Water Resources Management (2025) 39:779–794
https://doi.org/10.1007/s11269-024-03990-x

Advancing Reservoir Water Level Predictions: Evaluating Conventional, Ensemble and Integrated Swarm Machine Learning Approaches

Issam Rehamnia[1] · **Amin Mahdavi-Meymand**[2]

Received: 11 July 2024 / Accepted: 17 September 2024 / Published online: 25 September 2024

Abstract

Accurate estimation of reservoir water level fluctuation (*WLF*) is crucial for effective dam operation and environmental management. In this study, seven machine learning (ML) models, including conventional, integrated swarm, and ensemble learning methods, were employed to estimate daily reservoir *WLF*. The models comprise multi-linear regression (MLR), shallow neural network (SNN), deep neural network (DNN), support vector regression (SVR) integrated with homonuclear molecules optimization (HMO) and particle swarm optimization (PSO) meta-heuristic algorithms, classification and regression tree (CART), and random forest (RF). These models were trained and evaluated using in situ data from three embankment dams in Algeria: the Kramis dam, the Bougous dam, and the Fontaine Gazelles dam. Performance evaluation was conducted using statistical indices, scatter plots, violin plots, and Taylor diagrams. The results revealed superior prediction accuracy for the Fontaine Gazelles dam compared to Kramis and Bougous dams. Particularly, the RF, DNN, and SVR-HMO models exhibited consistent and excellent predictive performance for *WLF* at the Fontaine Gazelles dam with *RMSE* values of 0.502 m, 0.536 m, and 0.57 m, respectively. The RF model demonstrates remarkable accuracy across all three case studies. This can be attributed to the ensemble structure of RF, as evidenced by the results. This study underscores the significance of considering factors such as seepage flow intensity in understanding *WLF* variability. Furthermore, the proposed ML models offer promising capabilities in *WLF* prediction, highlighting their potential utility in enhancing reservoir management practices and addressing the limitations of traditional regression models. Keys words. Embankment dam, Water level fluctuations, Seepage, Artificial neural network, meta-heuristic algorithm.

Keywords Embankment dam · Water level fluctuations · Seepage · Srtificial neural network · Meta-heuristic algorithm

Extended author information available on the last page of the article

1 Introduction

A storage dam is one of the main hydraulic structures in water management. Storage dams may be constructed for various purposes, including producing green energy, providing fresh water for drinking, agriculture, industrial purposes, controlling floods, and supporting fisheries. Hence, accurate estimation of *WLF* in dams emerges as a fundamental necessity for effective management of these diverse needs (Altunkaynak and Şen 2007). The storage dams significantly change the ecosystems and monitoring the *WLF* is crucial for environmental management projects. The prediction of *WLF* poses a complex challenge due to high non-linearity and involves various hydrological and hydraulic aspects, including direct precipitation, evaporation, groundwater exchange, river flows, and seepage. Additionally, factors like wind can influence water levels, further complicating the problem. In pursuit of this objective, various methodologies have been devised to assess and forecast *WLF* in dams. These include radar altimetry, echo sounder, and sonar techniques, each offering distinct approaches elucidated in studies like those conducted by Su et al. (2015), Shang et al. (2019), Bio et al. (2022), and Bandini et al. (2020). Additionally, numerical models and conceptual hydrological models, as discussed by Knotters and Bierkens (2000), Gong et al. (2016), and Zhang et al. (2017), have been developed for this purpose. However, many of these formulas, techniques, and models have drawbacks. Traditional models, for instance, often grapple with timescale conversion issues and may struggle to accurately predict lake level variations on shorter timescales (Zhang et al. 2017). In addition to the traditional and hydrological models mentioned, numerous studies have explored artificial intelligence (AI) approaches for predicting *WLF* problems (Ünes et al. 2015; Emami and Parsa 2020; Malekpour and Tabari, 2020).

Despite the broad spectrum of AI applications in the water and environmental fields (Rehamnia et al. 2021, 2023), these models have particularly stood out for their capability to predict water fluctuations across various domains. These studies encompass a wide range of phenomena, including sea level (Wang et al. 2020), lake levels (Wen et al. 2019; Buyukyildiz et al. 2014), groundwater level (Kow et al. 2024; Yadav et al. 2017), dam piezometric water level (Ziggah et al. 2022), and dams water level and capacity (Emami and Parsa 2020; Malekpour and Tabari, 2020; Das et al. 2016; Ünes et al. 2015). The models used in these studies offer valuable insights into the dynamics of water systems, aiding in better management and decision-making processes. Buyukyildiz et al. (2014) employed multi-layer perceptron neural network (MLPNN), adaptive neuro-fuzzy inference system (ANFIS), radial basis neural network (RBNN), SVR, and MLPNN tuned PSO to forecast the monthly variations of water level in Lake Beysehir in Turkey. Results showed that the SVR outperformed its counterparts, achieving a notably high correlation coefficient of $R^2=0.9988$. Ünes et al. (2015) demonstrated the superiority of autoregressive moving average (ARMA) and conventional autoregressive (CAR) models over MLR in predicting the daily water levels at Millers Ferry Dam, located in Alabama, USA. They also noticed that the MLPNN model performed well in all conditions and outperformed the other three models. Das et al. (2016) proposed a modeling strategy based on hybrid Bayesian network for predicting daily water levels of Massanjore dam in India. The potential of this model was assessed against autoregressive integrated moving average (ARIMA) and nonlinear approaches. The proposed methodology offers a more effective and timely mechanism compared to alternative models. Yadav et al. (2017) evaluated the results of extreme learning machine (ELM) and

