



Prediction of karst spring discharge using LSTM with Bayesian optimisation hyperparameter tuning: a laboratory physical model approach

Portia Annabelle Opoku^{1,2} · Longcang Shu^{1,2} · Theophilus Ansah-Narh³ · Patrick Banahene⁴ · Kouassi Bienvenue Mikael Onan Yao^{1,2} · Albert Kwame Kwaw⁵ · Shuyao Niu^{1,2}

Received: 29 March 2023 / Accepted: 27 June 2023 / Published online: 5 September 2023
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

Abstract

Estimating spring discharge in karst aquifers is challenging due to non-linear and non-stationary hydrological processes caused by spatial and temporal variations. This study mimicked the phenomenon by simulating spring discharge using a laboratory physical model. The hydrological processes adopted in the simulation include systems such as infiltration, fissure-conduit, and drainage. We then recorded spring discharge and precipitation values from the simulated model along side the corresponding air temperature and humidity—in order to analyse the time series behaviour of the system. To estimate spring discharge from the simulation, a deep learning algorithm is developed taking temperature, humidity and precipitation as the input. In this work, the Bayesian optimisation was used to sweep through a range of hyperparameter values to search for the top 5 optimal training options for a Long Short Term Memory (LSTM) neural network. In addition, XGBoost was employed to identify the key predictors of spring discharge, resulting in enhanced predictability. The results show that LSTM-1, LSTM-2, LSTM-3, and LSTM-4 are the recommended recurrent neural network designs for predicting spring discharge using all three input parameters. LSTM-1, LSTM-2, and LSTM-3 network architectures are optimal for utilising two input variables: precipitation intensity and temperature. LSTM-5 has shown that a single parameter is inadequate for estimating spring discharge. The LSTMs yielded an RMSE value of ~ 0.04 , as well as a R^2 value of $\sim 98.01\%$. The study showed that using different input parameters, the suggested LSTM model can effectively simulate spring discharge in a karst environment.

Keywords Karst spring discharge · Laboratory physical model · Hydrological parameters · LSTM · Bayesian optimisation

Introduction

Karst springs are naturally occurring formations that serve as a principal water source for local communities around the world. The hydrological disposition of a karst system is mostly identified by the complex spatial and temporal variable processes of recharge, such as precipitation, surface runoff, infiltration, groundwater flow, conduits and fissure zones, as well as anthropogenic factors (Telesca et al. 2013; Yu et al. 2018). Figure 1 presents a 3D illustration of the hydrological behaviour of a karst aquifer. Generally, karst springs are a direct reflection of the state of groundwater in the aquifers that feed them, reflect the variability of groundwater storage in the basins, and have a direct impact on the streams and other surface water bodies into which they discharge, as well as all dependent ecosystems. Additionally, spring water replenishes the majority of rivers. These springs have significance for the survival of microhabitats in

✉ Longcang Shu
lcshu@hhu.edu.cn

¹ College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

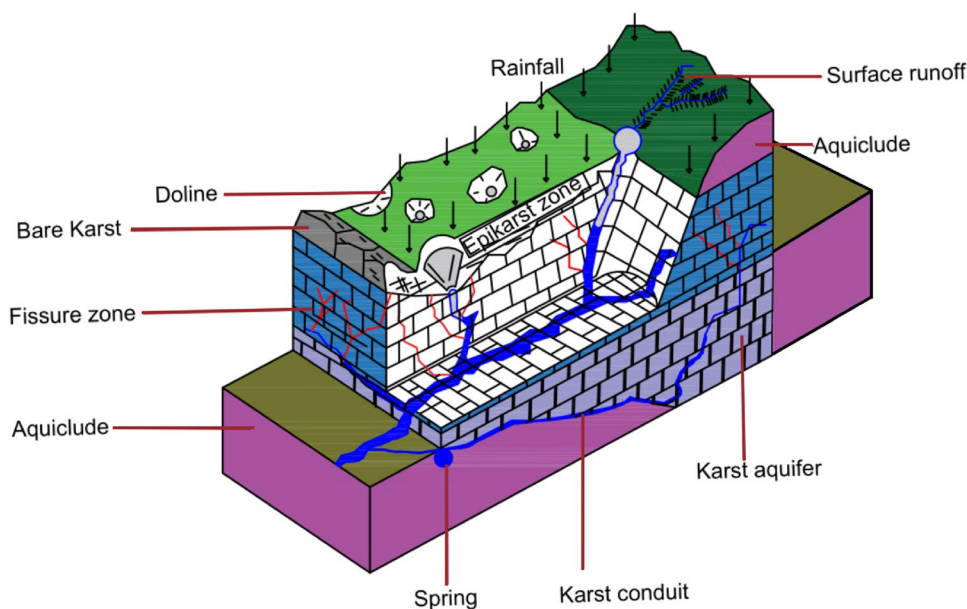
² State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China

