AQUA
Water Infrastructure, Ecosystems and Society

 AQUA — Water Infrastructure, Ecosystems and Society Vol 72 No 7, 1228 doi: 10.2166/aqua.2023.042

A review of artificial intelligence methods for predicting gravity dam seepage, challenges and way-out

Priyanka Ashok Garsole [a,*], Shantini Bokil[a], Vijendra Kumar[a], Arunabh Pandey[a] and Niraj S. Topare[b]

[a] Department of Civil Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India
[b] Department of Chemical Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India
*Corresponding author. E-mail: priyankagarsole24@gmail.com

PAG, 0000-0002-0589-968X

ABSTRACT

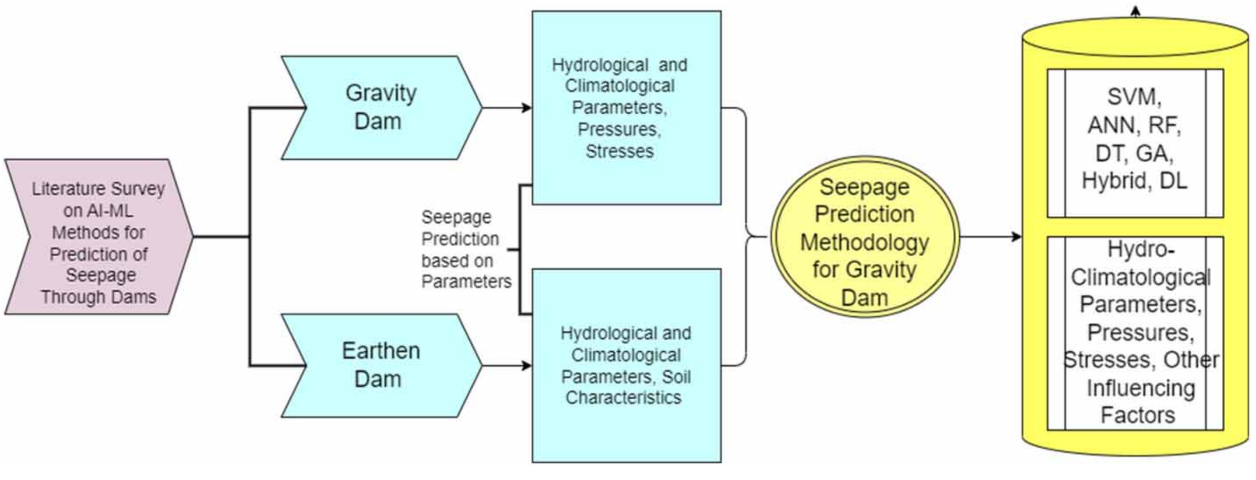
Seepage is the phenomenon of water infiltrating through a gravity dam's foundation, causing erosion and weakening the dam's construction over time. If not properly managed, this can eventually lead to the dam's catastrophic failure, posing a significant danger to public safety and the environment. As a result, precise seepage prediction in gravity dams is essential for ensuring their safety and stability. This review paper looks at the use of artificial intelligence (AI) techniques for predicting seepage in gravity dams, as well as the challenges and possible solutions. The paper identifies and suggests potential solutions to the challenges connected with using AI for seepage prediction, such as data quality and model interpretability. The paper also covers future research paths, such as the creation of advanced machine learning algorithms and the improvement of data collection and processing. Overall, this review gives insight on the current state of the art in using AI to predict gravity dam seepage and recommends methods to improve the accuracy and reliability of such models.

Key words: artificial intelligence, hydro-climatological parameters, pressures, seepage predictions, stresses

HIGHLIGHTS

- AI methods for predicting gravity dam seepage reviewed, with challenges and solutions.
- The review provides an overview of using AI for seepage prediction in gravity dams.
- AI challenges addressed with suggested solutions for improved seepage prediction.
- Standardizing data collection and improving quality reduces errors in prediction models.
- Insights for dam safety practitioners, improving seepage.

GRAPHICAL ABSTRACT



ABBREVIATIONS

AFI	adaptive fuzzy identification
ACR	anthropomorphic causal reasoning
AI	artificial intelligence
ANN	artificial neural network
ABC	artificial bee colony
ANFIS	adaptive-network-based fuzzy inference system
BEM	boundary element method
CART	classification and regression tree
CB	CatBoost
CFD	computational fluid dynamics
C LSTM	convolutional long short-term memory
CNN	convolutional neural network
CVC	conventional vibrated concrete
DBM	deep Boltzmann machine
DBN	deep belief network
DNN	deep neural network
DL	deep learning
DST	decision support tool
DT	decision tree
FDM	finite difference method
FEM	finite element method
FFA	feed forward algorithms
FFNN	feed forward neural network
FFANN	feed forward artificial neural network
FFMLP	feed forward multi-layer perceptron
GAN	generative adversarial networks
GEP	genetic expression programming
GEVR	grout-enriched vibrated RCC
GP	genetic programming
GPR	Gaussian process regression
HTRT	hydrostatic thermal rainfall time
IBL	instance-based learning
MAE	mean absolute error
MARS	multivariate adaptive regression splines
MIC	maximal information coefficient
ML	machine learning
MLP	multi-layer perceptron
MSE	mean squared error
R	adjusted R (squared error)
RBF	radial basis function
RBM	restricted Boltzmann machine
RCC	reinforced cement concrete
ReLU	rectified linear unit
RMSE	root mean squared error
RNN	recurrent neural network
RVM	relevance vector machine
SEEP/W	finite element method software
SVM	support vector machine

1. INTRODUCTION

Gravity dams are large hydraulic structures that are usually built across rivers or valleys to create reservoirs or generate hydroelectric power. They are built to withstand the weight of the water they hold as well as the forces produced by water pressure, wind, and seismic (Bernier *et al.* 2016). The potential of seepage, or the flow of water through the dam's structure or foundation, is one of the main concerns with gravity dams. Seepage can cause several problems that jeopardize the dam's integrity and safety (Al-Juboori & Datta 2019). It causes erosion, which weakens the dam's structure and increases the risk of collapse (Messerklinger 2014). Seepage through gravity dams can lead to erosion, uplift pressure, piping, and saturation, which compromise the dam's stability and safety. Managing seepage is crucial, requiring monitoring and analysis, and implementing seepage control measures (Nourani *et al.* 2021). During the construction of a gravity dam, several measures can be taken to reduce direct uplift pressure, such as constructing cut-off walls under the upstream face, drainage channels between the dam and foundation, and pressure grouting the foundation (Wang *et al.* 2019). Despite these safety measures, seepage can still occur in gravity dams, although generally less than in earthen dams. However, there are cases where significant seepage occurs, which can affect the yearly storage of water and hydropower electricity generation. Figure 1 depicts a real-time picture of serious seepage around the gallery. These locations give engineers important details about the dam's behavior well as the position and level of seepage. Making choices about the dam's maintenance and repairs can be done so that its long-term stability and safety are ensured (Hu *et al.* 2018).

As a result, predicting gravity dam seepage behavior is essential for ensuring their safety and reliability (Kang *et al.* 2017). Predictions of seepage can be used to identify possible failure mechanisms, optimize dam design and construction, and guide maintenance and repair decisions (Wahl 2004). Accurate and dependable seepage predictions can also reduce the likelihood of catastrophic failures, which can have severe economic, environmental, and social consequences (Sevieri *et al.* 2021). Therefore, a precise seepage prediction in gravity dams is essential for ensuring both their longevity and safety (Al-Juboori & Datta 2019) Here are a few typical seepage prediction techniques for gravity dams. *Empirical techniques*: Based on the reservoir water level, the size of the dam, and the characteristics of the soil, empirical techniques can be used to determine the seepage flow through the dam body (Su *et al.* 2016b). The Manning formula, the Hazen formula, and the Bligh formula are a few examples of these techniques. These equations are used in the preliminary planning stage to estimate the seepage flow through the data (Li & Wang 2019). Given that the dam is fully operational during peak season, it is important to conduct manual measurements and apply empirical methodologies to estimate the seepage occurring through the dam. *Analytical techniques*: Analytical techniques compute seepage through the dam's body and foundation using mathematical formulae. The Dupuit–Forchheimer method (Castro-Orgaz & Giráldez 2012), which assumes steady-state flow through the dam, and the finite element method (FEM) (Athani *et al.* 2015), which simulates uneven seepage through the dam and substructure, are two examples of these techniques. Analytical methods produce a detailed examination of a scenario in which an instant judgment cannot be made unless and until the whole analysis is performed.

Numerical techniques: Numerical techniques utilizing computer simulations for example (Seth & Pandey 2009), predict seepage in gravity dams. With these techniques, the equations governing the passage of water through the dam and foundation are broken up into smaller components and solved (Hu & Ma 2016). The non-uniform seepage through the foundation and dam, the effects of construction joints, and the impact of changing reservoir levels are simulated (Zhang *et al.* 2021b). The numerical methods deal with continuous data input to models and runtime outputs that may be used to make quick choices. *Field methods*: In the field, tools like piezometers, flow meters, and pressure sensors (Russoniello & Michael 2015) are used to measure the seepage flow through the foundation and dam. The data from these observations are crucial for validating and adjusting seepage prediction models (Xu *et al.* 2020). Previously, empirical and analytical methodologies were used to forecast seepage through dams. These methods are more time-consuming and labor-intensive than numerical methods, because numerical methods, such as computer simulations for seepage prediction, can be effectively

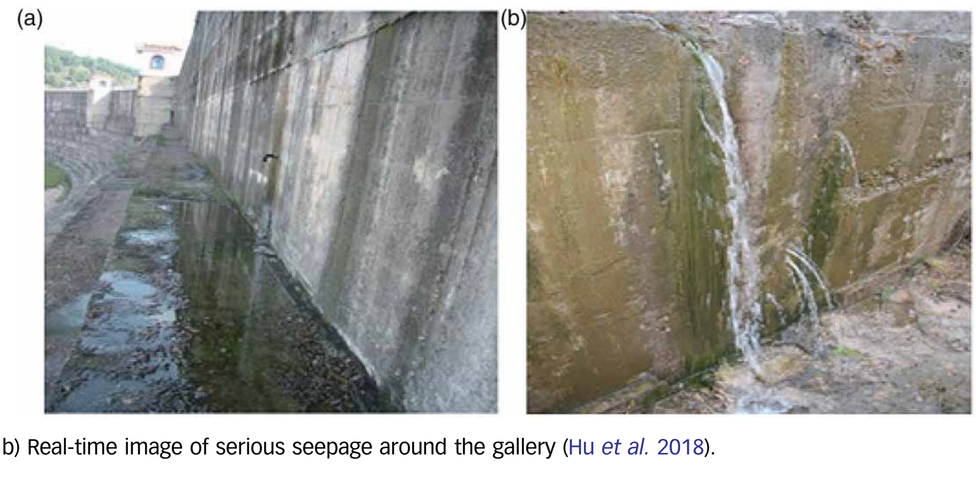


Figure 1 | (a and b) Real-time image of serious seepage around the gallery (Hu *et al.* 2018).

used for the system (Sermet *et al.* 2020), where seepage prediction and control measures can be effectively implemented as part of the automation process using various sensor-mounted instruments and a strong decision support system can be built.

