHICH A.P., GOYAL R., DADHICH, P.N. (2021). Assessment and prediction of groundwater using geospatial and ANN modeling, Water resources management, Vol. 35, pp. 2879-2893.

DAS U.K., ROY P., GHOSE D.K. (2019). Modeling water table depth using adaptive neuro-fuzzy inference system, ISH journal of hydraulic engineering, Vol. 25, Issue 3, pp. 291-297.

DEB S. (2024). Optimizing hydrological exploration through GIS-based groundwater potential zoning in Gomati district, Tripura, India, Larhyss Journal, No 60, pp. 231-256.

CHIBANE B., ALI-RAHMANI S.E. (2015). Hydrological based model to estimate groundwater recharge, real- evapotranspiration and runoff in semi-arid area, Larhyss Journal, No 23, pp. 231-242.

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, Vol. 148, pp. 181-191.

EL FELLAH IDRISSI B., CHERAI B., HINAJE S., MEHDI K. (2017). Climatic variability and its influence on water resources in the northern part of the middle atlas Moroccan: the case of the Sefrou and the Anoceur causses, Larhyss Journal, No 32, pp. 155-179. (In French)

EL MOUKHAYAR R., BAHIR M., CARREIRA P. (2015). Estimation of groundwater recharge in arid region through hydrochemistry and isotope: a case study Kourimat basin Morocco, Larhyss Journal, No 23, pp. 87-104.

FERREIRA C. (2001). Gene expression programming a new adaptive algorithm for solving problems, Complex systems, Vol. 13, Issue 2, pp. 87-129.

FULADIPANAH M., AZAMATHULLA H.M., TOTA-MAHARAJ K., MANDALA V., CHADEE A. (2023). Precise forecasting of scour depth downstream of flip bucket spillway through data-driven models, Results in engineering, Vol. 20, Paper ID 101604.

FULADIPANAH M., MAJEDI ASL M., JAFARINIA R. (2020). Application and assessment of svm algorithm to simulate the geometry of scour hole downstream of a siphon spillway, Iranian journal of irrigation & drainage, Vol. 14, Issue 3, pp. 1032-1045.

FULADIPANAH M., MAJEDIASL M. (2021). Assessment of the geometric shape of bridge pier on the scour depth using the support vector machine, Jwss-isfahan university of technology, Vol. 24, Issue 4, pp. 197-210.

GAALOUL N. (2015). Modeling of underground flows in unsaturated porous medium: application to artificial recharge by treated wastewater - Korba coastal water table (cap-bon, Tunisia), Larhyss Journal, No 21, pp. 181-190. (In French)

HAOUCHINE A., HAOUCHINE F.Z., LABADI A. (2015). Climatic change and anthropic activities: impacts on coastal aquifers in Algeria, Larhyss Journal, No 24, pp. 227-241. (In French)

HOUNTONDJI B., CODO F.P., AINA M.P. (2020). Characterization of hydrogeological conditions from the Monzoungoudo groundwater reservoir in Benin, Larhyss Journal, No 41, pp. 223-232. (In French)

JEIHOUNI E., ESLAMIAN S., MOHAMMADI M., ZAREIAN M.J. (2019). Simulation of groundwater level fluctuations in response to main climate parameters using a wavelet–ann hybrid technique for the shabestar plain, iran, Environmental earth sciences, Vol. 78, Issue 10, Paper ID 293.

KANTHARIA V., MEHTA D., KUMAR V., SHAIKH M.P., JHA S. (2024). Rainfall–runoff modeling using an Adaptive Neuro-Fuzzy Inference System considering soil moisture for the Damanganga basin, Journal of Water and Climate Change, Vol. 15, Issue 5, pp. 2518-2531.

KAUSHIK V., KUMAR M. (2023). Water surface profile prediction in non-prismatic compound channel using support vector machine (SVM), AI in civil engineering, Vol. 2, Issue 6, pp. 1-12.