SVR models for monthly groundwater level prediction. The findings demonstrated that the highest accuracy was achieved by ELM model with $R^2=0.901$ and $R^2=0.862$ for first and second station, respectively. Nadiri et al. (2019) studied the skill of three fuzzy logic methods—Sugeno, Mamdani, and Larsen— for predicting groundwater level variations in the Duzduzan–Bilverdi area in Tabriz, Iran. Following the modeling process, the Sugeno model outperformed other models. The study of Wang et al. (2020) involved a detailed comparison of the accuracy of artificial neural network (ANN), wavelet ANN (WANN), and wavelet ANFIS (WANFIS) models in forecasting sea surface levels. The findings demonstrated that WANFIS exhibits notable efficiency. Moreover, the feasibility of integrating wind information into the models was highlighted. A modeling study conducted by Malekpour and Tabari (2020) involved an evaluative comparison of SVR, ANFIS, ANN, and RBFNN models in predicting monthly water levels. These models were also combined with a supervised intelligence committee machine (SICM). Among all models, including both regular and hybrid ones, the SICM-ANN exhibited superior performance with a correlation coefficient of R^2 0.97. Emami and Parsa (2020) conducted a study on Shaharchay dam, situated in the northwest of Iran. They indicated that the imperialist competitive algorithm (ICM) is a powerful tool for solving reservoirs optimization functions. Malekpour et al., (2022) compared the potential of wavelet SVR integrated with the teaching learning-based algorithm (WSVR-TLBO) with SVR, ANN, ARIMA, and generalized regression neural network (GRNN) models in predicting reservoir water level. The results demonstrated that the hybrid model (WSVR-TLBO) was more accurate than all other models. Özdoğan-Sarıkoç et al. (2022) compared the ANN, long short-term memory (LSTM), and SVR models in modeling reservoir volumes in the Ladik and Yedikir reservoirs and found that LSTM was the best among the three models. Ziggah et al. (2022) compared the potential of seven models including SVR, M5 tree, group method of data handling (GMDH), MLPNN, RBNN, and Gaussian process regression (GPR) in predicting dam piezometers. The methods were tested by using the data collected from four distinct piezometers, each located at separate locations within the X dam in Ghana. The findings indicated that the GMDH model gives the best result among all the methods applied. Azad et al. (2022) applied the seasonal autoregressive integrated moving average (SARIMA), ANN and Hybrid SARIMA-ANN models for predicting Red Hills reservoir water level. They demonstrated that the SARIMA-ANN hybrid model outperformed the other two models, with an $R^2=0.845$. Yuan et al. (2022) used data from 19 different stations along the Yangtze River and employed a clustering method and the LSTM to model *WLF*. The clustering method divided the data into six clusters, and then an LSTM was trained on each cluster. The results indicated the efficiency of the suggested methodology in predicting *WLF*. Liang et al. (2023) applied the extreme gradient boosting tree (XGBoost) for reconstructing the water levels in several North American lakes. The XGBoost model showed the highest accuracy compared to the other models. Huang et al. (2023) used a RF method to model the *WLF* in the downstream of reservoirs. They suggested the application of RF for modeling similar environmental problems. Stefenon et al. (2023) suggested wavelet sequence-to-sequence LSTM for prediction of *WLF* in dams. The results indicated that the used LSTM works efficiently as a decision maker for reservoir management. Kim et al. (2024) compared the potential of DNN and LSTM for predicting flood water level. The LSTM model, with a correlation coefficient of 0.98, was ranked as the best model. Dai et al. (2024) developed a hybrid DNN for forecasting *WLF*. The results indicated that the developed model achieves good accuracy in comparison with the state-of-the-art baselines. Gao et al. (2024) used an improved version of PSO for tuning the LSTM parameters in predicting *WLF* in front of the check gate. The results approved the skill of meta-heuristic algorithms in enhancing the capability of LSTM.

The main objective of the current study is to model the daily *WLF* across three embankment dams located in Algeria: the Kramis dam, the Bougous dam, and the Fontaine Gazelles dam. Despite the challenges posed by limited available data across diverse geographical locations, this research endeavors to employ two distinct sets of combination variables for predicting WLF. The first set of predictors, which includes the seepage rate and piezometric measurements, was used for the Bougous and Fontaine Gazelles dams. The second set of predictors, which comprises drainage measurements obtained from gutters (channels) and the seepage rate, was used for the Kramis dam. algorithms such as MLR, CART, and shallow neural networks (SNN), as well as advanced ones including DNN, RF, SVR integrated with HMO, and SVR integrated with PSO developed for this purpose. The contributions and novelties of this study can be highlighted through several key points such as computational methods, modeling process, and comprehensive comparison. The HMO is recently introduced meta-heuristic algorithm (Mahdavi-Meymand and Zounemat-Kermani 2022). In this study, the integration of SVR with HMO was applied for the first time to model an engineering issue. This study suggests SVR-HMO as a new computational method for solving environmental and hydrologic problems. The inputs used for modeling *WLF* are unique due to the selection of three distinct case studies. Hence, the modeling methodology is different from that in the literature. Finally, the performances of novel models were compared with different types of algorithms. This comparison might guides readers for their future direction of studies.

2 Material and Methodology

2.1 Machine Learning Methods

In this study, the MLR and six ML models, including SNN, DNN, CART, RF, and SVR integrated with PSO and HMO, were developed for predicting daily *WLF*. The MLPNN is the most well-known and commonly applied ANN for simulating complex problems. The theory of MLPNN is inspired by the function of a biological neuron (Sun et al. 2010). The structure of an MLPNN consists of three layers, and each layer is made up of several nodes called neurons. The neurons of the first layer return the inputs of the problem, and thus, the number of neurons is equal to the number of inputs. The second layer is called middle or hidden layer. An MLPNN with one hidden layer is referred to as an SNN, and a DNN is an MLPNN with several hidden layers (Mahdavi-Meymand and Sulisz 2023). The last layer generally consists of one neuron that returns the output of the network. In this study, the number of neurons and layers were determined by a trial and error methodology. An SNN with 20 neurons and a DNN with five layers, each containing 20 neurons, were considered. The Bayesian optimization algorithm was applied to calculate the weights and biases.

Cortes and Vapnik (1995) introduced the support vector machine (SVM) as an innovative method for solving classification problems. SVR is a type of SVM algorithm is intended for regression tasks. In practical applications of SVR determining the optimal values for the regularization index, insensitive loss parameter, and kernel function parameter poses a challenge. The meta-heuristic algorithms offer an attractive solution for addressing this challenge. In this study two swarm algorithm including PSO and HMO were used to optimize the SVR parameters. HMO with inspiring Bohr atomic model, has fewer tendency to fall into local solutions while still being able to find them (Mahdavi-Meymand and Zounemat-Kermani 2022).

CART is a commonly used decision tree method for classification and regression problems (Breiman et al. 1984). CART algorithm splits a dataset into multiple sub-datasets by generating a graphic that resembles a tree. The sub-datasets are more homogeneous than the original dataset. The tree is generated based on decision rules present in the nodes. A linear regression model is fitted to each terminal node's data points following the generation of the tree and the clustering of the data.