These approaches, however, have drawbacks and are ineffective when dealing with complicated issues (Nourani *et al.* 2021). As a result, applying artificial intelligence (AI) models to address these issues has gained more attention. AI models like neural networks, decision trees (DTs), fuzzy logic, and support vector machines (SVMs) are trained on large amounts of data to identify patterns and make predictions (Zhang *et al.* 2021c). The non-linear relationships between input and output factors, which are typical in dam seepage and pore water pressure issues, can be handled by these models (Wang *et al.* 2018). They can also take into consideration complexities and unknowns, like the variability of soil characteristics.

One of the main advantages of AI models is their ability to learn from data and improve over time (Sharghi *et al.* 2018). For example, neural networks are trained on historical data from previous dam projects to predict the behavior of new dams (Kumar & Yadav 2021). As more data is collected and added to the model, its predictions can become more accurate and reliable. Another advantage of the AI model is its ability to analyze and interpret large amounts of data quickly. It saves time and resources compared to traditional numerical and analytical methods, which can be time-consuming and require significant computational power (Lai *et al.* 2021).

Hybrid methods, which combine numerical methods, analytical methods, and AI models, are also becoming more prevalent in solving dam seepage and pore water pressure problems. These methods are leveraging the strengths of each approach to provide more accurate and efficient solutions (Belmokre *et al.* 2019). For example, a hybrid method might use a numerical method to simulate the flow of water through a dam and predict pore water pressures at different locations (Rehamnia *et al.* 2021). Then, an AI model could be used to analyze the output data and identify patterns that could be used to optimize the design of the dam or predict future behavior.

1.1. Need of the review

Predicting gravity dam seepage is crucial for several reasons. Firstly, seepage can compromise the safety of gravity dams by increasing uplift pressure and reducing the stability of the dam structure. By identifying potential failure points, engineers can take preventative measures to ensure dam safety. Secondly, predicting seepage can help identify cost-effective solutions to reduce seepage and maintenance costs. Thirdly, seepage can result in the loss of water and harm the environment. A study predicting seepage can help minimize the environmental impact by identifying measures to reduce seepage. Fourthly, with advances in technology, new methods are being developed to predict seepage through gravity dams. Lastly, a study predicting seepage can help optimize the design of gravity dams to minimize seepage and ensure long-term stability. Excessive seepage through earthen dams and gravity dams can lead to significant water loss from the reservoir and increase the uplift pressure on the dam, which can fail. If the tension in the dam increases, it can cause cracks to develop, typically at the toe of the dam, resulting in ineffective base width. Early prediction of seepage and effective control measures can help reduce water loss.

1.2. Scope of the review

This review has analyzed past studies on seepage analysis and predictions for gravity dams. While extensive research has already been performed on seepage through earthen dams, there is a need to establish a methodology for predicting seepage through gravity dams based on the obtained methodology and other prediction analyses. To achieve this, it is important to study seepage patterns through gravity dams, perform FEM analysis of seepage flow through dams, and understand the amount of seepage occurring through specific sections of gravity dams. These efforts will aid in the development of effective measures for seepage control in gravity dams. While reviewing the seepage analysis studies more emphasis is given to utilizing AI technologies for prediction and analysis. Numerous studies have utilized AI algorithms for seepage prediction and have found that these methods produce effective results.

1.3. Research objective

The research objective of this review paper is on predicting seepage through gravity dams to provide a comprehensive overview of the current state-of-the-art methods and techniques for seepage prediction, as well as identify future research needs. Additionally, it is required to evaluate the effectiveness of existing methods and identify any gaps or limitations in current research.

1.4. Structure of the paper

The structure of this review paper is as follows: Section 2 describes a survey of the literature. Section 3 discusses the current status of AI methods for seepage prediction and applications. Section 4 focuses on the challenges and way forward. Section 5 discusses potential opportunities and future directions. Finally, in Section 6, we conclude and summarize the study's main findings.

2. LITERATURE REVIEW

The problem of seepage through gravity dams is grabbing attention nowadays addressing the issue of seepage through earthen dams. Seepage through earthen dams if left uncontrolled and untreated, leads to piping through the dam body or foundation, undermining and sloughing off the downstream toe (Garg 2005). Failures in earthen dams may result in the collapse of the entire structure. As a result, several case studies have demonstrated that earthen dams experience seepage. Several strategies for controlling seepage have been proposed by researchers for solving these cases, including early predictions of seepage and various controls for seepage. There have been many cases of earthen dams where an excessive amount of seepage has occurred and addressed at earlier stages to apply several seepage control measures on it such as the provision of impervious core, filter layers, adequate drainage system, seepage blankets, slope protection, and monitoring and maintenance. Refaiy *et al.* (2021) demonstrated variations in the geometry of the earthen dams and performed sensitivity analysis on the earth dam to evaluate the drain's most effective geometric parameters. For downstream drainage, the most effective parameter is found to be the length of the drain downstream. An increase in the length of the drain downstream increases the seepage discharge and reduces the pore water pressure, which is another parameter affecting the seepage through the dam. These control measures are typically suggested or provided after case-specific studies. Omofunmi *et al.* (2017) presented an approach to various seepage control measures of earthen dams for all the seepage conditions for homogeneous, zoned, and diaphragm-type earthen embankments; resulting in the use of various dam safety instruments use such as pore water pressure, settlement gauges, horizontal movement devices, and seismic conductivity measures.

Talking about the overall scenario of the dams in general, there are some conditions or factors responsible for seepage occurrences through the dam irrespective of the dam material. Among these hydrological parameters, climatological parameters, geometric features of the dam, and some other factors such as pore water pressure, uplift pressure, dam structural deformations, and other developed stresses and pressures have been enlisted to be major influencing parameters for seepage occurrences through dams. Many researchers have demonstrated the models showing the effect of a single or two or more above-stated parameters on dam seepage. These case studies are included in the literature.

The hydro-climatological parameters (Omar *et al.* 2019) were studied to find out the effect of those parameters on dam seepage. Ishfaque *et al.* (2022b) presented a study on seepage occurring through the earth and rockfill dam that was thoroughly investigated using hydro-climatological factors such as temperature, rainfall, water input, sediment inflow, and reservoir level using data-driven models such as CatBoost (CB), random forest (RF), SVM, and artificial neural network (ANN). The accuracy of seepage prediction is influenced by other parameters in addition to the hydro-climatological parameters and to determine the correlation between input and output parameters and obtain the best-fit model, Shapley Additive Explanations (SHAP) method was used to comprehend the impact of each parameter on the output. According to SHAP summary plots, the reservoir level was reported as the most important parameter for determining the average seepage, and a direct association between reservoir level and average seepage was detected. Ishfaque *et al.* (2022a) implemented several deep learning methods for seepage prediction through clay and rock dam. Seepage prediction was carried out using several hydro-climatological parameters, geophysical, and engineering characteristics for peak-to-peak inflows. Various time-series models were built on data such as recurrence neural network (RNN) and long short-term memory (LSTM). Structural health monitoring in the form of seepage prediction projected damage occurrences with minimal MSE and minimal losses and high accuracy of prediction. Lee *et al.* (2018) studied the effect of rainfall on earthen dam seepage, as rainfall is one of the major climatological parameters affecting the hydrological cycle and ultimately the seepage. The seepage study was conducted on a steady-state rockfill dam with the application of a digital filter that helped to measure the rainfall-induced infiltration resulting in seepage. The seepage through the rockfill dam was effectively reduced using the digital filter with ease in the seepage measurement process. Another study (Sur *et al.* 1999) focused on understanding the effect of hydrological parameters such as rainfall, and evaporation in the water loss process (i.e. seepage) through small water harvesting structures (WHS) in northern India. Records of harvested water show that the major mode for loss of water occurred through evaporation and seepage through the dam (i.e. around 61–86%) and that affected the annual production in farms due to loss in quantity of the planned irrigation water.

Several researchers have addressed another contributing component, variable water level, and have extended their study in the topic to execute AI models, traditional numerical models, or image processing models (Omar *et al.* 2022a). The research conducted by Wang *et al.* (2018) presents the monitoring models for the base flow effect and daily variation of dam seepage elements (i.e. reading of the installed device piezometer) to validate the impact of daily variations of water levels in the reservoir. Sensitivity analysis performed based on support vector regression (SVR) gives optimized lag time and influencing period and enhanced the model performance by jittering and ensembling by around 30%. Tayfur *et al.* (2005) proposed an approach to seepage and water level predictions for seepage estimation using FEM and ANN. Adequate performance of ANN was observed in comparison to FEM while prediction of water level in the earthen dam and prediction of seepage. ANN builds site-specific models using a user-friendly and easier approach to model building. Zhang *et al.* (2021a) studied the flow through the dam body to ascertain the permeability coefficient of critical dam sites. A time series model was created using the dataset to estimate seepage through the dam body based on characteristics such as different combinations of permeability coefficient and water levels (Omar *et al.* 2020) upstream and downstream. Generated data is used to study and forecast the permeability coefficient of significant portions of the dam using AI-ML algorithms and numerical models. More effectiveness and accuracy were found in the performance of models which were built by using AI techniques as compared to the other conventional methods, as various AI models have already demonstrated acceptable performances while prediction algorithms were run on the datasets.

Pore water pressure plays an important role in seepage through soil or rock materials (Singh *et al.* 2023a). Pore water pressure is the pressure exerted by the water within the pore spaces of soil or rock materials. When water flows through soil or rock materials, it generates a gradient in pore water pressure that can affect the stability of the material. An approach for the determination of the pore water pressure of an earthen dam was proposed (Kashef 1965) and the analysis of vertical dam core and earth dams or sloping embankments was performed under steady-state conditions. El Bilali *et al.* (2022) predicted daily pore pressure using ANN, MLR, SVR, AdaBoost, and non-linear autoregressive exogenous (NARX) algorithms proposed to combine the response delay of the dam to hydraulic load. The performance of these models was compared with conventional hydrostatic seasonal time (HST). Beiranvand & Komasi (2021) proposed a cut-off wall method to control seepage. The numerical simulations were performed for full reservoir condition (FRC), and a drop in the pore water pressure in the downstream part was observed. The data of instrumentation and numerical fitted using the multivariate regression analysis and a good coefficient of determination was obtained. Liu *et al.* (2023) performed a dual analysis of lag sensitivity of influencing factors based on MIC optimization RF algorithm presented for seepage predictions, whereas study findings revealed that the water pressure component (Omar *et al.* 2021) is the primary influencing element of seepage, the rainfall component comes in second. The sensitivity of the factors was quantitatively analyzed by %IncMSE and IncNodePurity, these two parameters of the RF algorithm. This study shows that the MIC-RF model has high accuracy, high robustness, and strong adaptability while predicting the seepage of the dam. The result shows MIC-RF models can be used as a seepage safety monitoring model. Wei *et al.* (2021) validated the performance of algorithms for pore-water pressure time-series prediction after working on machine learning (ML) by applying the ANN algorithms and concluded with the performance of double-layer gated recurrent unit (GRU) is time-cost-consuming and comes with little accuracy improvements over the performance of single-layer GRU model. Beiranvand & Rajaee (2022) examined the pore water pressure focused on AI-based single and hybrid models in the prediction of seepage and pore water pressure of dams. The time and cost found to be reduced by using various AI methods like ANN, SVR, RF, and feed forward neural network (FFNN) have been used. Overall the adaptive-network-based fuzzy inference system (ANFIS), SVM, classification and regression tree (CART), instance-based learning (IBL), artificial bee colony (ABC), teaching-learning based optimization (TLBO), genetic programming (GP), and gene-expression programming (GEP) have been used. It has been observed that among the studied hybrid models, 81.25% of models have used ANN and 31.25% of models have used genetic algorithm (GA).