³ Ghana Space Science and Technology Institute, Ghana Atomic Energy Commission, Box LG 80, Legon-Accra, Ghana

⁴ Key Laboratory of Integrated Regulation and Resource Development on Shallow Lakes, Ministry of Education, College of Environment, Hohai University, Nanjing 210098, China

⁵ School of Earth Sciences and Engineering, Hohai University, Nanjing 210098, China

Fig. 1 A schematic drawing of a heterogeneous karst aquifer system showing the primary source of aquifer recharge (rainfall) and recharge from runoff from non-karst rocks (surface runoff), infiltration, spring discharge outlet, porosity, and flow (characterised by conduits and fissures) (Ding et al. 2020; Williams 2009)



river systems (Hartmann et al. 2021; Ikard and Pease 2019; Kareem et al. 2022).

When the world's population distribution is compared to the distribution of carbonate rocks, karst groundwater accounts for approximately 20% of the Earth's land surface and can be accountable for about a quarter of the world's water supplies (Ding et al. 2020; Ford and Williams 2007; Parise et al. 2004). Carbonate rocks underlie 3×10^6 km² of the land surface in Europe and provide water for major cities, including Bristol, London, Paris, Montpellier, Rome, and Vienna (Andreo 2012; Biondić et al. 1998). In addition, karst water is crucial for the survival of more than 100 million individuals in South China (Ding et al. 2020).

Predicting spring flow in karstic aquifers is challenging due to their heterogeneous nature and non-linear relationship. Furthermore, insufficient observational data on spring flow hinders prediction studies and analysis. Overexploitation and inadequate water resource management have caused a spring dry-up. The discharge of springs has decreased due to increased abstraction to meet domestic, industrial, and agricultural water supply demands (Mohammed et al. 2022). The drying up of particular springs underscores the need for improved comprehension of regional groundwater systems and efficient management strategies to ensure the uninterrupted supply of groundwater resources and restore spring flow (An et al. 2019). Karst aquifers exhibit a range of conduits, cracks, and pores that possess hierarchical permeability structures and flow pathways. The Niangziguan Springs in China serve as an instance of aquifers (An et al. 2019, 2020).

Human activities, on the other hand, have decreased karst spring discharge. Previous research discovered that anthropogenic effects influenced or impacted karst hydrological

processes in the Niangziguan Springs catchment more than climate change effects (An et al. 2019, 2020). However, it is still unclear how human activity affects the magnitude and periodicity of spring discharge due to the strong non-linear trend and non-stationary fluctuations (An et al. 2019, 2020; Hao et al. 2016).

Physical-based numerical models are commonly used to forecast spring discharge. The governing equations are applied to compute spring flow, considering relevant assumptions and boundary conditions. These models require significant amounts of accurate data for parameter calibration. In certain uncharted regions, acquiring precise and adequate data presents challenges, leading to suboptimal model efficacy and increased uncertainty (Hassanzadeh et al. 2020; Khorrami et al. 2019; Milly et al. 2008; Tavakol-Davani et al. 2019). However, owing to the complexities caused by random uncertain boundary conditions of precipitation, extreme karst aquifer heterogeneity, hydrological modelling errors such as initial conditions, model parameterisation, structure, and anthropogenic impacts, interpreting and predicting spring discharge data using only a physical model is a challenge for understanding the contributions of each component (Khorrami et al. 2019; Quan et al. 2022; Tavakol-Davani et al. 2019; Hartmann et al. 2021; Jeannin et al. 2021).

Various prediction models, including statistical learning, artificial neural networks (ANNs), and grey systems have been developed and applied to address uncertainties in spring flow simulation and solve practical issues (Alameer et al. 2020; Quan et al. 2022). Therefore, in this work, the prediction model adopted the techniques in ANN.