RF is an ensemble of several decision trees within a single structure (Breiman 2001). RF employs a bagging strategy in its structure to combine the trees. The regression trees are generated on a subset of the training data selected by a random procedure. It is highly possible that some data points are selected several times (Lahouar and Slama 2015). In this study, the CART algorithm was used as the base model for generating the trees, with the number of trees set to 100. All models were developed in the MATLAB environment.

2.2 Modeling Process and Data

This study delves into the prediction of *WLF* utilizing data from three prominent earthen dams situated in Algeria: the Kramis dam, the Bougous dam, and the Fontaine Gazelles dam. The first dam, the Kramis, boasts a rockfill structure with a central clay core, strategically positioned on the Kramis River in Mostaganem city. Standing tall at a height of 61.5 m, it offers an annual useful water storage capacity of approximately 33.38 million cubic meters. The Bougous dam, an embankment structure with a vertical clay core, holds its ground in the northeastern part of Algeria, near the Algerian-Tunisian border within El-Taref Province. Towering at a height of 71.4 m, it provides an annual useful water storage capacity of 56 million cubic meters, with a regularized volume exceeding 71 hm3/year. Lastly, the Fontaine Gazelles dam, featuring a compacted alluvium embankment with a clay core, resides on the El Hai River in Biskra city. Standing at a height of 42.5 m, it boasts an annual useful water storage capacity of about 48.5 million cubic meters, with a regularized volume surpassing 14 million cubic meters per year. Figure 1 shows the location of considered dams.

The described methods were used to model the *WLF* of dams. The *WLF* of Kramis was simulated using five predictors, including seepage flow (S) and drainage (h) of four different gutters. Two gutters are located on the upstream side and two on the downstream side, extending from the left bank to the right bank of the dam. In each channel, there is a piezometric probe to measure the water level. The data used from this dam spanned from December 31, 2005, to December 30, 2018. The considered inputs of Fontaine Gazelles include S and the piezometric levels of piezometers P_7 to P_{10}, which were recorded from January 14, 2010, to May 29, 2017. For the Bougous dam, the inputs were S and the piezometric levels of piezometers P_1 to P_6. The data cover a period from February 16, 2012, to March 30, 2019. For all three cases, the data were randomly divided to four groups. A 4-fold modeling strategy was used to train the models. In this modeling process, the models are trained using three datasets, with the remaining one used for final assessing. In this study, a portion

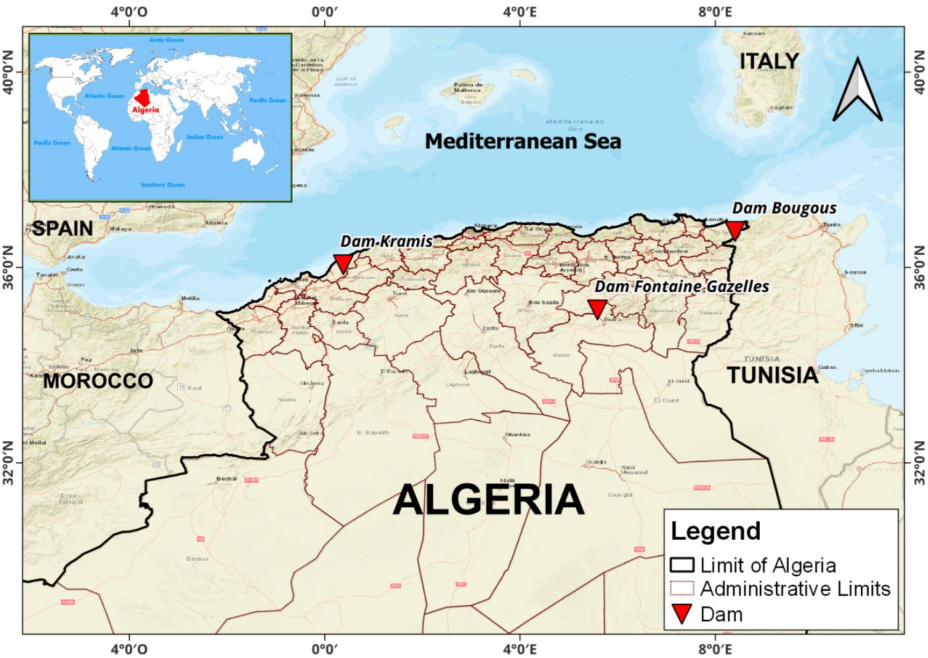

Fig. 1 The location of Kramis, Bougous, and Fontaine Gazelles embankment dams, Algeria

of the training datasets, specifically 15% of all data, were utilized to mitigate overfitting. In the 4-fold modeling methodology employed, the models for each dam were trained four times with different training data. The average performance of the models was considered to assess the results. Table 1 summarizes the dataset used in this study.

2.3 Performance Criteria

In this study, scatter plot, Taylor diagram, violin plot, and five statistical indices were used to analyze the performance of applied models in predicting daily reservoir *WLF*. The used indices include root mean square error (*RMSE*), coefficient of determination (R^2), mean absolute error (*MAE*), Nash-Sutcliffe efficiency (*NSE*), and index of agreement (*IA*), which can be calculated using the following equations:

$$RMSE = \sqrt{\frac{1}{N}\sum {}_{i=1}^{N} \left(M_i - P_i\right)^2} \tag{1}$$

$$R^2 = \left[\frac{\sum {}_{i=1}^{N}\left(M_i - \bar{M}\right)\left(P_i - \bar{P}\right)}{\sqrt{\sum {}_{i=1}^{N}\left(M_i - \bar{M}\right)^2}\sqrt{\sum {}_{i=1}^{N}\left(P_i - \bar{P}\right)^2}}\right]^2 \tag{2}$$