Several authors have performed the analysis of seepage through gravity dams. Zhang *et al.* (2020) proposed a combined approach to model development for the prediction of seepage through a gravity dam. Back propagation neural network (BPNN) was combined with GAs for improving the accuracy of seepage prediction. The study resulted in more accurate seepage projections for the near future. Another approach of seepage prediction based on dam deformation analysis performed on an Algerian roller-compacted gravity dam has obtained adequate results. Belmokre *et al.* (2019) examined the effect of dam deformation on seepage and thoroughly analyzed the seepage and dam deformation with statistical models, SVM and RF. The performance of the models has been evaluated by using the RMSE and MAE criteria. The data recorded using flowmeters from 2011 to 2018 of water levels and water temperature was fed to algorithms. For the prediction of seepage flow rates, these models were applied to the seepage data. Prediction concluded that the RF regression model performs better compared to the SVM. Dam deformation was found to be another influencing factor for dam seepage predictions.

Material properties play an important role in the stability of the structure. Similarly, material decomposition has a significant effect on dam seepage. Gravity dams are typically constructed using concrete, which is a durable and long-lasting material. However, over time, concrete can degrade and deteriorate due to various factors, such as exposure to weathering, freeze–thaw cycles, and chemical reactions. Zhang *et al.* (2021b) conducted a study to determine the primary reason for seepage and the evolution of its features while considering the effect of calcium leaching. For studying the hydro-chemo behavior of materials made of cement, a novel mathematical model was suggested. It was concluded that a 100-day leaching period would cause an increase in leakage amounts of the dam body and foundation of 48 and 17 times, respectively, and a 40.8% rise in uplift pressure. In an aquatic environment, the breakdown of calcium compounds in cement-based materials results in seepage, which eventually raises the uplift pressure along the dam support and malfunctions the grout curtain. Yue *et al.* (2022) investigated the decline in shear strength that led to a decrease in dam stability because of the calcium-leaching impact. The dam's factor of safety (FOS) is greatly impacted by the hydraulic conductivity of the grout curtain and rock base. Cancienne *et al.* (2008) developed the bank stability and toe erosion model (BSTEM) (Omar *et al.* 2022b) to ascertain the relationship between bank shear strength, soil pore water pressure, bank angle, root reinforcements, and undercutting of seepage. As per the BSTEM, seepage undercutting can be avoided when the soil-water pore system is entirely saturated. Salem (2019) conducted seepage computations and analysis of slope stability by using the GEOSTUDIO program in terms of SEEP/W and SLOPE/W. To calculate seepage, variables like hydraulic gradient, soil flexibility, particle size, and capillary tension were analyzed. The effect of core permeability, core width, core base thickness, and core penetration on seepage, pressure head, and exit-gradient, upstream and downstream slope stabilities on seepage through earthen embankments with or without core was studied, and it was concluded that an increase in side slope (H:V), increase in downstream slope stability, decreasing the permeability of the inner core, increase in core thickness, and core penetration. Seepage discharge through a homogeneous earthen dam with a core has been investigated by Abdul Jabbar Jamel (2018) using SEEP/W and ANN models concludes the estimated seepage discharge by using ANN with a single hidden layer and by using SEEP/W model was nearly similar, which shows better efficiency of models. Some authors have performed the CFD analysis to understand the seepage behavior through the dam wall.

Many authors have targeted using AI-ML techniques in the field of water resources engineering as these techniques have proven acceptable performances and prediction accuracy is also obtained when considering all the relevant parameters in the phase of model building. The applications of deep learning (DL) in hydrology and water resources for the prediction of various hydro-climatological parameters were discussed by Sit *et al.* (2020) where DNN was used to predict the temperature and geothermal springs and wells using hydrogeochemical data, including chemical concentrations. Most of the automated forecasting models have been developed on water level predictions and stream flow. Early prediction systems help us to create awareness about upcoming hazards. Such an intelligent system was proposed (Sermet & Demir 2018) for knowledge generation and communication about flooding. Presented flood AI designed for flood alerts or intimations on communications channels including web-based systems, agent-based chatbots, and smartphone applications automated workflows, smart/automated home devices for the improvements in societal preparedness for extreme events. Also, the gaming framework examined by Sermet *et al.* (2020) for decision support on hydrological hazards developed the web-based decision support system tool (DST) for shared multi-hazard analysis including floods, drought, soil erosion, water pollution, etc.

Seepage occurring through the gravity dam is also a severe problem. In some cases, dams having a functioning age of more than 100 years can fail, and seepage can be enlisted as one of the major reasons. Computational methods such as CFD and analysis using FEM can be used to find the behavior of seepage, and seepage flow patterns through the gravity dam wall. Cheng *et al.* (2018) simulated the complex field of gravity dam foundation using the computational fluid dynamics (CFD) approach; where the three-dimensional (3D) air–water two-phase seepage mathematical model coupled with the volume of fluid (VOF) method was used. Grouting curtains and drain holes as a seepage control measure played an appreciable role in resisting and draining the seepage water, respectively. By considering the hydraulic gradient, as the opening degree of drainage holes decreases, the hydraulic gradient of grout curtains decreases and the water head near the dam base and uplift pressure on the dam foundation increases. However, the CFD analysis is more critical and time-consuming while considering the seepage patterns and flow of water through the dam body, as the width of the dam is observed to be more at a section where the seepage analysis is required to be performed. Authors have applied the AI-ML methods for seepage predictions through gravity dams. Salazar *et al.* (2017) carried out dam safety analysis by focusing on seepage and uplift, using the hydrostatic-season-time (HST) method and built various predictive models ANN and SVM using air temperature, reservoir levels, the temperature of the dam body, stresses, and displacements. This research proposed the criteria for the selection of input parameters, and training–testing split mechanism. The study concluded the highly valuable performance of ML models SVM, KNN, and boosted regression tree (BRT) compared to other applied methods for dam safety.

This article has reviewed the analysis and prediction of seepage through earthen and gravity dams. Moreover, the prediction cases for earthen dams and other seepage cases have been referred to. The approach used for the prediction of seepage through earthen dams was found very useful to design the framework for proposing the outline of research for gravity dams as the hydro-climatological parameters, pore pressures, and dam deformations are observed to be some of the influencing parameters for gravity dam seepage predictions.

Here, this review focuses on the application of AI-ML algorithms for seepage predictions and other cases of seepage, leakage, water loss, etc., similar problems where the base remains the same. Literature has been reviewed and the overview of research has been given on challenges faced for the prediction and analysis of seepage. Many authors have presented their study on seepage prediction through earthen dams and have successfully implemented soft computing techniques such as ANN, RF, MLP, SVM, GP, FFNN, GRU, LSTM, DT, EKF-ANN, ANFIS, and ARIMA models. For validation of the performances of AI-ML algorithms, various AI-ML algorithms were applied to problems in water resources engineering. ML algorithms have some limitations for predictions as there may have some manual errors in data collection and can lead to misleading information which can lead to dam failure. Some algorithms have been performed explicitly for the prediction and analysis of seepage through earthen dams. Several AI-ML algorithms have been successfully implemented for various water resources engineering applications.

There are proven acceptable performances of Random Forest (RF) algorithms in around 20–30% of the studies reviewed. The performance of RF algorithms was found better as compared to other ML algorithms, the study concludes RF algorithms can be widely used for the prediction of seepage. The acceptable performance of the DL algorithms was observed in the referred literature. The DL algorithms can be effectively implemented for seepage predictions. The existing literature says that overall many AI techniques give acceptable results. Depending on the data availability, catchment characteristics (Nagure & Shahapure 2021), and varieties observed in the data the corresponding AI-ML technique can be further recommended. Several ensemble techniques (Nourani *et al.* 2018; Abba *et al.* 2019) were used for predictions, and have ascertained performances as compared to the use of single AI models. The referred literature is presented in the table in the section below and a graphical representation is given for a better understanding of the frequently used and successfully implemented technologies for seepage prediction.

3. AI METHODS FOR SEEPAGE PREDICATION AND APPLICATIONS

A key component of seepage prediction in gravity dams is seepage modeling. Developing a seepage prediction model for gravity dams involves several steps. First, the problem must be defined and relevant data gathered. Next, a numerical model must be developed using traditional numerical modeling techniques or a data-driven approach (Chen *et al.* 2020) such as ML algorithms (Sharma *et al.* 2022). The model must be calibrated and validated using field measurements to ensure accuracy. Once validated, the model can be used to make predictions and evaluate different scenarios. Finally, the model should be continually refined and updated based on new data and observations to improve its predictive capabilities (Wang *et al.* 2017). Careful consideration and calibration are essential for accurate seepage prediction in gravity dams.

Creating traditional numerical modeling techniques to replicate the water flow through the dam and its foundation is known as seepage modeling. For gravity dams, some of the frequently used traditionally used seepage modeling methods include:

(a) A numerical technique called the FEM can be used to simulate seepage in a gravity dam. When using FEM, the domain is divided into smaller components, and the water movement through each component is then solved (Salem 2019).
(b) Finite difference method (FDM): FDM is yet another numerical technique for modeling seepage in gravity dams. FDM solves the movement of water between neighboring nodes by dividing the domain into a grid of nodes (Fu *et al.* 2016).

(c) Boundary element method (BEM): BEM is a numerical technique that can be used to model seepage through domain boundaries. BEM solves for the flow of water at the domain's boundary rather than within the domain itself (Demetracopoulos & Hadjitheodorou 1996).
(d) Analytical solutions: Analytical solutions are mathematical equations that can be used to model seepage in gravity dams. Analytical solutions can provide a simplified approach to seepage modeling, but may not capture the full complexity of the system (Castro-Orgaz & Giráldez 2012).

AI has grown in popularity among academics as a way to save time and resources as compared to traditional numerical and analytical approaches, which may be time-consuming and need substantial computer capacity. AI technologies are utilized to increase the accuracy and efficiency of seepage modeling, and they may be combined with standard numerical modeling techniques to give more robust and accurate forecasts (Fatima & Pasha 2017). One of the most important uses of AI in seepage modeling is the creation of a data-driven model (Sharghi *et al.* 2019). ML algorithms are used in data-driven models to evaluate massive volumes of data and uncover patterns, connections, and variables (Xu & Liang 2021). Based on the input data, these models may subsequently be used to forecast seepage behavior.