KHEBIZI H., BENLAOUKLI B., CHAOUCHE M., CHACHA B. (2023). The Ghout of El Oued in Algeria: a patrimony and a natural hydroagrarian alarm system to advance, Larhyss Journal, No 55, pp. 107-124.

KOUASSI A.M., KOUAME K.F., SALEY M.B., BIEMI J. (2013). Impacts of climate change on groundwater of crystalline and Crystalophyllian basement aquifers in west Africa: case of the N'zi-Bandama watershed (Ivory Coast), Larhyss Journal, No 16, pp. 121-138. (In French)

KOUSSA M., BERHAIL S. (2021). Evaluation of spatial interpolation techniques for mapping groundwater nitrates concentrations - case study of Ain Elbel-sidi Makhlouf syncline in the Djelfa region (Algeria), Larhyss Journal, No 45, pp. 119-140.

LAGHZAL A., SALMOUN F., BOUDINAR B., ARGAZ A. (2019). Potential impact of anthropogenic activities on groundwater in the Tangier-Tetouan-Alhoceima region (Morocco), Larhyss Journal, No 34, pp. 39-50.

LATER F., LABADI A.S. (2024). Origin of the alluvial aquifer's groundwater in wadi Biskra (Algeria), Larhyss Journal, No 57, pp. 145-158.

LEON L.P., AZAMATHULLA H., FELIX P., PRASAD C.V.S.R. (2023). Prediction of stiffness modulus of bituminous mixtures using the applications of multi-expression programming and gene expression programming, Road materials and pavement design, Vol. 24, Issue 9, pp. 2192-2208.

LI H., LU Y., ZHENG C., YANG M., LI S. (2019). Groundwater level prediction for the arid oasis of northwest china based on the artificial bee colony algorithm and a back-propagation neural network with double hidden layers, Water, Vol. 11, Issue 4, p. 860.

MAJEDI-ASL M., FULADIPANAH M., ARUN V., TRIPATHI R.P. (2022). Using data mining methods to improve discharge coefficient prediction in piano key and labyrinth weirs, Water supply, Vol. 22, Issue 2, pp. 1964-1982.

MEHTA D., DHABUWALA J., YADAV S.M., KUMAR V., AZAMATHULLA H.M. (2023). Improving flood forecasting in Narmada River basin using hierarchical clustering and hydrological modelling, Results in Engineering, Vol. 20, Paper ID 101571.

MERONI E., PIÑEIRO G., GOMBERT P. (2021). Geological and hydrogeological reappraisal of the Guaraní aquifer system in the Uruguayan area, Larhyss Journal, No 48, pp. 109-133.

MOHAMMED K.S., SHABANLOU S., RAJABI A., YOSEFVAND F., IZADBAKHSH M.A. (2023). Prediction of groundwater level fluctuations using artificial intelligence-based models and GMS, Applied water science, Vol. 13, Issue 2, pp. 1-14.

MOZAFFARI S., JAVADI S., MOGHADDAM H.K., RANDHIR T.O. (2022). Forecasting groundwater levels using a hybrid of support vector regression and particle swarm optimization, Water resources management, Vol. 36, Issue 6, pp. 1955-1972.

NAKOU T.R., SENOU L., ELEGBEDE B., CODO F.P. (2023). Climate variability and its impact on water resources in the lower mono river valley in Benin from 1960 to 2018, Larhyss Journal, No 56, pp. 215-234.

NGOUALA M.M., MBILOU U.G., TCHOUMOU M., SAMBA-KIMBATA M.J. (2016). Characterization surface water - groundwater aquifer in coastal watershed of the republic of Congo Loémé, Larhyss Journal, No 28, pp. 237-256. (In French)

NICHANE M., KHELIL M.A. (2015). Climate change and water resources in Algeria - vulnerability, impact and adaptation strategy, Larhyss Journal, No 21, pp. 25-33. (In French)

NOORI R., KHAKPOUR A., OMIDVAR B., FAROKHNIA A. (2010) comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistics, Expert systems with applications, Vol. 37, pp. 5856-5862.