Table 1 Summarized data details of studied dams

Fontaine Gazelles	*Statistical parameter*	Inputs parameters						Output	
		S (l/s)	P_6 (m)	P_7 (m)	P_8 (m)	P_9 (m)	P_{10} (m)	*WFL* (m)	
	Min	2.1	31.41	19.85	8.62	0.9	5.18	374.55	
	Max	6.35	33.84	21.29	9.37	1.72	5.58	383.44	
	Avg	3.57	31.76	20.24	9.04	1.36	5.4	378.8	
	Sd	1.4	0.23	0.2	0.18	0.19	0.09	2.56	
Bougous	Statistical parameter	Inputs and output parameters							Output
		S (l/s)	P_1 (m)	P_2 (m)	P_3 (m)	P_4 (m)	P_5 (m)	P_6 (m)	*WFL* (m)
	Min	0.01	23.12	54.32	1.5	17.06	23.02	0	122.52
	Max	64.11	56.37	57.75	19.79	18.64	36.75	1.17	232.21
	Avg	11.59	35.25	56.39	7.61	17.77	23.58	0.07	135.21
	Sd	16.37	9.33	0.74	4.99	0.39	0.98	0.26	7.47
Kramis	Statistical parameter	Inputs and output parameters						Output	
		S (l/s)	h_1 (m)	h_2 (m)	h_3 (m)	h_4 (m)		*WFL* (m)	
	Min	0.04	0	0	1	0		86.6	
	Max	1.15	23.66	27.04	36.7	34.63		108.3	
	Avg	0.23	6.52	0.92	6.4	3.05		98.69	
	Sd	0.1	2.52	1.48	3.63	2.48		2.29	

P represents piezometer, *h* represents drainage measurements obtained from gutters, *Min* denotes minimum values, *Max* denotes maximum values, *Avg* denotes average values, and *Sd* denotes standard deviation

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |M_i - P_i| \tag{3}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(M_i - P_i)^2}{\sum_{i=1}^{n}\left(M_i - \bar{M}\right)^2}$$

$$IA = 1 - \frac{\sum_{i=1}^{N}(M_i - P_i)^2}{\sum_{i=1}^{N}\left(\left|M_i - \bar{M}\right| + \left|P_i - \bar{M}\right|\right)^2} \tag{4}$$

where N represents the number of data set, M and P denote the values of measured and predicted points, respectively, $\overline{M}$ is the average of M, and $\overline{P}$ is the average of P.

3 Results

The simulation results of the developed models for the Kramis dam are presented in Table 2. For this case, most of the seven models displayed acceptable predictions of *WLF* during the study period, although exceptions were observed with models such as RF and CART, which exhibited higher correlations. During the training phase, the models demonstrated the *RMSE* values ranging from 2.055 m to 0.347 m, the *MAE* values ranging from 1.387 m to 0.141 m, the *IA* values ranging from 0.57 to 0.994, and the *NSE* values ranging from 0.216 to 0.978. The SVR-PSO and SVR-HMO model yielded the highest accuracy with the lowest average *RMSE* of 0.347 m and 0.349 m, respectively, followed by CART, RF, DNN, SNN, and MLR. During the testing phase, the RF model exhibited the most accurate predictions,

Table 2 Performances of different methods in modeling *WLF* for the Kramis, Bougous, and Fontaine Gazelles embankment dam

Dam	Model	Training phase					Testing phase				
		RMSE	R^2	*MAE*	*NSE*	*IA*	*RMSE*	R^2	*MAE*	*NSE*	*IA*
Kramis	RF	0.698	0.914	0.368	0.91	0.974	0.995	0.814	0.507	0.809	0.942
	SNN	1.368	0.652	0.957	0.652	0.884	1.508	0.575	1.019	0.564	0.857
	DNN	1.196	0.732	0.831	0.732	0.916	1.422	0.624	0.931	0.611	0.879
	CART	0.543	0.944	0.289	0.944	0.985	1.055	0.793	0.486	0.785	0.942
	SVR-HMO	0.349	0.977	0.141	0.977	0.994	1.41	0.613	0.56	0.61	0.853
	SVR-PSO	0.347	0.978	0.153	0.978	0.994	1.421	0.609	0.561	0.606	0.85
	MLR	2.055	0.216	1.387	0.216	0.57	2.058	0.194	1.399	0.19	0.561
Bougous	RF	1.39	0.891	0.82	0.884	0.965	2.235	0.725	1.391	0.693	0.904
	SNN	2.038	0.764	1.302	0.764	0.931	2.613	0.63	1.746	0.591	0.877
	DNN	1.73	0.798	1.114	0.798	0.936	2.42	0.692	1.495	0.653	0.897
	CART	1.003	0.941	0.511	0.941	0.985	2.797	0.604	1.559	0.551	0.875
	SVR-HMO	0.871	0.965	0.821	0.958	0.988	2.852	0.558	2.008	0.513	0.837
	SVR-PSO	2.175	0.744	1.835	0.735	0.921	3.042	0.543	2.194	0.462	0.844
	MLR	2.954	0.516	2.201	0.516	0.82	3.254	0.464	2.353	0.38	0.786
Fontaine Gazelles	RF	0.303	0.986	0.229	0.986	0.996	0.502	0.963	0.375	0.957	0.989
	SNN	0.535	0.956	0.405	0.955	0.988	0.594	0.948	0.447	0.942	0.985
	DNN	0.486	0.969	0.38	0.965	0.991	0.536	0.956	0.417	0.955	0.988
	CART	0.271	0.988	0.198	0.988	0.997	0.592	0.947	0.396	0.945	0.986
	SVR-HMO	0.162	0.996	0.145	0.996	0.999	0.57	0.953	0.34	0.949	0.987
	SVR-PSO	0.139	0.997	0.123	0.997	0.999	0.577	0.951	0.337	0.948	0.986
	MLR	0.575	0.948	0.451	0.948	0.987	0.632	0.936	0.492	0.933	0.983

with the lowest *RMSE* of 0.995 m and *MAE* of 0.507 m, followed closely by the CART model with an *RMSE* of 1.055 m and *MAE* of 0.452 m. Furthermore, the RF and CART models again demonstrated superior predictive accuracy when assessed using R^2. The R^2 values clearly demonstrate that the predicted results of MLR are far from the actual values, confirming that *WLF* is a complicated problem and requires robust algorithms.

To have a visual comparison, the scatter plots of developed models in predicting *WFL* of Kramis dam for the best k-fold are displayed in Fig. 1(a). The graphs of algorithms indicate that the predicted *WLF* is acceptable. For all of models, the scatter points are located near the 1:1 line. Although the training points of SVR-HMO and SVR-PSO are nearly perfectly located on the 1:1 lines, the testing points are more scattered than expected. In the training and testing phase, the DNN points are more scattered than those of other advanced ML models. The strong correlation between the predicted values of RF and the actual values is evident. RF can be introduced as the best method for the Kramis embankment dam.