ANNs are training based-models designed to have a functional mapping between input predictor and output

prediction variables when relevant historical big data is processed (Anand and Oinam 2022; Roshani et al. 2022). This approach has proven to produce more accurate predictions of hydrological events than the traditional calibrated numerical hydrological models (Sharafati et al. 2020; Akano and James 2022), although they are unable to explain the underlying processes, being data-driven rather than process-driven models. This provides an opportunity for improving prediction accuracy by coupling with other robust schemes such as Bayesian optimisation.

ANNs have been employed by several researchers to model and forecast non-stationary hydrological processes (He et al. 2020; Meng et al. 2019; Niu and Feng 2021; Singh et al. 2023; Xie et al. 2019; Yan et al. 2021; Khorram and Jehbez 2023; Haggerty et al. 2023). Optimal use of ANNs in karst aquifer systems can improve model prediction accuracy. An ANN typically consists of two primary components: The network comprises numerous interconnected artificial neurons. These links extract relevant parameters from a dataset. An artificial neural network (ANN) model does not require a data distribution assumption prior to feature engineering. ANN finds applications in diverse domains. RNNs are more appropriate for modelling dynamic systems than feed-forward neural networks, as Bailer-Jones et al. (1998) and Tsung (2010) suggested. The issue frequently encountered when training RNNs with numerous hidden layers is the exploding gradient problem, as Hochreiter identified in 1998. The problem arises during the weight adjustment of the RNN for gradient computation through backpropagation (Hochreiter 1998).

To address the problem of exploding gradient in RNN, long short-term memory (LSTM) network was developed using the activation functions of the forget gate in addition to the cell state (Sahar and Han 2018; Thonglek et al. 2019). LSTM, a type of artificial neural network, has feedback connections, making it ideal for data processing and prediction. Recent studies by hydrologists (Yang et al. 2019; Yin et al. 2021) have exhibited the potential of LSTM in the application of hydrological sciences. For example, Kratzert et al. (2018) used LSTM to model rainfall-runoff because of its ability to learn and store long-term dependencies between the provided input and output of the network. Using the forget gate and cell state activation functions, long-short-term memory (LSTM) networks were developed to limit gradient explosion in RNNs (Sahar and Han 2018; Thonglek et al. 2019).

LSTM is an artificial neural network with feedback connections, making it ideal for data processing and prediction. Recent studies by hydrologists (Yang et al. 2019; Yin et al. 2021) have exhibited the potential of LSTM in applying hydrological sciences. For example, Kratzert et al. (2018) used LSTM to model rainfall-runoff because of its ability to learn and store long-term dependencies between the

provided input and output of the network. They found that LSTM produced better model performance than the traditional RNNs. Like other machine learning (ML) models, RNN models may have several vital hyperparameters that must be optimised to improve model performance (Cheng et al. 2019). As a result, hyperparameter optimisation is a critical step (Ma et al. 2020). Therefore, Bayesian Optimisation (BO) was used in this study to set the network size. BO is more appropriate when objective functions are computationally expensive to calculate due to its higher efficiency compared to alternatives such as genetic algorithms (Alizadeh et al. 2021; He et al. 2019; Lin et al. 2020).

This study aimed to achieve two objectives to comprehend the karst spring discharge in the Jinan basin. A 3D physical laboratory model was put up to replicate the Jinan spring basin (see Fig. 2). Subsequently, rainfall-runoff experiments were performed to collect data on spring discharge. Second, LSTM coupled with BO was used to produce the best LSTM-RNN model that predicts the spring discharge in the simulated area using precipitation intensity, temperature, and humidity. Statistical metrics (the root mean square error (RMSE), the mean absolute error (MAE), the mean absolute percentage error (MAPE), Pearson product-moment correlation coefficients (ρ), and the coefficient of determination (R^2)) were used to assess the performance of the top 5 LSTM models developed with different input combinations.

Materials and methods

Study area

The Jinan Karst System shown in Fig. 2, is located in the midwestern Shandong Province of Northern China. The springs are important historical sites in China. Jinan City has been known as “Spring City” since ancient times due to its fractured-karst springs. However, due to Jinan’s social and economic development, urbanisation and over-exploitation have negatively impacted the level of karst groundwater in recent years, leading to a substantial decrease in the spring flow rate (Zhu et al. 2020).