3.1. Data-driven method for seepage predication

ML algorithms are used to uncover patterns and correlations in data in data-driven solutions for seepage prediction. Based on historical data such as water levels, rainfall, soil qualities, and other pertinent elements, these approaches forecast dam behavior. The data-driven strategy for forecasting seepage entails gathering and analyzing important data, selecting key components and an appropriate ML algorithm, training the model using historical data, and evaluating its accuracy and generalization. A flow diagram of this procedure is shown in Figure 2. After being trained, the model may make predictions on new data. The methodology offers advantages over traditional methods, but its effectiveness is based on the quality and amount of training data.

The exact stages required in using data-driven methodologies for seepage prediction are as follows:

(a) Data collection: Relevant dam and environment data must be collected and prepared for the study. Data on dam geometry, soil properties, water levels, rainfall, and other pertinent variables can be included. Enough care should be taken while considering these variables, and analyzing their relative importance as suggested by Seth *et al.* (2018).
(b) Data pre-processing: The acquired data must be cleaned and processed to remove any noise, missing values, or other issues that might impair the accuracy of the model.
(c) Feature engineering: Recognize and choose relevant characteristics or inputs that are likely to impact seepage. These traits might be obtained from data or expert knowledge.
(d) Model selection: An appropriate ML technique must be chosen based on the nature of the issue and the available data. Neural networks, SVMs, DTs, and RFs are some of the most often utilized seepage prediction algorithms.

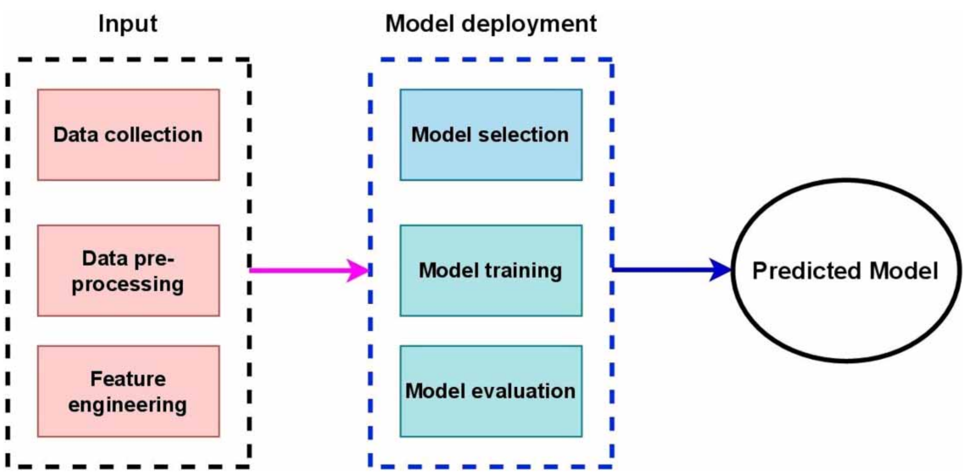


Figure 2 | Flow diagram of data-driven seepage prediction methods.

(e) Model training: Using previous data, the chosen model is trained, and its parameters are fine-tuned to maximize its performance.
(f) Model evaluation: The trained model is assessed for accuracy and generalization using validation data.
(g) Model deployment: After it has been validated, the model is utilized to make predictions on new data.

3.2. Classification of data-driven method and application

For gravity dam seepage prediction, there are two types of data-driven methods: supervised learning and unsupervised learning. Supervised learning is the process of using labeled data to train a model to predict a target variable. A model must be developed using historical data on the dam's characteristics, hydrological data, and seepage rates to predict seepage rates from gravity dams. ANNs (Zhang *et al.* 2021a), SVMs (Su *et al.* 2016a), RFs (Li *et al.* 2021), and extreme learning machines (ELMs) (Lei *et al.* 2022) are some of the most often utilized supervised learning approaches for this purpose. But predicting gravity dam seepage also necessitates finding connections and patterns in unlabeled data. Unsupervised learning techniques like principal component analysis (PCA) (Xu & Wu 2022), k-means clustering (Song *et al.* 2021), and self-organizing map (SOM) are useful in this situation. These techniques can assist in identifying patterns and structure in dam physical and hydrological data that might be important for predicting seepage.

There are several techniques to anticipate seepage in gravity dams using the data-driven approach. One method is to use ML algorithms to forecast seepage rates based on input data such as dam geometry, soil properties, and hydrological data. Another method is to use ML algorithms to identify critical variables that are essential for seepage prediction, which can then be used to guide traditional seepage prediction methods. Furthermore, data-driven techniques have been used for other gravity dam seepage prediction tasks, such as finding anomalous seepage rates and identifying potential failure modes based on historical data patterns. These applications show how data-driven methods can help us better comprehend the seepage behavior of gravity dams and improve our ability to predict and manage seepage-related risks. Figure 3 shows the various soft computing methods effectively used for seepage analysis.

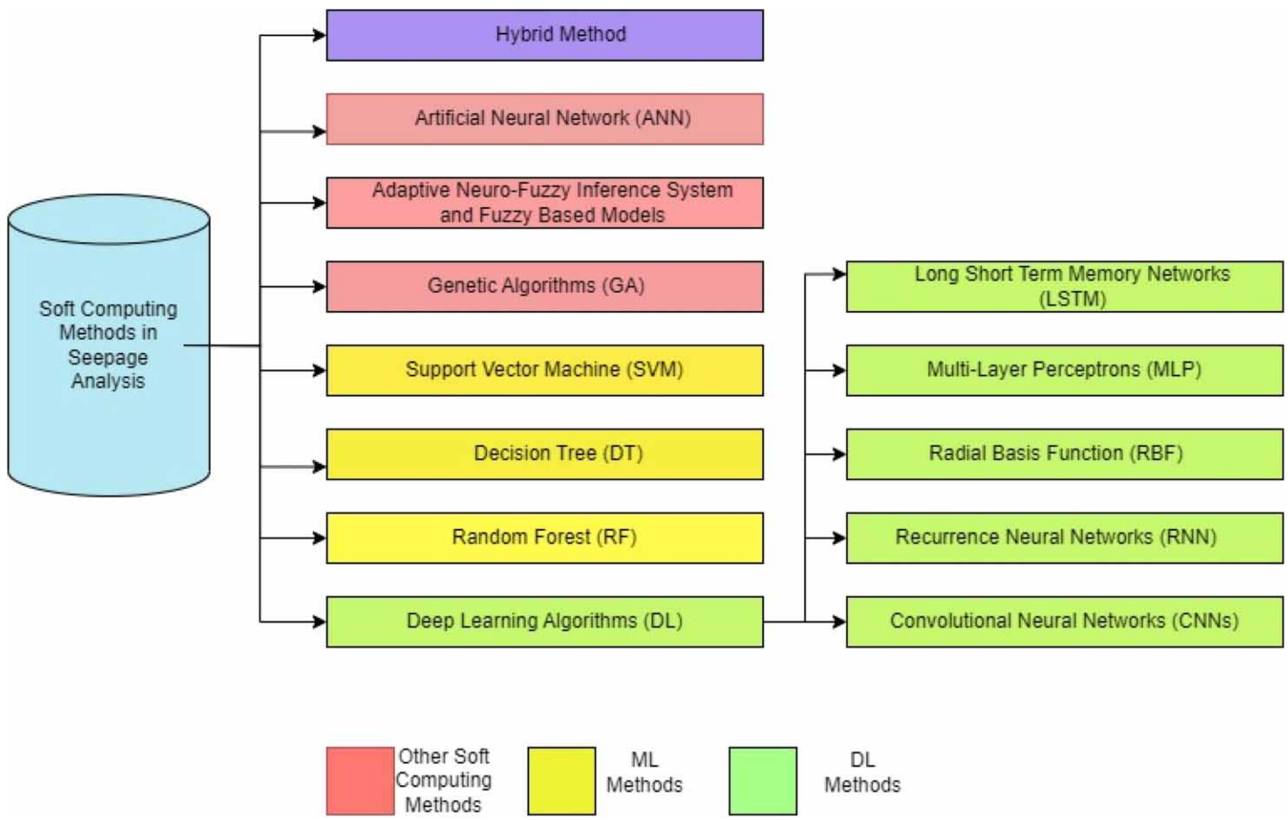


Figure 3 | Various soft computing methods effectively used for seepage analysis.

3.2.1. Artificial neural network (ANN)

Many researchers use ANNS for various types of non-linear problem-solving. The ANN structure is made up of three major units: the input layer, the hidden layer, and the output layer. When applied to non-linear issues, ANN has generated satisfactory results. Dam seepage analysis studies (Tayfur *et al.* 2005; Abdul Jabbar Jamel 2018) proved ANN for seepage prediction and found the model performed better than finite element analysis (FEM). In another study, rainfall-runoff modeling was used effectively (Rajurkar *et al.* 2004) to obtain significant findings using an error minimization approach. Zhang *et al.* (2021a) performed a comparative analysis of ANN and FEM for analyzing the group of five representative grouting schemes. To reduce errors, the error back propagation (EBP) technique is utilized, and several activation functions, including the linear function, Sigmoid function, ReLU function, Tanh function, and SoftMax function, can be used. Other ANN approaches include feed forward back propagation (FFBP) and radial basis function (RBF). Several such heuristic optimization techniques were used for performing optimizations (Kumar & Yadav 2022). Another study (Al-Saati *et al.* 2021) validated the long-term periodicity by applying various derived ANN models to historical data of Hindiya Barrage and obtained the projections for upcoming years utilizing Box-Jenkins models. These models have obtained similar predicted and actual values while testing and obtained acceptable performance validated using RMSE.

3.2.2. Adaptive neuro-fuzzy inference system (ANFIS) and fuzzy-based models

To maximize performance, ANFIS is an integrated system that integrates neural networks with fuzzy inference algorithms. Jang (1993) gives a thorough explanation of ANFIS techniques. ANFIS approach was used to perform seepage analysis for an earthen dam (Sharghi *et al.* 2019; Nourani *et al.* 2021) and utilizing a variety of soft computing techniques with successful results.

3.2.3. Genetic algorithms (GAs)

It is one of the often-used techniques for resolving both limited and unconstrained optimization issues. The concepts of genetics and natural selection underlie how GAs function (Yang 2021). GA can handle complex optimization problems by following the standard sequence of Initialization, Evaluation, Selection, Crossover, and Mutation. These advanced GA ultimately sift through the probabilities of crossover and mutation and may assess numerous distinct solutions to find the best solution to a problem (Katoch *et al.* 2021). In the context of earthen dams, GA was successfully utilized to anticipate water levels (Liu & Li 2015). These strategies appear to be used to solve search and optimization problems.