Table 2 also presents the results of simulation for the Bougous embankment dam. Despite achieving acceptable results during the learning phase, the models struggle to effectively predict *WLF* in the testing phase. In the training phase, SVR-HMO exhibited the best results, with the lowest average *RMSE* of 0.871 m, *MAE* of 0.821 m, and the highest R^2 of 0.965 and *NSE* of 0.958. However, in the testing phase, the performance of SVR-HMO did not meet expectations ($R^2 = 0.558$). The MLR with an R^2 less than 0.5 demonstrated inferior results when compared with other techniques. In the testing data sets, the RF outputs, with the lowest *RMSE* of 2.235 m, *MAE* of 1.391 m, and the highest R^2 of 0.725, *NSE* of 0.693, and *IA* of 0.904, exhibited the highest correlation with the predicted values. In contrast to

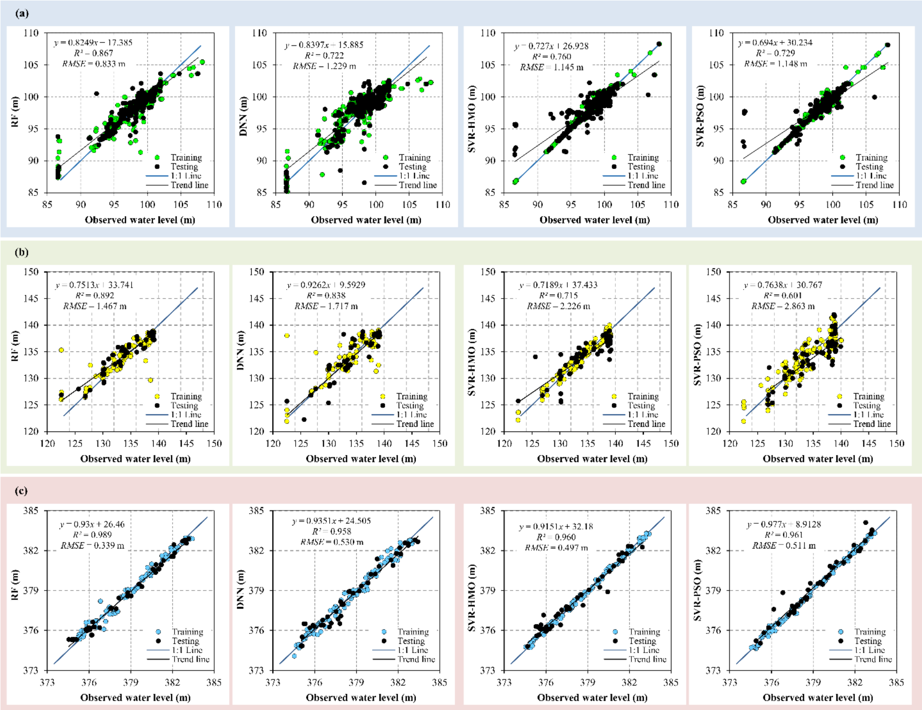

Fig. 2 Performance plots of the applied methods during both the training and testing phases for (**a**) Kramis embankment dam, (**b**) Bougous embankment dam, and (**c**) Gazelles embankment dam

the Kramis dam, DNN with *RMSE* of 2.42 m outperforms CART in this particular dam. The testing results corroborate that while the accuracies of SVR-HMO and SVR-PSO are comparable, SVR-HMO proves to be marginally more precise. Another significant finding is that, in comparison to other advanced ML models, the SNN with an *RMSE* of 2.613 m, ranks as the least accurate method. Figure 2(b) showcases the scatter plots for the models applied to the Bougous dam, representing the best achieved scenario. There is a strong correlation between the predicted values and actual values for the RF and DNN. The scatter plots clearly reveal the superiority of the HMO over the PSO in finding optimal parameters of SVR.

The results of the last dam, involving the Fontaine Gazelles embankment dam, are presented in Table 2; Fig. 2(c). The results demonstrate outstanding performance across all models during both the training and testing phases. Particularly noteworthy is the performance of CART, SVR-HMO, and SVR-PSO, which exhibit highly accurate predictions of *WLF* within this dam. The average results confirm that, similar to the two previous dams, the RF model with an *RMSE* of 0.502 m stands out as the most accurate method for predicting *WLF* at the Fontaine Gazelles dam. The scatterplots indicate that all predicted *WLFs* are precise. The scatter plots show that all model points are close to the 1:1 line during both the training and testing phases, with the RF model demonstrating notably better performance than the others. The scatter plots show that the SVR-based models have achieved an excellent fit to the training data, indicating the robust capabilities of meta-heuristic algorithms in finding optimal values of ML models.

For additional evaluation the skill of the applied methods, violin plots illustrating the trend variation calculated against the observed values for the best scenarios are shown for the testing phase (Fig. 3). The violin shape of predicted *WLF* within each method exhibits considerable variability across the three dams. This variability is most pronounced for the Kramis and Bougous dams. Additionally, it is clear from the graphs that none of the seven methods demonstrate the capability of predicting *WLF* with high precision. This indicates that the models may have failed to encompass certain values, and the shape of the violins suggests that the outliers are poorly estimated. However, the violin shape of RF and CART in these two dams is more similar to the predicted violin shape, confirming their higher performance compared to other models. It can be inferred that in the Fontaine Gazelles dam, all models predict WLF very effectively. The RF model, in particular, is adept at capturing nonlinear behavior compared to the MLR, DNN, SVR-HMO, SVR-PSO, and CART approaches.

The estimation accuracy of all proposed models was visually depicted using Taylor diagrams (Fig. 4). These diagrams present the standard deviation (σ), the correlation coefficient (r), and the centered *RMSE* (*CRMSE*) in a single plot. Taylor diagrams enable a better visual understanding of the power and accuracy of each model. On average, for the Kramis dam, the r values for all models are situated below 0.820, with the exception of the RF model, which slightly exceeded 0.90. Meanwhile, the σ value of the CART model was the closest to the measured value (illustrated by the red dotted line in Fig. 4). The MLR point is far from the exact point indicating its poor accuracy. The performance points of the other methods are situated in close proximity, suggesting that their performance is comparable. For the Bougous Dam, the r values of the seven models generally fell within the range of 0.700 to 0.800, except for the MLR model, which ranged between 0.400 and 0.500. Concerning the Fontaine Gazelles Dam, the Taylor diagram shows that all ML models attained high r values while maintaining low *CRMSE*. In general, the r values of the seven models were situated within the range of 0.95 to 0.99, indicating that all seven models closely approximated the measured value.