Jinan is abundant in easily dissolvable Cambrian and Ordovician carbonate rocks. Because of the geological structure, lithology, climate, and other factors, many karst fractures and pipelines have been produced, offering a favourable conduit for groundwater transport and ample storage area. Despite safeguards such as limited mining and artificial groundwater recharge in Jinan, the spring is still in danger of drying out (Zhang et al. 2018).

The main problem is the difficulty in the prediction of spring flow in the field. Hence, there is not enough observation data on spring flow data. Even the data available is insufficient to make any meaningful inference. This

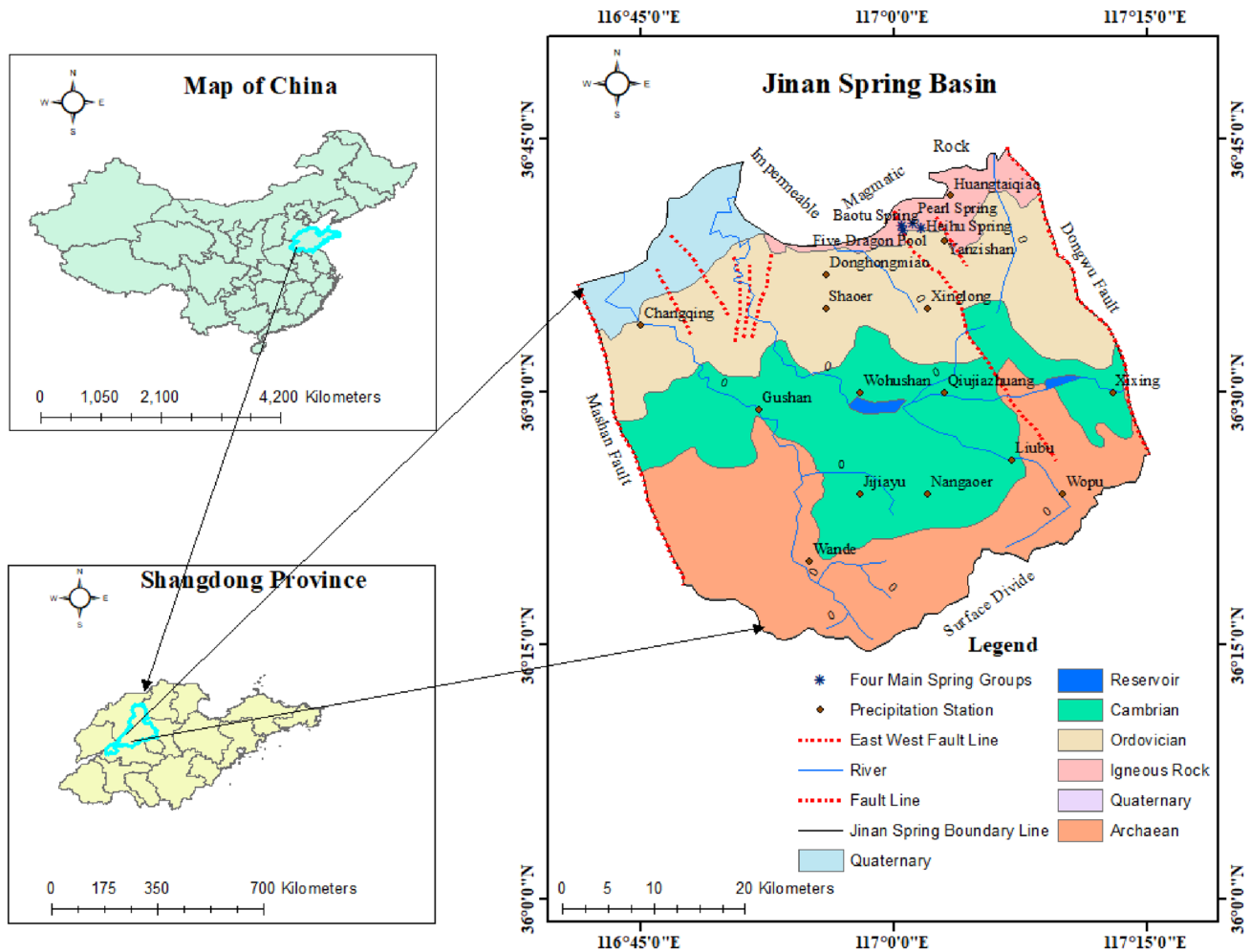


Fig. 2 A geographical map of the Jinan spring basin showing the four main spring groups (Baotu, Heihu, Five Dragon, and Pearl) and the Dongwu and Mashan Faults

is due to the highly heterogeneous nature and non-linear relationship of karstic aquifers (Mohammed et al. 2022). Therefore, this work used the physical model to aid in the simulation of karstic spring flow predictions.