3.2.4. Support vector machine (SVM)

SVMs are supervised learning systems that are commonly used for classification, regression, and outlier identification (Singh *et al.* 2023b). When the dataset has high dimensionality, this approach may be efficiently implemented. Several problem statements in hydraulics and water resources are found to be multi-dimensional (Meng & Li 2017). This approach can successfully predict and does not require a big dataset for training to achieve accuracy and memory efficiency since it employs a subset of training points in the decision function (support vectors). This function enumerates several types of Kernel functions. SVM evaluates probabilities using a five-fold CV technique. SVR is a more advanced approach for resolving regression problems. Kernel functions utilized in SVM include linear, polynomial, sigmoid, and RBF (Smola & Vishwanathan 2008). The SVR approach was used for streamflow forecasting (Garsole & Rajurkar 2015), and the RBF Kernel function improved model performance. SVM Kernel function-based sensitivity analysis was utilized by Wang *et al.* (2018) and Gaur *et al.* (2021) to get optimal daily fluctuation in reservoir water levels.

3.2.5. Decision tree (DT)

DTs are non-parametric supervised learning approaches for classification and regression problems (Humbird *et al.* 2019). It is a tree-like structured model that learns basic decision rules to anticipate the most relevant value of the target feature (Al-Turjman *et al.* 2022). More sophisticated decision rules and a better-fit model may be obtained as the depth of the tree increases (Merghadi *et al.* 2020). These models are capable of handling both categorical and numerical data. One such study on stream flow forecasting (Londhe & Dixit 2011) successfully applied the DT model that was generated from it, known as the M5 Model Tree, which is utilized only for regression problems. In place of a categorical split in DTs, branching operations in M5 Model Trees employ linear regression methods.

3.2.6. Random forest (RF)

The RF is a common ensemble approach that combines the performance of numerous trees driven by diverse sources of data to get the optimal output. In the case of RF algorithms, the two crucial parameters are the number of trees to grow and the number of variables to split on (Belmokre *et al.* 2019). Breiman (2001) provides a comprehensive description of RF algorithms. Li *et al.* (2021) presented an approach for dam safety monitoring by building RF models. Comparatively faster and more accurate performance was obtained using RF algorithms as it poses high accuracy of prediction and better robustness due to uniformity of calculation. When working with RF algorithms, it is best practice to fine-tune various parameters to obtain model accuracy (Zhou *et al.* 2020). Some of the parameters that are typically changed while utilizing RF algorithms include maximum depth, estimators, maximum features, minimum sample split, random state, and the number of folds to conduct cross-validation (CV). Parameters are fine-tuned to achieve the required accuracy based on the issue description and data. Regression and classification issues are solved with RFs. Biases are frequently somewhat enhanced to reduce variation (Smola & Vishwanathan 2008).

3.2.7. Deep learning algorithms (DL)

The DL method contains several processing layers. DL approaches have been widely employed in image processing, speech recognition, visual object recognition, object detection, and a variety of other applications (LeCun *et al.* 2015). Deep convolutional networks have been widely employed in image processing, with several applications in hydrology and water resource engineering. Krizhevsky *et al.* (2017) have shown how to process satellite imagery using DL for high-resolution images, a method that differs from conventional image processing techniques. Sediment content in the reservoir was researched by Bejestan & Nouroozpou (2007), since it may be one of the influencing or regulating variables for seepage. To develop the decision support system for the river basin, where deep neural networks were embedded with a geographic information system (GIS) (Rohmat *et al.* 2019; Sharma *et al.* 2023), utilized DL algorithms.

Table 1 lists a detailed overview of the above-mentioned AI techniques and algorithms that have been used to anticipate seepage through dams. Figure 4 gives an overview of AI-ML techniques used in recent years. The graphical representation shows AI techniques ANN and its derived methods and RF have been used multiple times and given acceptable performance in the last five years.

4. CHALLENGES AND WAY-OUT

Quantity and quality of data: The data-driven method has demonstrated encouraging results in predicting seepage in gravity dams, several challenges need to be addressed. One of the most important concerns is the quantity and quality of data. Data availability and quality are significant challenges in creating accurate and reliable predictive models for gravity dam seepage prediction using the data-driven method. The quality and quantity of data used to train ML algorithms significantly influence their accuracy. Incomplete, inconsistent, or biased data can lead to unreliable model predictions. In the case of gravity dam seepage prediction, the data needed to train a model may include dam geometry, soil properties, water levels, rainfall, and other relevant variables. Obtaining accurate and complete data for all of these parameters can be difficult. Collecting extra measurements with sensors or other data collection methods is one approach to address data availability issues. Sensors, for example, can be installed in and around the dam to monitor water levels, temperature, and other environmental parameters. Satellite data and remote sensing techniques can also be used to collect data on rainfall, topography, and other pertinent parameters. More data collection can improve the models' precision and reliability.

Noise and errors in the data: To decrease data noise and errors utilizing cutting-edge data processing methods. Data smoothing techniques, for example, can be used to remove outliers and reduce the effect of noise on model predictions. Furthermore, data imputation methods can be used to fill in missing values in the data, reducing the effect of data sparsity on model predictions. Furthermore, data quality can be improved by ensuring that it is collected and processed following standardized procedures. This can aid in reducing mistakes and inconsistencies in the data and ensuring its suitability for use in ML models.

Uncertainties in the model: In many cases, the data used to train ML models contains uncertainties such as measurement errors, modeling assumptions, or model parameter estimates. These uncertainties can spread throughout the algorithm, affecting predictability. Using probabilistic ML models that can account for uncertainties is one method to address uncertainties. Bayesian ML techniques, for example, are used to create models that incorporate uncertainty in model parameters and output

Table 1 | Literature surveys of artificial intelligence techniques used for seepage predictions

Authors and year	Type of dam/ specific case study	Analysis	Analytical tools	Analysis outcome	Remarks
Tayfur *et al.* (2005)	Earthfill Dam in Poland	Seepage predictions for earthen dams and water levels estimation	Finite element method (FEM), artificial neural network (ANN)	Adequate accuracy of ANN was observed while predicting the water level in piezometers on the earthen dam and seepage	Performance of ANN seepage predictions found better compared to FEM
He *et al.* (2012)	Xiaowan, Arch Dam	Dam foundation seepage simulation	Feed forward neural network (FFNN), finite element analysis (FEM)	Moderate performance of models was observed when proposed methods were applied for the simulation of seepage and permeability measurement	The average performance of ANN
Liu & Li (2015)	Slope stability	Stability analysis and water-seepage modeling	Genetic algorithms (GAs), finite element method (FEM), and limit equilibrium method (LEM)	The solution was provided using GA for 1D groundwater flow in an unconfined aquifer	GA was effectively used for ground water level (GWL) predictions
Garsole & Rajurkar (2015)	Streamflow forecasting	Prediction of streamflow based on historical data collected for rainfall and runoff	Support vector regression (SVR)	Observed better performance of SVRs for streamflow forecasting while considering limited parameters	Considerable performance of SVR
Saraf & Regulwar (2016)	Upper Godavari River Basin, India	Predictions and projections of temperature and precipitation	Hadlycentre Coupled Climate Model (HadCM3), Coupled Global Climate Model (CGCM3), Statistical Down-Scaling Model (SDSM), Multiple Regression (MLR), Stochastic Weather Generator (SWG)	Multiple regression techniques have been used effectively used for applications such as downscaling of temperature and precipitation	Projected downscaled values can be effectively used for further predictions of seepage
Abdul Jabbar Jamel (2018)	Homogeneous Earthen Dam	Investigation and estimation of seepage discharge through a homogeneous earthen dam with core	ANN and SEEP/W models	Discharges calculated by ANN with a single hidden layer and SEEP/W model are nearly the same and hence can be effectively used	The performance of ANN has proven to be better for the estimation of seepage
Yadav *et al.* (2018)	Research overview	Analysis of hybrid framework for optimization	A hybrid framework for lean six sigma (LSS), fuzzy AHP-PROMETHEE	Combining fuzzy sets in the analytical hierarchy process (AHP) gives the optimum solution	The performance of the fuzzy system can give inputs to researchers for better results in the optimization process
Sermet & Demir (2018)	Flood automation	An intelligent system for knowledge generation and communication about flooding	Artificial intelligence (AI), natural language processing (NLP), PHP Hypertext Pre-processor	Framework has designed communication-based channels for flood intuitions using AI	AI-based flood alert systems have been successfully designed
Wang *et al.* (2018)	Concrete gravity dam	Concrete gravity dam seepage modeling	Support vector regression (SVR)	Optimized SVM models show the impact of the daily variation of the reservoir water level and rainfall on the daily variation of the piezometric tube water level following the normal distribution	–
Rohmat *et al.* (2019)	Irrigated River Basin	Study of the impact of best management practices (BMP) on stream-aquifer exchange and water law compliance of irrigated river basin using DL	Deep learning (DL)	Deep learning has been effectively implemented in decision support systems. An embedded DNN and GIS-based basin scale decision support system was developed to access the impacts	DL methods can be effectively used with GIS tools for building an efficient decision support system
Belmokre *et al.* (2019)	Algerian Roller Compact Gravity Dam	Seepage and dam deformation analysis with statistical models	Support vector regression (SVR), random forest (RF)	Prediction concludes the better performance of the random forest regression model provides	Better performance of RF models observed in

				comparatively than the support vector machine	comparison with SVR model
Sharghi *et al.* (2019)	Sattarkhan Earthen Dam, Northwest Iran	Earthfill dam seepage modeling	FFNN, ANFIS, SVR	A combination of Jittering and Ensembling (especially neural ensemble) has enhanced the model performance by 30% in the testing phase	The ensemble technique performs better provided used as per requirements
Sit *et al.* (2020)	Review	Applications of deep learning in hydrology and water resources engineering	Artificial intelligence (AI), deep learning (DL), machine learning (ML), Convolutional LSTM, ANN, CNN, deep belief network, deep Boltzmann machine, deep neural network, FFNN, generative adversarial networks, restricted Boltzmann machine, recurrent neural network, support vector machines, support vector regression, rectified linear unit, multi-layer perceptron	Deep learning tools are implicitly used for decision-making applications. Algorithms leading to decision-making or predictions are built using deep learning	Deep learning algorithms have been used for various applications in hydrology and water resources engineering
Chen *et al.* (2020)	Canal	Water leak detection in canal sections	Convolutional neural network (CNN)	More computationally efficient performance of CNN over traditional deep learning algorithms. Reduction in the number of channels has enhanced the performance	–
Tao & Zheng (2020)	Case study	Seepage damage identification	Adaptive fuzzy identification (AFI) algorithm and anthropomorphic causal reasoning (ACR)	The study approach shows the consistency of the proposed methods in real-time moisture content assessment	–
Nourani *et al.* (2021)	Review	Applications of soft computing methods for seepage modeling	ANN, adaptive neuro-fuzzy inference system and fuzzy-based models (ANFIS-FBM), SVM, GP, DL, non-linear autoregressive model with exogenous inputs, random forest (RF), decision tree (DT), hybrid models	Soft computing models are efficient and reliable tools for seepage modeling, especially when they are linked to data pre- or post-processing techniques	Soft computing methods can be used for seepage prediction modeling
Rehamnia *et al.* (2021)	Fontaine Gazelles Dam, Algeria	Embankment dam seepage modeling	Extended Kalman filter feed forward neural network (EKF-ANN), MLP, RBF-NN, RF	The excellence performance of EKF-ANN over MLP, RF, and RBF-NN was observed	–
Zhang *et al.* (2021a)	Deqing Dam and Hydropower Station	Seepage modeling through the base of a dam	Feed forward neural network (FFNN)	Condition-specific average accuracy in prediction obtained by 3D finite element method model of complex geological bodies combined with ANN for the model of inversion of seepage field of the dam	–
Wei *et al.* (2021)	Natural soil slope, Earthen Embankment	Pore water pressure estimation	RNN, long short-term memory (LSTM), GRU, MLP	GRU and LSTM models can provide better accuracy and more robust, precise prediction considering the performance of standard RNNs	GRU and LSTM give robust and precise predictions for pore water pressure predictions