4 Discussion

In this study, the performance of seven ML models in predicting the *WLF* of three different dams was analyzed. The outcomes of the developed models are compared with the leading results from other AI-based studies carried out by researchers. The results indicated that the performance of the ML models in two dams was not perfect, as was expected. Previous studies, including those by Ziggah et al. (2022), Liu et al. (2023), Leyla et al. (2023), and Dayal et al. (2024), have underscored the high accuracy of ML models in modeling *WLF* in reservoirsWhile it may initially appear that ML methods were ineffective in modeling these two dams, it is important to discuss two crucial points. Firstly, all the developed ML models were more accurate than MLR, which serves as a suitable benchmark for comparing the efficiency of models. Secondly, the considered inputs have affected the results. Ziggah et al. (2022) used four inputs—rainfall, temperature, pressure, and modulus—to predict the water level in dam piezometers. Liu et al. (2023) used the previous time steps to model

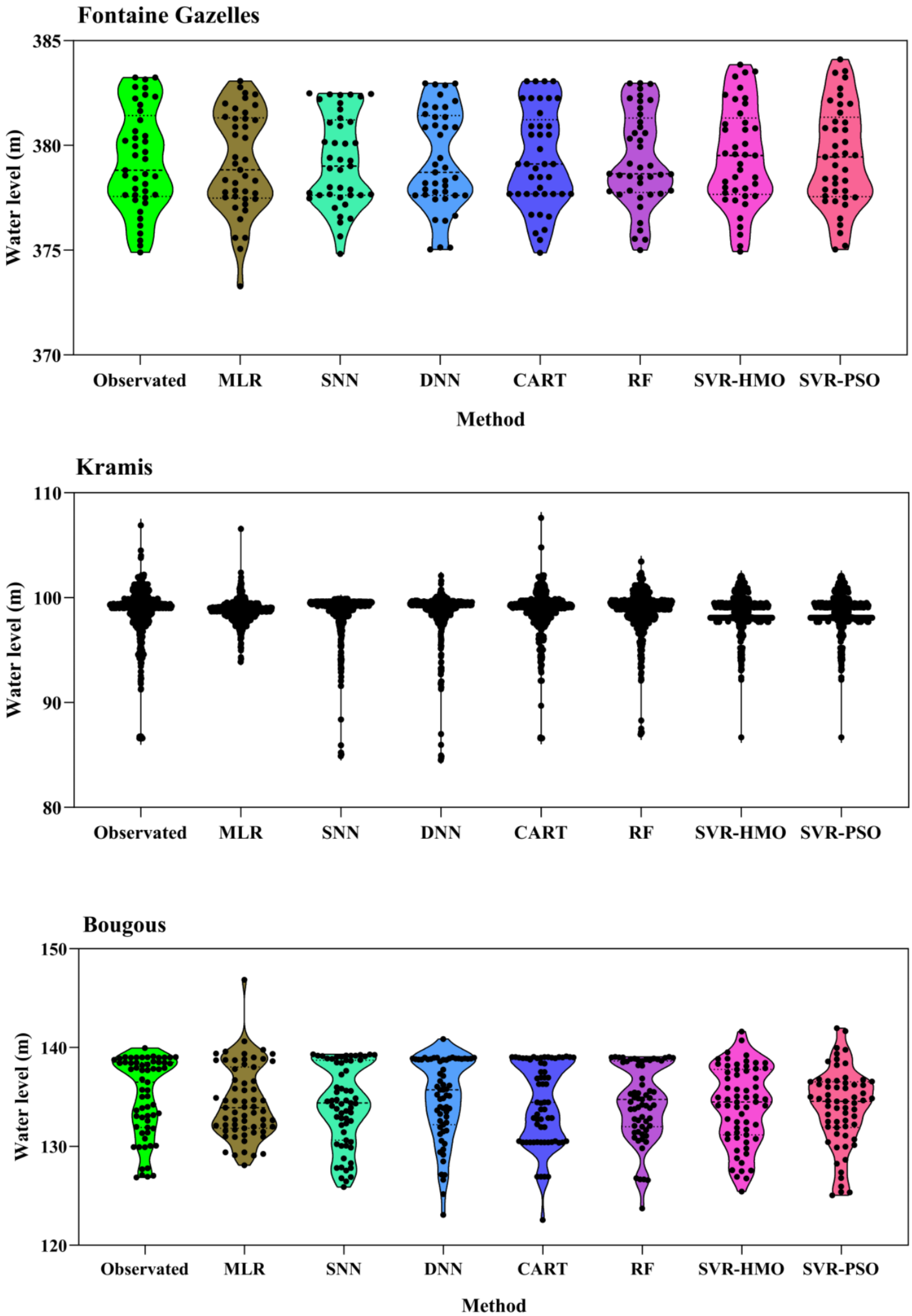

Fig. 3 Violin plots of observed and model-predicted *WLF* for the three dams

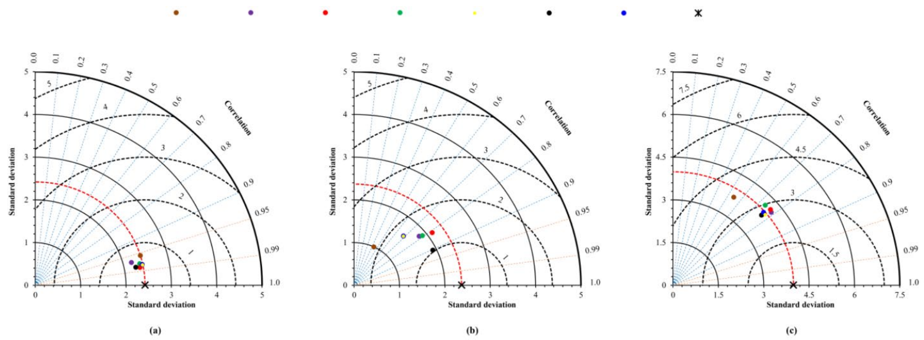

Fig. 4 Taylor diagram of observed *WLF* and model-predicted *WLF* for the three dams: (**a**) the Fontaine Gazelles, (**b**) the Kramis, and (**c**) the Bougous

water level. Leyla et al. (2023) considered *WLF* as predictor to model the piezometers water level. Dayal et al. (2024) considered several inputs including wind, maximum temperature, inflow, outflow, humidity, and visibility to forecast *WLF*. The inputs of this study differ from those of the mentioned studies, which is one of the main reasons for the less-than-perfect results of the Kramis and Bougous dams. The results for the Kramis dam are excellent, indicating that the selected inputs are crucial for achieving accurate predictions.