OUHAMDOUCH S, BAHIR M., CARREIRA P., CHKIR N., GOUMIH A. (2016). Climate change impact on Hauterivian aquifer of Essaouira basin (Morocco), Larhyss Journal, No 25, pp. 269-283. (In French)

QURESHI H.U., ABBAS I., SHAH S.M.H., TEO F.Y. (2024). Hydrologic evaluation of monthly and annual groundwater recharge dynamics for a sustainable groundwater resources management in Quetta city, Pakistan, Larhyss Journal, No 60, pp. 27-53.

RAJPUT D.C., MISTRY K.P., BHORANIYA J.K., UMRIGAR J.N., WAIKHOM S.I. (2023). Assessing the decadal groundwater level fluctuation - a case study of Gujarat, India, Larhyss Journal, No 54, pp. 175-191.

REMINI B. (2018). The foggaras of the oasis of Ghardaia (Algeria): the sharing of flood waters, Larhyss Journal, No 36, pp. 157-178. (In French)

SATTARI M.T., MIRABBASI R., SUSHAB R.S., ABRAHAM J. (2018). Prediction of groundwater level in Ardebil plain using support vector regression and m5 tree model, Groundwater, Vol. 56, Issue 4, pp. 636-646.

SEIFI A., EHTERAM M., SINGH V.P., MOSAVI A. (2020). Modeling and uncertainty analysis of groundwater level using six evolutionary optimization algorithms hybridized with ANFIS, SVM, and ANN, Sustainability, Vol. 12, Issue 10, pp. 1-42.

SRIVASTAVA D.K., SHUKLA A., JEMNI D. (2023). Prediction of ground water level in rajasthan state using machine learning, Procedia computer science, Vol. 218, pp. 1702-1711.

UMRIGAR J., MEHTA D.J., CALOIERO T., AZAMATHULLA H.M., KUMAR V. (2023). A comparative study for provision of environmental flows in the Tapi River, Earth, Vol. 4, Issue 3, pp. 570-583.

VADIATI M., RAJABI YAMI Z., ESKANDARI E., NAKHAEI M., KISI O. (2022). Application of artificial intelligence models for prediction of groundwater level fluctuations: case study (tehran-karaj alluvial aquifer), Environmental monitoring and assessment, Vol. 194, Issue 9, Paper ID 619.

VU M.T., JARDANI A., MASSEI N., FOURNIER M. (2021). Reconstruction of missing groundwater level data by using Long Short-Term Memory (LSTM) deep neural network, Journal of Hydrology, Vol. 597, Paper ID 125776.

YAO Z., WANG Z., WU T., LU W. (2024). A hybrid data-driven deep learning prediction framework for lake water level based on fusion of meteorological and hydrological multi-source data, Natural Resources Research, Vol. 33, Issue 1, pp. 163-190.

YI S., KONDOLF G.M., SANDOVAL SOLIS S., DALE L. (2024). Groundwater Level forecasting using machine learning: a case study of the Baekje Weir in Four Major Rivers Project, South Korea, Water Resources Research, Vol. 60, Issue 5, Paper ID e2022WR032779.

YOON H., JUN S.C., HYUN Y., BAE G.O., LEE K.K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer, Journal of hydrology, Vol. 396, Issues 1-2, pp. 128-138.

ZEGAIT R., REMINI B., BENSAHA H., (2021). Groundwater vulnerability assessment in the M'zab valley - southern Algeria, Larhyss Journal, No 48, pp. 211-234.

ZELLA L., SMADHI D. (2010). Water shortage in Arab countries and the need for the use of non-conventional water, Larhyss Journal, No 8, pp. 149-166. (In French)