(*Continued.*)

Table 1 | Continued

Authors and year	Type of dam/ specific case study	Analysis	Analytical tools	Analysis outcome	Remarks
Ishfaque *et al.* (2022b)	Tarbela Dam, Pakistan (Earthen Dam)	Seepage prediction and understanding the effect of hydro-climatological parameters and other factors on prediction using SHAP analysis	ANN, RF, SVM, and CatBoost (CB), Shapley Additive Explanations (SHAP)	Machine learning models are reliable in predicting and understanding dam seepage. Manual errors in data collecting and analysis are highly prone to errors	ML models work well in cases where adequate data is available
Beiranvand & Rajaee (2022)	Review	Prediction of seepage and pore water pressure in dams	Single models' artificial neural networks (ANNs), support vector regression (SVR), random forest (RF), feed forward neural network (FFNN), and Hybrid models such as adaptive-network-based fuzzy inference system (ANFIS), support vector machine (SVM), classification and regression tree (CART), instance-based learning (IBL), artificial bee colony (ABC), teaching-learning based optimization (TLBO), genetic programming (GP), and gene-expression programming (GEP)	The complexity of indirectly determining the issue of dam seepage measurement is resolved using AI techniques. The shortcomings of numerical methods and the complexity of dam behavior can be effectively overcome by using hybrid or single AI models	The performance of AI techniques has overcome the shortcomings of numerical methods effectively
Bouchehed *et al.* (2023)	Embankment Dams	Predictions of seepage through embankment dams	Support vector regression (SVR), relevance vector machine (RVM), Gaussian process regression (GPR), genetic programming (GP), ANN, multi-layer perceptron neural network (MLPNN), multivariate adaptive regression splines (MARS)	The performance of model RVM with R and NSE values of ~0.909 and ~0.767 while the performance of SVR was poor as compared to RVM with R and NSE values of ~0.780 and ~0.600	RVM performance was observed to be better than SVM

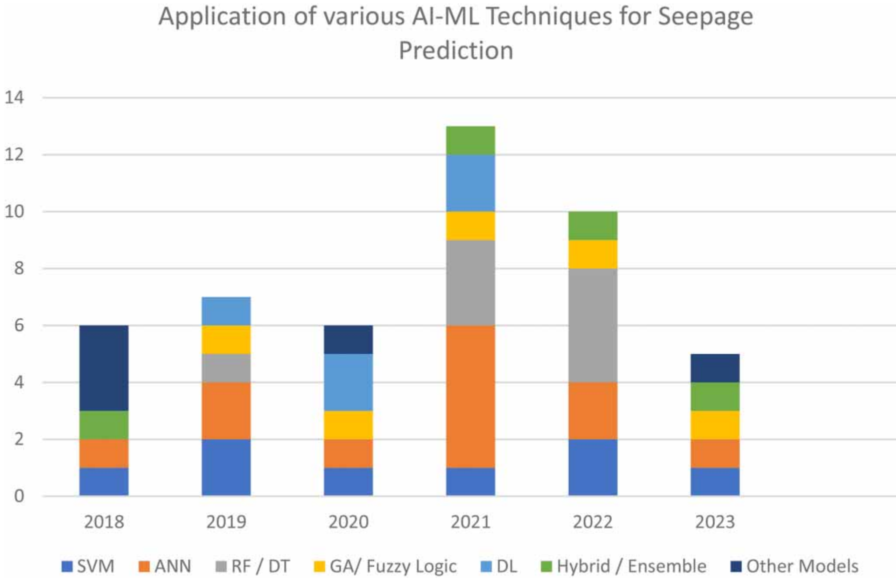


Figure 4 | Various AI-ML methods effectively applied for dam seepage prediction.

predictions as probability distributions. Models can provide more reliable predictions and better quantify the degree of trust in the predictions by accounting for uncertainties.

Interpretability and explainability: In recent years, the interpretability and explainability of ML models have become increasingly essential. This is particularly true in decision-making processes, where the reasoning underlying the model's predictions is critical. In the case of gravity dam seepage prediction, decision-makers must understand how the model generates its predictions to correctly evaluate the risk of dam failure and take appropriate measures. Many ML algorithms, on the other hand, are regarded as black boxes due to their absence of interpretability and explainability. In other words, understanding how the model produces predictions or which features have the biggest impact on the model's output might be challenging. This can pose a serious issue in a variety of real-world scenarios, such as gravity dam seepage prediction, when the stakes are high and informed choices are essential. Several approaches can be adopted to address this challenge. One method is to use visualization methods to assist in interpreting the model's output. Heat maps or saliency maps, for example, can be used to highlight which features have the greatest influence on the model's output, making it easier to comprehend how the model arrived at its prediction. Model-agnostic methods, such as LIME (Local Interpretable Model-Agnostic Explanations), can also be used to describe the model's behavior at the local level. LIME creates simple, interpretable models that can describe complex ML model predictions.

Complexity: The need for a compromise between model complexity and interpretability is important in gravity dam seepage prediction. DL models, for example, can provide more accurate predictions, but they may be more challenging to interpret and explain. Simpler models, such as linear regression models, may be more interpretable, but they may not correctly capture the complexity of the seepage process. As a result, it is critical to strike a balance between model complexity and interpretability that satisfies the application's requirements. One solution is to use hybrid models, which combine the strengths of numerous models, such as DTs and neural networks, to provide accurate and interpretable predictions.

Updating and improvement in the model per new dataset: Updating and improving the model as a new dataset: ML models are trained on a particular dataset, and their accuracy may deteriorate over time as new data becomes accessible. As a result, the models must be continuously monitored and updated to ensure that they stay accurate and reliable. This can be accomplished through the use of techniques such as online learning or active learning, which enable the model to be continuously updated based on new data.

Deployment and implementation of models: ML models are deployed and implemented in real-world applications. Problems such as model scalability, data privacy, and digital security may affect the model's deployment and implementation. For example, large-scale deployment of ML models may necessitate substantial computational resources, whereas ensuring data privacy and security may necessitate the use of encryption techniques or secure data sharing protocols. Furthermore, the models must be incorporated into existing decision-making processes and workflows, which may necessitate significant changes in the culture and practices of the organization. It is crucial to collaborate with specialists from a variety of fields; including data science, civil engineering, and decision-making, to develop a complete approach that incorporates ML models with current practices and processes to address these difficulties. Install and operate ML models effectively, this may include developing specific tools and platforms as well as educating staff on how to utilize them. Furthermore, clear guidelines and procedures for data sharing, privacy, and security must be established to ensure that the models are implemented responsibly and ethically.

In summation, while the data-driven method has shown potential in predicting seepage in gravity dams, several challenges remain. Data availability and quality, interpretability, and explainability, model complexity and scalability, and deployment and implementation are examples of these. Addressing these challenges will necessitate a collaborative effort involving experts from various disciplines, as well as the development of specialized tools and platforms to aid in the deployment and implementation of ML models.

5. FUTURE DIRECTION

There are several potential future study directions in the area of data-driven methods for gravity dam seepage prediction.

(a) One possible path is the development of more advanced ML algorithms capable of capturing the complex interactions between various factors influencing seepage behavior in gravity dams. DL algorithms, for example, can be used to create more accurate and dependable models by capturing non-linear relationships between variables. Furthermore, ensemble methods like RFs or gradient boosting can be used to combine the strengths of numerous ML models to create more accurate and reliable predictions.

(b) Physical models with data-driven models: Although physical models can provide useful insights into gravity dam seepage behavior, they are computationally expensive and may not capture all of the complex interactions between various factors. It may be possible to create hybrid models that can provide accurate and reliable predictions while still capturing the essential physics of the system by combining physical models with data-driven models.

(c) Accuracy and reliability of the models: There is also a need to improve the interpretability and explainability of the models to improve their accuracy and dependability. As previously stated, interpretability and explainability are critical for decision-making processes in which the reasoning behind the model's predictions is critical. The development of more advanced visualization techniques that can help users better understand how the model arrives at its predictions is one possible path for future research. Furthermore, more work is required to create explainable AI (XAI) methods capable of producing transparent and interpretable models.

(d) Real-time monitoring and feedback: Real-time monitoring and feedback are being integrated into the seepage prediction system. It may be possible to improve the accuracy and reliability of the models over time by constantly monitoring key parameters such as water levels, temperature, and soil moisture. Furthermore, real-time feedback can be used to inform choices about gravity dam operation and maintenance.

(e) Finally, more comprehensive experimental testing and validation of data-driven models are required. While the data-driven method has shown encouraging results in predicting seepage in gravity dams, more research is required to validate the models under a variety of conditions and compare their performance to traditional methods. To guarantee that the models can be applied to novel and unanticipated situations, they should also be tested using data from other dams.

In general, the creation of accurate and reliable data-driven models for gravity dam seepage prediction has the potential to raise the security and effectiveness of these significant buildings. To resolve these problems and fully realize the promise of these strategies, further study is needed.

6. CONCLUSIONS

The state-of-the-art review included an overview of the difficulties in using AI to anticipate seepage in gravity dams as well as possible future possibilities. Gravity dam seepage prediction is a challenging issue that calls for precise and reliable models. Although the data-driven approach has produced encouraging results, there are still problems that need to be solved. Addressing issues with data availability and quality, model complexity, and model interpretability and explainability is necessary to increase forecast accuracy and dependability. The emphasis is on building cutting-edge ML algorithms that can accurately capture the complexity of the seepage process. Data errors and inconsistencies may also be decreased by increasing the quantity and quality of the data used to train the models and by developing consistent methods for data collecting and processing. Another crucial topic for future study is enhancing the models' interpretability and explicability. Furthermore, it is possible to look into developing models that are easier to understand and interpret, like rule-based models. Finally, gravity dams, which are significant infrastructure assets, may be made safer and more stable by addressing these problems and upgrading the models. Future research in this area may enhance the ability to predict seepage in gravity dams, which is essential for guaranteeing the security and dependability of these vital structures. Researchers and professionals working on the subject of dam safety and infrastructure management will find this review to be a relevant resource.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Abba, S. I., Nourani, V. & Elkiran, G. 2019 Multi-parametric modeling of water treatment plant using AI-based non-linear ensemble. *Journal of Water Supply: Research and Technology-AQUA* **68** (7), 547–561.