Ozdemir et al. (2023), in a review study, highlighted that DNN are the most accurate models for *WLF* modeling. However, the results of this study confirm that the RF model is significantly more accurate than the DNN. Even in some cases, CART provides more accurate *WLF* predictions than DNN. The underlying cause of this discrepancy in results warrants analysis from two distinct perspectives. Firstly, RF is an ensemble learning method characterized by a hybrid structure, which enhances its ability to identify trends within functions. Secondly, some DNNs incorporate previous time steps as predictors, significantly influencing the outcomes. For a fair comparison, other models should also be evaluated under the same conditions.

In this study, HMO was introduced to optimally determine the SVR parameters. Mahdavi-Meymand et al. (2024) reviewed the application of evolutionary ML models in hydrology and indicated that, in general, PSO is a strong algorithm for training ML models. The results of the current study showed that HMO is slightly more accurate than PSO. HMO can be an alternative approach for modeling hydrology and environmental problems. Moreover, incorporating other recent algorithms in training ML models may contribute to a better understanding of trends in engineering problems.

It is important to note that DNN and SNN models follow the same ANN approach. DNN typically has multiple hidden layers for feature extraction, while SNN has one hidden layer, leading to a simpler architecture. The DNN requires more time for training; however, a properly structured DNN in this study effectively predicted WLF. It is worth noting that increasing the number of hidden layers can raise the risk of overfitting.

5 Conclusion

Accurate estimation of WLF in embankment dams is of paramount importance due to its direct impact on dam stability, as well as effective water resource and environmental management. This study has demonstrated the effectiveness of various ML models, including RF, SNN, DNN, CART, SVR integrated with HMO and PSO meta-heuristic algorithms, and MLR, in predicting daily reservoir *WLF*. Through the analysis of in-situ data from three embankment dams in Algeria—the Kramis, the Bougous, and the Fontaine Gazelles—this study has revealed notable variations in prediction accuracy across different dams. Particularly, the Fontaine Gazelles dam exhibited superior predictive performance compared to Kramis and Bougous. The main reason for this superiority is the selected input parameters. On the other hand, the P is more important than h. In line with predictions, the ML models outshone MLR in terms of accuracy. In the testing phase for all three dams, the RF model, with an average *RMSE* of 1.244 m, outperformed other ML models. Given that the RF is an ensemble learning method, its superiority is justified. The DNN and CART, with average *RMSE* values of 1.459 m and 1.483 m respectively, were ranked as the next most accurate models. Nonetheless, their accuracy in predicting WLF in the Kramis and Bougous dams is suboptimal ($R^2 < 0.7$). The HMO as an optimal parameter tuner for SVR, outperformed PSO, underscoring the value of employing novel optimization algorithms. The constraints and boundaries of this study are twofold. Firstly, the performance of the developed models may be influenced by various site-specific factors and environmental conditions, which could affect their generalizability to other dam sites. Secondly, the study primarily focuses on embankment dams in Algeria, and the applicability of the models to dams in different regions with distinct hydrological characteristics warrants further investigation. In conclusion, while the proposed AI-based models offer promising capabilities in enhancing reservoir management practices, further research is necessary to validate their performance across a broader range of dam sites and environmental conditions. Additionally, incorporating additional variables and refining model parameters could potentially improve prediction accuracy and expand the scope of application in future studies.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Issam Rehamnia, Modeling part were performed by Amin Mahdavi-Meymand, the first draft of the manuscript was written by Issam Rehamnia and Amin Mahdavi-Meymand. All authors read and approved the final manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data Availability Data will be made available upon reasonable request.

Declarations

Ethical Standards The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Consent to Publish All authors agree to the publication of the manuscript and any supplementary materials.

Consent to Participate All participants provided informed consent to participate in this study.

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

References

Altunkaynak A, Şen Z (2007) Fuzzy logic model of lake water level fluctuations in Lake Van, Turkey. Theoret Appl Climatol 90:227–233
Azad AA, Sokkalingam R, Daud H, Adhikary SK, Khurshid H, Mazlan SNA, Rabbani MBA (2022) Water Level Prediction through Hybrid SARIMA and ANN models based on Time Series Analysis: Red Hills Reservoir Case Study. Sustainability 14(3):1843
Bandini F, Sunding TP, Linde J, Smith O, Jensen IK, Köppl CJ, Butts M, Bauer-Gottwein P (2020) Unmanned aerial system (UAS) observations of water surface elevation in a small stream: comparison of radar altimetry, LIDAR and photogrammetry techniques. Remote Sens Environ 237:111487
Bio A, Gonçalves JA, Magalhães A, Pinheiro J, Bastos L (2022) Combining low-cost sonar and high-precision global navigation satellite system for shallow water bathymetry. Estuar Coasts 45:1000–1011
Breiman L (2001) Random forests. Mach Learn 45(1):5–32
Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and Regression Trees. CRC Press, New York
Buyukyildiz M, Tezel G, Yilmaz V (2014) Estimation of the change in lake water level by artificial intelligence methods. Water Resour Manage 28:4747–4763
Cortes C, Vapnik V (1995) Support vector network. Mach Learn 20:273–297
Dai R, Wang W, Zhang R, Yu L (2024) Multimodal deep learning water level forecasting model for multiscale drought alert in Feiyun River basin. Expert Syst Appl 244:122951
Das M, Ghosh SK, Chowdary VM, Saikrishnaveni A, Sharma RK (2016) A probabilistic nonlinear model for forecasting daily water level in reservoir. Water Resour Manage 30:3107–3122
Dayal A, Bonthu S, Saripalle P, Mohan R (2024) Deep learning for Multi-horizon Water level forecasting in KRS reservoir, India. Results Eng 21:101828
Emami S, Parsa J (2020) Comparative evaluation of imperialist competitive algorithm and artificial neural networks for estimation of reservoirs storage capacity. Appl Water Sci 10(7):1–13
Gao L, Ha D, Ma L, Chen J (2024) The prediction model of water level in front of the check gate of the LSTM neural network based on AIW-CLPSO. J Comb Optim 47(2):1–17
Gong Y, Wang X, Hu BX, Zhou Y, Hao C, Wan L (2016) Groundwater contributions in water-salt balances of the lakes in the Badain Jaran Desert, China. J Arid Land 8:694–706
Huang J, Qin H, Zhang Y, Hou D, Zhu S, Ren P (2023) Short-term prediction method of reservoir downstream water level under complicated hydraulic influence. Water Resour Manage 37(11):4475–4490
Kim D, Han H, Lee H, Kang Y, Wang W, Kim HS (2024) Predicting Flood Water Level using combined hybrid model of rainfall-runoff and AI-Based models. KSCE J Civ Eng 28:1580–1593
Knotters M, Bierkens MF (2000) Physical basis of time series models for water table depths. Water Resour Res 36(1):181–188
Kow PY, Liou JY, Sun W, Chang LC, Chang FJ (2024) Watershed groundwater level multistep ahead forecasts by fusing convolutional-based autoencoder and LSTM models. J Environ Manage 351:119789
Lahouar A, Slama JBH (2015) Day-ahead load forecast using random forest and expert input selection. Energy Conv Manag 103:1040–1051
Leyla H, Nadia S, Bouchrit R (2023) Modeling and predictive analyses related to piezometric level in an earth dam using a back propagation neural network in comparison on non-linear regression. Model Earth Syst Environ 9(1):1169–1180