The laboratory physical model

The 3D physical model can simulate surface runoff, pore water runoff, and fissure karst water runoff in karst areas. The model comprised a test sand tank, precipitation replenishment system, mulch, rift karst zone medium, water level monitoring system, and flow monitoring system. The upper infiltration system represented the unsaturated zone, while the lower fissure-conduit system represented the saturated zone. Figure 3 displays a schematic representation of the physical model. The labelled parts are shown in Table 1.

Experimental set-up

The physical model can investigate surface runoff, pore water runoff, and spring discharges in the karst region. The study enhances the theoretical understanding of karst water flow and establishes a scientific basis for practically analysing complex karst groundwater resources. It is crucial for developing and utilising water resources in karst areas. Figure 4a–k illustrates the sequential process for setting up the laboratory experiment.

Figure 4a displays the test sand tank. The sand tank is a micro-sloped box that acts as the model domain and is inclined at a 2° tilt. To control the initial conditions, allow all the water to drain thoroughly after each experiment. The tank has multiple outflow holes with no flux boundaries for monitoring runoffs and water levels. The inside and outside of the tank walls are coated with impermeable waterproof paint, preventing water seepage outside the tank's domain.

Fig. 3 Schematic diagram of the physical laboratory model

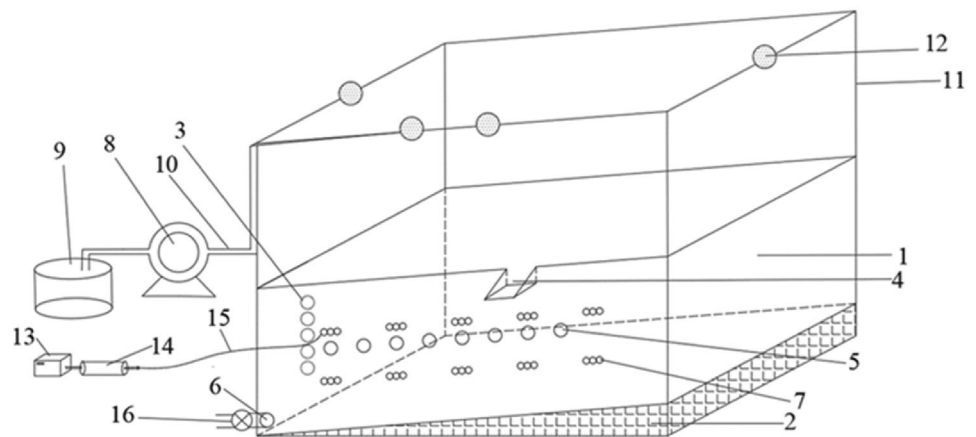


Table 1 The parts of the schematic diagram of the laboratory physical model

1. Test sand tank with dimension ns of $2.5 \times 1.3 \times 0.9$ m	9. Water tank
3. Bottom of the tank with an inclination of about 2°	10. Rubber hose (for transferring simulated Rainfall to the sand tank)
4. Side face aperture (diameter of 0.02 m) for controlling the initial water levels	11. Bracket or boundary of a sand tank
5. Surface runoff outlet	12. The nozzle of the hose pipe (for spraying simulated Rainfall)
6. Pore water outlets	13. Paperless recorder
7. Fissure karst water outlet (spring discharge)	14. Pressure sensor
8. Pressure measuring holes	15. Rubber tube
9. Variable frequency pump (to simulate Rainfall)	16. Flux flow meter

It also has a surface flow outlet at the top of the sidewall to measure surface runoff and a spring discharge hole at the bottom, connected by conduits.

The fissure-conduit system is constructed of 10 cm concrete cube bricks. Layer one (see Fig. 4b) contains two connected conduits with a cross-sectional area of 30 cm^2 serving as the fracture zone and a hydraulic conductivity of 0.05 m/d. The bricks are arranged systematically with a small gap of 2.5 mm in between them. Construction consists of the pore, fissure, and pipe layers. These segments simulate the flow as it passes through a series of water-conducting fissure zones and is then eventually discharged through two large fissures at the bottom. The second layer in Fig. 4c, showcases the fissure zone designed to aid flow movement. For the third layer (in Fig. 4d) two conduits with dimensions of $5.2 \times 146 \times 10$ cm and $5.4 \times 153 \times 10$ cm are designed to mimic the fracture zone. As depicted in Fig. 4e multiple conduits were constructed to improve the drainage system in the installation. In layer 5 (refer to Fig. 4f), a piece of cloth is added to the set-up to block sand from filling up the space created.