Abdul Jabbar Jamel, A. 2018 Investigation and estimation of seepage discharge through homogenous earth dam with core by using SEEP/W model and artificial neural network. *DJES* **11** (3), 54–61.

Al-Juboori, M. & Datta, B. 2019 Performance evaluation of a genetic algorithm-based linked simulation-optimization model for optimal hydraulic seepage-related design of concrete gravity dams. *Journal of Applied Water Engineering and Research* **7** (3), 173–197.

Al-Saati, N. H., Omran, I. I., Salman, A. A., Al-Saati, Z. & Hashim, K. S. 2021 Statistical modeling of monthly streamflow using time series and artificial neural network models: Hindiya Barrage as a case study. *Water Practice and Technology* **16** (2), 681–691.

Al-Turjman, F., Hussain, A. A., Alturjman, S. & Altrjman, C. 2022 Vehicle price classification and prediction using machine learning in the IoT smart manufacturing era. *Sustainability* **14** (15), 9147.

Athani, S. S., Shivamanth, Solanki, C. H. & Dodagoudar, G. R. 2015 Seepage and stability analyses of earth dam using finite element method. *Aquatic Procedia* **4**, 876–883.

Beiranvand, B. & Komasi, M. 2021 An investigation on performance of the cut off wall and numerical analysis of seepage and pore water pressure of Eyvashan earth dam. *Iranian Journal of Science and Technology, Transactions of Civil Engineering* **45** (3), 1723–1736.

Beiranvand, B. & Rajaee, T. 2022 Application of artificial intelligence-based single and hybrid models in predicting seepage and pore water pressure of dams: a state-of-the-art review. *Advances in Engineering Software* **173**, 103268.

Bejestan, M. S. & Nouroozpou, S. 2007 Use of image processing technique to estimate sediment concentration. *Journal of Applied Sciences* **7** (20), 3096–3100.

Belmokre, A., Mihoubi, M. K. & Santillan, D. 2019 Seepage and dam deformation analyses with statistical models: Support vector regression machine and random forest. *Procedia Structural Integrity* **17**, 698–703.

Bernier, C., Padgett, J. E., Proulx, J. & Paultre, P. 2016 Seismic fragility of concrete gravity dams with spatial variation of angle of friction: Case study. *Journal of Structural Engineering* **142** (5). https://doi.org/10.1061/(ASCE)ST.1943-541X.0001441

Bouchehed, A., Laouacheria, F., Heddam, S. & Djemili, L. 2023 Machine learning for better prediction of seepage flow through embankment dams: Gaussian process regression versus SVR and RVM. *Environmental Science and Pollution Research* **30** (9), 24751–24763.

Breiman, L. 2001 Random forest. *Machine Learning* **45** (1), 5–32.

Cancienne, R. M., Fox, G. A. & Simon, A. 2008 Influence of seepage undercutting on the stability of root-reinforced streambanks. *Earth Surface Processes and Landforms* **33** (11), 1769–1786.

Castro-Orgaz, O. & Giráldez, J. V. 2012 Steady-state water table height estimations with an improved pseudo-two-dimensional Dupuit-Forchheimer type model. *Journal of Hydrology* **438–439**, 194–202.

Chen, J., Tang, P., Rakstad, T., Patrick, M. & Zhou, X. 2020 Augmenting a deep-learning algorithm with canal inspection knowledge for reliable water leak detection from multispectral satellite images. *Advanced Engineering Informatics* **46**, 101161.

Cheng, Z., Wang, X., Ren, B., Lv, P. & Zhu, X. 2018 Simulation of complex seepage field of a gravity dam foundation using a CFD-based approach. *Journal of Hydraulic Engineering* **144** (3). https://doi.org/10.1061/(ASCE)HY.1943-7900.0001428

Demetracopoulos, A. C. & Hadjitheodorou, C. 1996 Seepage from surface canals by boundary element method. *Journal of Irrigation and Drainage Engineering* **122** (1), 40–48.

El Bilali, A., Moukhliss, M., Taleb, A., Nafii, A., Alabjah, B., Brouziyne, Y., Mazigh, N., Teznine, K. & Mhamed, M. 2022 Predicting daily pore water pressure in embankment dam: Empowering machine learning-based modeling. *Environmental Science and Pollution Research* **29** (31), 47382–47398.

Fatima, M. & Pasha, M. 2017 Survey of machine learning algorithms for disease diagnostic. *Journal of Intelligent Learning Systems and Applications* **09** (01), 1–16.

Fu, J. C., Hsu, M. H. & Duann, Y. 2016 Development of roughness updating based on artificial neural network in a river hydraulic model for flash flood forecasting. *Journal of Earth System Science* **125** (1), 115–128.

Garg, S. K., 2005 *Irrigation Engineering and Hydraulic Structures: Vol. II*. (Garg, S. K., ed.; 19th edn.) Khanna Publication, Delhi, India.

Garsole, P., Rajurkar, M. P., 2015 Streamflow forecasting by using support vector regression. In: *HYDRO 2015 INTERNATIONAL, 20th International Conference on Hydraulics, Water Resources and River Engineering, IIT Roorkee* (Garsole, P. & Rajurkar, M. P., eds). IIT Roorkee. 17–19 December 2015. Available from: https://www.academia.edu/19881517/Streamflow_Forecasting_by_using_Support_Vector_Regression

Gaur, S., Johannet, A., Graillot, D. & Omar, P. J. 2021 Modeling of groundwater level using artificial neural network algorithm and WA-SVR model. In: *Groundwater Resources Development and Planning in the Semi-Arid Region*. Springer International Publishing, New York, USA, pp. 129–150.

He, J., Chen, S.-H. & Shahrour, I. 2012 A revised solution of equivalent permeability tensor for discontinuous fractures. *Journal of Hydrodynamics* **24** (5), 711–717.

Hu, J. & Ma, F. 2016 Comprehensive investigation method for sudden increases of uplift pressures beneath gravity dams: Case study. *Journal of Performance of Constructed Facilities* **30** (5). https://doi.org/10.1061/(ASCE)CF.1943-5509.0000874

Hu, J., Ma, F. & Wu, S. 2018 Comprehensive investigation of leakage problems for concrete gravity dams with penetrating cracks based on detection and monitoring data: A case study. *Structural Control and Health Monitoring* **25** (4), e2127. 1–18.

Humbird, K. D., Peterson, J. L. & Mcclarren, R. G. 2019 Deep neural network initialization with decision trees. *IEEE Transactions on Neural Networks and Learning Systems* **30** (5), 1286–1295.

Ishfaque, M., Dai, Q., Haq, N. u., Jadoon, K., Shahzad, S. M. & Janjuhah, H. T. 2022a Use of recurrent neural network with long short-term memory for seepage prediction at Tarbela Dam, KP, Pakistan. *Energies* **15** (9), 3123.

Ishfaque, M., Salman, S., Jadoon, K. Z., Danish, A. A. K., Bangash, K. U. & Qianwei, D. 2022b Understanding the effect of hydro-climatological parameters on dam seepage using Shapley Additive Explanation (SHAP): A case study of earth-fill Tarbela Dam, Pakistan. *Water* **14** (17), 2598.

Jang, J.-S. R. 1993 ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics* **23** (3), 665–685.

Kang, F., Liu, J., Li, J. & Li, S. 2017 Concrete dam deformation prediction model for health monitoring based on extreme learning machine. *Structural Control and Health Monitoring* **24** (10), e1997.

Kashef, A.-A. I. 1965 Seepage through earth dams. *Journal of Geophysical Research* **70** (24), 6121–6128.

Katoch, S., Chauhan, S. S. & Kumar, V. 2021 A review on genetic algorithm: Past, present, and future. *Multimedia Tools and Applications* **80** (5), 8091–8126.

Krizhevsky, A., Sutskever, I. & Hinton, G. E. 2017 Imagenet classification with deep convolutional neural networks. *Communications of the ACM* **60** (6), 84–90.

Kumar, V. & Yadav, S. M. 2021 Real-time flood analysis using artificial neural network, In: *Recent Trends in Civil Engineering* (Pathak, K. K., Bandara, J. M. S. J. & Agrawal, R., eds). Lecture Notes in Civil Engineering 77. Springer, Singapore. https://dx.doi.org/10.1007/978-981-15-5195-6_71.

Kumar, V. & Yadav, S. M. 2022 A state-of-the-art review of heuristic and metaheuristic optimization techniques for the management of water resources. *Water Supply* **22** (4), 3702–3728.

Lai, Y., Li, S., Guo, J., Zhu, Z. & Huang, X. 2021 Analysis of seepage and displacement field evolutionary characteristics in water inrush disaster process of karst tunnel. *Geofluids* **2021**, 1–20.

LeCun, Y., Bengio, Y. & Hinton, G. 2015 Deep learning. *Nature* **521** (7553), 436–444.

Lee, J.-W., Kim, J. & Kang, G.-C. 2018 Seepage behavior of earth dams considering rainfall effects. *Advances in Civil Engineering* **2018**, 1–9.

Lei, L., Zhou, Y., Huang, H. & Luo, Q. 2022 Extreme learning machine using improved gradient-based optimizer for dam seepage prediction. *Arabian Journal for Science and Engineering*. https://dx.doi.org/10.1007/s13369-022-07300-8

Li, M. & Wang, J. 2019 An empirical comparison of multiple linear regression and artificial neural network for concrete dam deformation modelling. *Mathematical Problems in Engineering* **2019**, 1–13.

Li, X., Wen, Z. & Su, H. 2021 An approach using random forest intelligent algorithm to construct a monitoring model for dam safety. *Engineering with Computers* **37** (1), 39–56.

Liu, Q. Q. & Li, J. C. 2015 Effects of water seepage on the stability of soil-slopes. *Procedia IUTAM* **17**, 29–39.

Liu, Y., Zheng, D., Wu, X., Chen, X., Georgakis, C. T. & Qiu, J. 2023 Research on prediction of dam seepage and dual analysis of lag-sensitivity of influencing factors based on MIC optimizing random forest algorithm. *KSCE Journal of Civil Engineering* **27** (2), 508–520.

Londhe, S. N. & Dixit, P. R. 2011 Forecasting stream flow using model trees. *International Journal of Earth Sciences and Engineering* **4** (6), 288–285.

Meng, J. & Li, H. 2017 An efficient stochastic approach for flow in porous media via sparse polynomial chaos expansion constructed by feature selection. *Advances in Water Resources* **105**, 13–28.

Merghadi, A., Yunus, A. P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D. T., Avtar, R. & Abderrahmane, B. 2020 Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews* **207**, 103225.

Messerklinger, S. 2014 Failure of a geomembrane lined embankment dam – case study. *Geotextiles and Geomembranes* **42** (3), 256–266.