Liang X, Song C, Liu K, Chen T, Fan C (2023) Reconstructing Centennial-Scale Water Level of large pan-arctic lakes using machine learning methods. J Earth Sci 34:1218–1230
Liu Z, Xu XH, Pan M, Loo CK, Li S (2023) Weighted error-output recurrent echo kernel state network for multi-step water level prediction. Appl Soft Comput 137:110131
Mahdavi-Meymand A, Sulisz W (2023) Development of aggregated random intelligent approach for the modeling of desalination processes. Desalination 567:116990
Mahdavi-Meymand A, Zounemat-Kermani M (2022) Homonuclear molecules optimization (HMO) meta-heuristic algorithm. Knowl Based Syst 258:110032
Mahdavi-Meymand A, Sulisz W, Zounemat-Kermani M (2024) Hybrid and Integrative Evolutionary Machine Learning in Hydrology: a systematic review and Meta-analysis. Arch Comput Methods Eng 31(3):1297–1340
Malekpour MM, Malekpoor H (2022) Reservoir water level forecasting using wavelet support vector regression (WSVR) based on teaching learning-based optimization algorithm (TLBO). Soft Comput 26:8897–8909. https://doi.org/10.1007/s00500-022-07296-1
Malekpour MM, Mohammad Rezapour Tabari M (2020) Implementation of supervised intelligence committee machine method for monthly water level prediction. Arab J Geosci 13(19):1049
Nadiri AA, Naderi K, Khatibi R, Gharekhani M (2019) Modelling groundwater level variations by learning from multiple models using fuzzy logic. Hydrol Sci J 64(2):210–226
Ozdemir S, Yaqub M, Yildirim SO (2023) A systematic literature review on lake water level prediction models. Environ Model Softw 163:105684
Özdoğan-Sarıkoç G, Sarıkoç M, Celik M, Dadaser-Celik F (2022) Reservoir volume forecasting using artificial intelligence-based models: Artificial neural networks, support Vector Regression, and long short-term memory. J Hydrol 616:128766
Rehamnia I, Benlaoukli B, Jamei M, Karbasi M, Malik A (2021) Simulation of seepage flow through embankment dam by using a novel extended Kalman filter based neural network paradigm: case study of Fontaine Gazelles Dam, Algeria. Measurement 176:109219
Rehamnia I, Benlaoukli B, Chouireb M, Prakash I, Amiri M, Pham BT (2023) Estimation of seepage flow using optimized artificial intelligent models. Geotech Geol Eng 41(4):2727–2739
Shang X, Zhao J, Zhang H (2019) Obtaining high-resolution seabed topography and surface details by co-registration of side-scan sonar and multibeam echo sounder images. Remote Sens 11(12):1496
Stefenon SF, Seman LO, Aquino LS, dos Santos Coelho L (2023) Wavelet-Seq2Seq-LSTM with attention for time series forecasting of level of dams in hydroelectric power plants. Energy 274:127350
Su H, Liu H, Wu Q (2015) Prediction of water depth from multispectral satellite imagery—the regression kriging alternative. IEEE Geosci Remote Sens Lett 12(12):2511–2515
Sun Y, Babovic V, Chan ES (2010) Multi-step-ahead model error prediction using time-delay neural networks combined with chaos theory. J Hydrol 395(1–2):109–116
Ünes F, Demirci M, Kisi O (2015) Prediction of Millers Ferry Dam Reservoir Level in USA using Artificial neural network. RPeriodica Polytech Civil Eng 59(3):309–318
Wang B, Wang B, Wu W, Xi C, Wang J (2020) Sea-water-level prediction via combined wavelet decomposition, neuro-fuzzy and neural networks using SLA and wind information. Acta Oceanol Sin 39:157–167
Wen J, Han PF, Zhou Z, Wang XS (2019) Lake level dynamics exploration using deep learning, artificial neural network, and multiple linear regression techniques. Environ Earth Sci 78(6):222
Yadav B, Ch S, Mathur S, Adamowski J (2017) Assessing the suitability of extreme learning machines (ELM) for groundwater level prediction. J Water Land Dev 32(1):103
Yuan Z, Liu J, Liu Y, Zhang Q, Li Y, Li Z (2022) A two-stage modelling method for multi-station daily water level prediction. Environ Model Softw 156:105468
Zhang J, Wang XS, Hu XN, Lu HT, Ma Z (2017) Research on the recharge of the lakes in the Badain Jaran Desert: Simulation study in the Sumu Fijaran Lakes area. J Lake Sci 29(2):467–479
Ziggah YY, Issaka Y, Laari PB (2022) Evaluation of different artificial intelligent methods for predicting dam piezometric water level. Model Earth Syst Environ 8(2):2715–2731

Authors and Affiliations

Issam Rehamnia[1] · **Amin Mahdavi-Meymand**[2]

✉ Amin Mahdavi-Meymand
a.mahdavi@ibwpan.gda.pl

Issam Rehamnia
issam.rehaminia@univ-annaba.dz

1 Department of Hydraulic, Faculty of Technology, Badji-Mokhtar Annaba University, P.O. Box 12, Annaba 23000, Algeria

2 Institute of Hydro-Engineering, Polish Academy of Sciences, Gdańsk, Poland