The experiment also included an infiltration system to simulate vegetation cover and the soil profile. Figure 4g shows how the soil layer is loamy and sandy. The loamy soil

has an average particle size of $57 \mu\text{m}$, and that of the sand is $753 \mu\text{m}$. The hydraulic conductivity of this layer is approximately 24 m/d, and the specific yield is 0.24. Additionally, the state of saturation was measured using the pressure sensor probes in layer 6.1. Layer 6.2 in Fig. 4h shows how the soil layer with buried pressure tubes is wholly covered. The vegetation cover is simulated using a cloth designed with cotton-like material to cover the soil layer with buried pressure tubes. Figure 4i displays the set-up of the simulated vegetation cover. Figure 4j shows the various positions of the passage for measuring surface runoff and pore water runoff conduits for the developed physical model.

Figure 4k shows the complete experimental set-up for this study. The precipitation recharge system comprises a water tank, a variable frequency pump, hosepipes, and four nozzles to simulate the process of natural rainfall supplying the surface karst region of the area. The flow rate is recorded at different times to measure the intensity of the rain.

Data description

Data from the unsaturated test scenario was used for the analysis of this study. Short-term rainfall was performed in the sand tank area utilising a frequency pump and a

Fig. 4 Experimental set-up of the physical model: **a** shows the sand tank with multiple outflow holes; **b–f** are the fissure conduit system at various layers; **g–j** describes the infiltration systems to mimic the vegetation cover and the soil profile, and **k** is the complete lab set-up for the experiment