Nagure, A. S. & Shahapure, S. S. 2021 Effect of watershed characteristics on a rainfall runoff analysis and hydrological model selection – a review. In *2021 International Conference on Computing, Communication and Green Engineering (CCGE)*, pp. 1–7.

Nourani, V., Elkiran, G. & Abba, S. I. 2018 Wastewater treatment plant performance analysis using artificial intelligence – an ensemble approach. *Water Science and Technology* **78** (10), 2064–2076.

Nourani, V., Behfar, N., Dabrowska, D. & Zhang, Y. 2021 The applications of soft computing methods for seepage modeling: A review. *Water* **13** (23), 3384. MDPI.

Omar, P. J., Dwivedi, S. B. & Dikshit, P. K. 2019 Temporal variability study in rainfall and temperature over Varanasi and adjoining areas. *Disaster Advances* **12** (1), 1–7.

Omar, P. J., Dwivedi, S. B. & Dikshit, P. K. S. 2020 Sustainable development and management of groundwater in Varanasi, India. In: *Lecture Notes in Civil Engineering*, Vol. 39. Springer, Singapore, pp. 201–209.

Omar, P. J., Gaur, S. & Dikshit, P. K. S. 2021 Conceptualization and development of multi-layered groundwater model in transient condition. *Applied Water Science* **11** (10), 162.

Omar, P. J., Rai, S. P. & Tiwari, H. 2022a Study of morphological changes and socio-economic impact assessment: A case study of Koshi River. *Arabian Journal of Geosciences* **15** (17), 1426.

Omar, P. J., Shivhare, N., Dwivedi, S. B. & Dikshit, P. K. S. 2022b Identification of soil erosion-prone zone utilizing geo-informatics techniques and WSPM model. *Sustainable Water Resources Management* **8** (3), 66.

Omofunmi, O., Kolo, J., Oladipo, A., Diabana, P. & Ojo, A. 2017 A review on effects and control of seepage through earth-fill dam. *Current Journal of Applied Science and Technology* **22** (5), 1–11.

Rajurkar, M. P., Kothyari, U. C. & Chaube, U. C. 2004 Modeling of the daily rainfall-runoff relationship with artificial neural network. *Journal of Hydrology* **285** (1–4), 96–113.

Refaiy, A. R., AboulAtta, N. M., Saad, N. Y. & El-Molla, D. A. 2021 Modeling the effect of downstream drain geometry on seepage through earth dams. *Ain Shams Engineering Journal* **12** (3), 2511–2531.

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.

Rohmat, F. I. W., Labadie, J. W. & Gates, T. K. 2019 Deep learning for compute-efficient modeling of BMP impacts on stream-aquifer exchange and water law compliance in an irrigated river basin. *Environmental Modelling & Software* **122**, 104529.

Russoniello, C. J. & Michael, H. A. 2015 Investigation of seepage meter measurements in steady flow and wave conditions. *Groundwater* **53** (6), 959–966.

Salazar, F., Morán, R., Toledo, M. Á. & Oñate, E. 2017 Data-based models for the prediction of dam behaviour: A review and some methodological considerations. *Archives of Computational Methods in Engineering* **24** (1), 1–21.

Salem, M. N. 2019 Analysis of seepage through earth dams with internal core. *International Journal of Engineering Research* **8** (08), 768–777.

Saraf, V. R. & Regulwar, D. G. 2016 Assessment of climate change for precipitation and temperature using statistical downscaling methods in Upper Godavari River Basin, India. *Journal of Water Resource and Protection* **08** (01), 31–45.

Sermet, Y. & Demir, I. 2018 An intelligent system on knowledge generation and communication about flooding. *Environmental Modelling & Software* **108**, 51–60.

Sermet, Y., Demir, I. & Muste, M. 2020 A serious gaming framework for decision support on hydrological hazards. *Science of The Total Environment* **728**, 138895.

Seth, D. & Pandey, M. K. 2009 A multiple-item inventory model for a non-stationary demand. *Production Planning and Control* **20** (3), 242–253.

Seth, D., Nemani, V. K., Pokharel, S. & Al Sayed, A. Y. 2018 Impact of competitive conditions on supplier evaluation: A construction supply chain case study. *Production Planning and Control* **29** (3), 217–235.

Sevieri, G., De Falco, A., Andreini, M. & Matthies, H. G. 2021 Hierarchical Bayesian framework for uncertainty reduction in the seismic fragility analysis of concrete gravity dams. *Engineering Structures* **246**, 113001.

Sharghi, E., Nourani, V. & Behfar, N. 2018 Earthfill dam seepage analysis using ensemble artificial intelligence based modeling. *Journal of Hydroinformatics* **20** (5), 1071–1084.

Sharghi, E., Nourani, V., Behfar, N. & Tayfur, G. 2019 Data pre-post processing methods in AI-based modeling of seepage through earthen dams. *Measurement* **147**, 106820.

Sharma, T., Prasad, S. K. & Mishra, P. 2022 Flood prediction analysis using supervised machine learning techniques. In *2022 IEEE Delhi Section Conference (DELCON 2022)*, Vol. 21, No. 11, pp. 552–560. ISSN: 0044-0477.

Sharma, K. V., Kumar, V., Singh, K. & Mehta, D. J. 2023 LANDSAT 8 LST Pan sharpening using novel principal component based downscaling model. *Remote Sensing Applications: Society and Environment* **30**, 100963.

Singh, A. N., Mudgal, A., Tripathi, R. P. & Omar, P. J. 2023a Assessment of wastewater treatment potential of sand beds of River Ganga at Varanasi, India. *AQUA – Water Infrastructure, Ecosystems and Society* **72** (5), 690–700.

Singh, K., Singh, B., Sihag, P., Kumar, V. & Sharma, K. V. 2023b Development and application of modeling techniques to estimate the unsaturated hydraulic conductivity. *Modeling Earth Systems and Environment*. https://doi.org/10.1007/s40808-023-01744-z.

Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y. & Demir, I. 2020 A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology* **82** (12), 2635–2670.

Smola, A., Vishwanathan, S. V. N., 2008 *Introduction to Machine Learning*, Vol. 252 (Smola, A. & Vishwanathan, S. V. N., eds.). Cambridge University Press, Cambridge, UK. ISBN 0 521 82583 0.

Song, J., Zhang, S., Tong, F., Yang, J., Zeng, Z. & Yuan, S. 2021 Outlier detection based on multivariable panel data and K-means clustering for dam deformation monitoring data. *Advances in Civil Engineering* **2021**, 1–11.

Su, H., Chen, Z. & Wen, Z. 2016a Performance improvement method of support vector machine-based model monitoring dam safety. *Structural Control and Health Monitoring* **23** (2), 252–266.

Su, H., Tian, S., Cui, S., Yang, M., Wen, Z. & Xie, W. 2016b Distributed optical fiber-based theoretical and empirical methods monitoring hydraulic engineering subjected to seepage velocity. *Optical Fiber Technology* **31**, 111–125.

Sur, H. S., Bhardwaj, A. & Jindal, P. K. 1999 Some hydrological parameters for the design and operation of small earthen dams in lower Shiwaliks of northern India. *Agricultural Water Management* **42** (1), 111–121.

Tao, K. & Zheng, W. 2020 An anthropomorphic fuzzy model for the time-spatial assessment of sandstone seepage damage. *Automation in Construction* **109**, 102989.

Tayfur, G., Swiatek, D., Wita, A. & Singh, V. P. 2005 Case study: finite element method and artificial neural network models for flow through Jeziorsko earthfill dam in Poland. *Journal of Hydraulic Engineering* **131** (6), 431–440.

Wahl, T. L. 2004 Uncertainty of predictions of embankment dam breach parameters. *Journal of Hydraulic Engineering* **130** (5), 389–397.

Wang, X., Yang, K. & Shen, C. 2017 Study on MPGA-BP of gravity dam deformation prediction. *Mathematical Problems in Engineering* **2017**, 1–13.

Wang, S. W., Xu, Y. L., Gu, C. S. & Bao, T. F. 2018 Monitoring models for base flow effect and daily variation of dam seepage elements considering time lag effect. *Water Science and Engineering* **11** (4), 344–354.

Wang, X., Yu, H., Lv, P., Wang, C., Zhang, J. & Yu, J. 2019 Seepage safety assessment of concrete gravity dam based on matter-element extension model and FDA. *Energies* **12** (3), 502.

Wei, X., Zhang, L., Yang, H.-Q., Zhang, L. & Yao, Y.-P. 2021 Machine learning for pore-water pressure time-series prediction: Application of recurrent neural networks. *Geoscience Frontiers* **12** (1), 453–467.

Xu, T. & Liang, F. 2021 Machine learning for hydrologic sciences: an introductory overview. *WIREs Water* **8** (5), 1–29.

Xu, Y. & Wu, Z. 2022 Parameter identification of unsaturated seepage model of core rockfill dams using principal component analysis and multi-objective optimization. *Structures* **45**, 145–162.

Xu, Q., Chen, J.-T. & Xiao, M. 2020 Analysis of unsteady seepage field and surrounding rock stability of underground cavern excavation. *Tunnelling and Underground Space Technology* **97**, 103239.

Yadav, G., Seth, D. & Desai, T. N. 2018 Application of hybrid framework to facilitate lean six sigma implementation: A manufacturing company case experience. *Production Planning & Control* **29** (3), 185–201.

Yang, X.-S. 2021 Genetic algorithms. In: *Nature-Inspired Optimization Algorithms*. Elsevier, Amsterdam, The Netherlands, pp. 91–100.

Yue, J., Shen, Z., Xu, L. & Zhang, K. 2022 Numerical analysis the calcium leaching effect in concrete and concrete-rock interface on concrete gravity dam anti-sliding stability. *Case Studies in Construction Materials* **17**, e01158.

Zhang, X., Chen, X. & Li, J. 2020 Improving dam seepage prediction using back-propagation neural network and genetic algorithm. *Mathematical Problems in Engineering* **2020**, 1–8.

Zhang, H., Song, Z., Peng, P., Sun, Y., Ding, Z. & Zhang, X. 2021a Research on seepage field of concrete dam foundation based on artificial neural network. *Alexandria Engineering Journal* **60** (1), 1–14.

Zhang, K., Xu, L., Qiu, L., Tan, J., Yang, C. & Zhang, K. 2021b Numerical analysis of concrete gravity dam seepage characteristics evolution considering the calcium leaching effect. *Advances in Civil Engineering* **2021**, 1–21.

Zhang, W., Li, H., Li, Y., Liu, H., Chen, Y. & Ding, X. 2021c Application of deep learning algorithms in geotechnical engineering: A short critical review. *Artificial Intelligence Review* **54** (8), 5633–5673.

Zhou, M., Lin, F., Hu, Q., Tang, Z. & Jin, C. 2020 AI-enabled diagnosis of spontaneous rupture of ovarian endometriomas: A PSO enhanced random forest approach. *IEEE Access* **8**, 132253–132264.

First received 28 February 2023; accepted in revised form 28 June 2023. Available online 7 July 2023