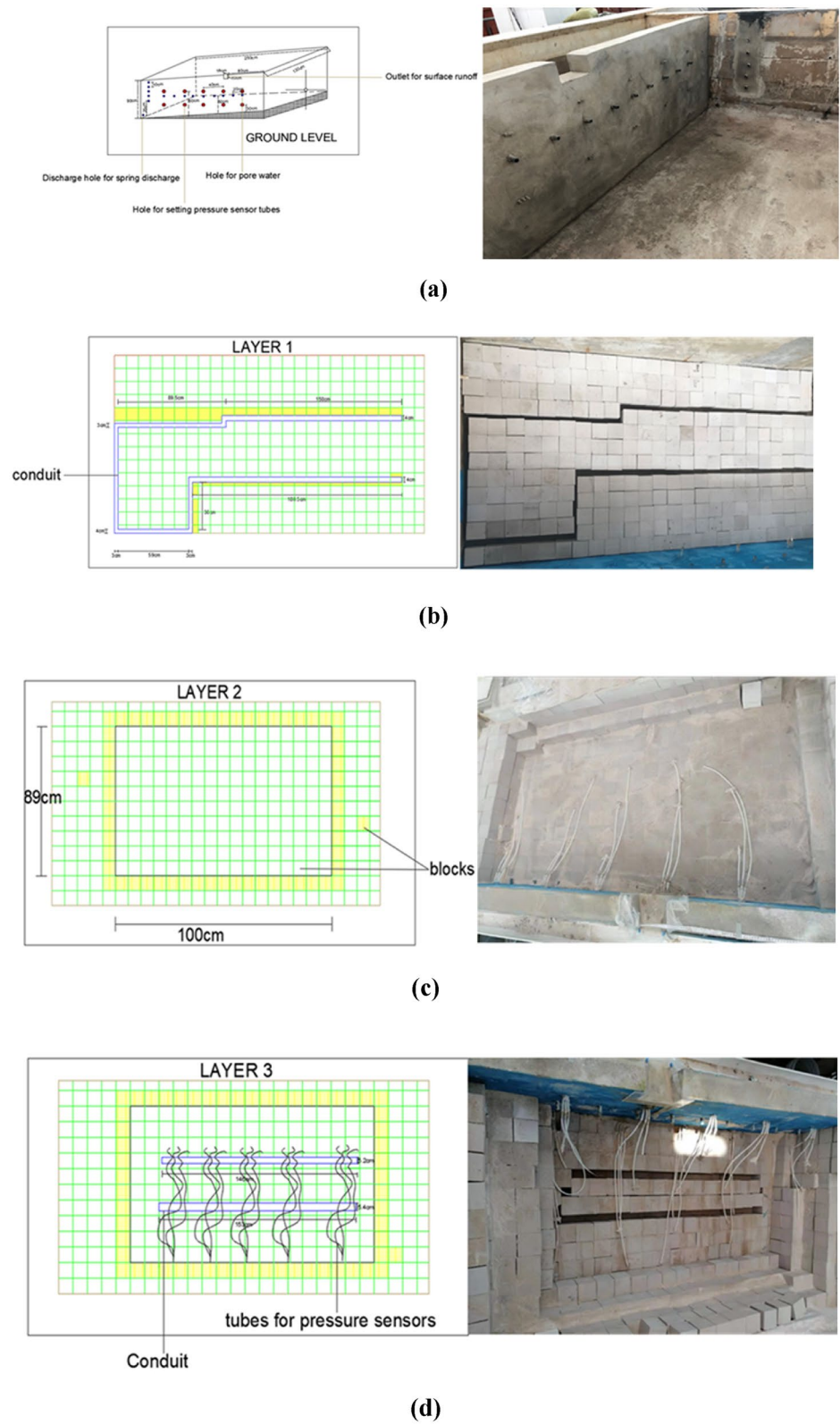
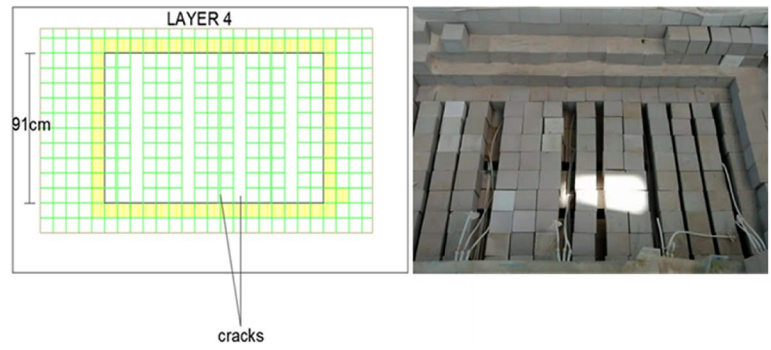
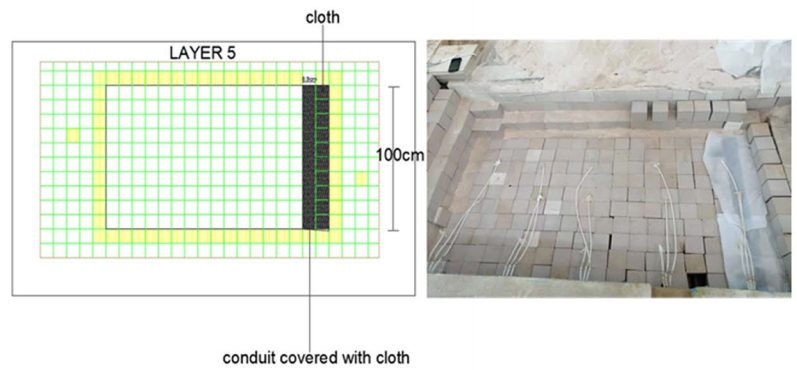


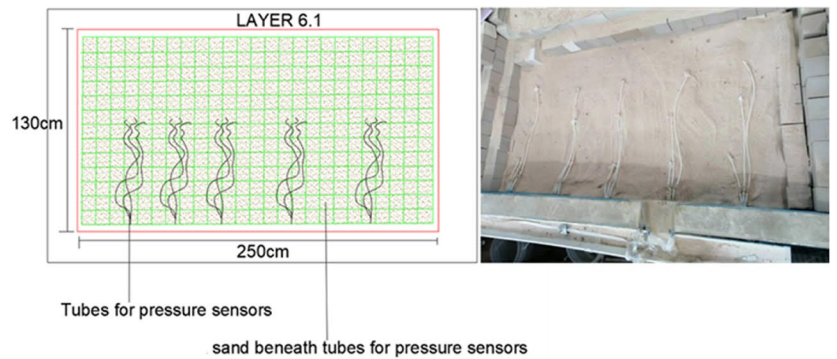
Fig. 4 (continued)



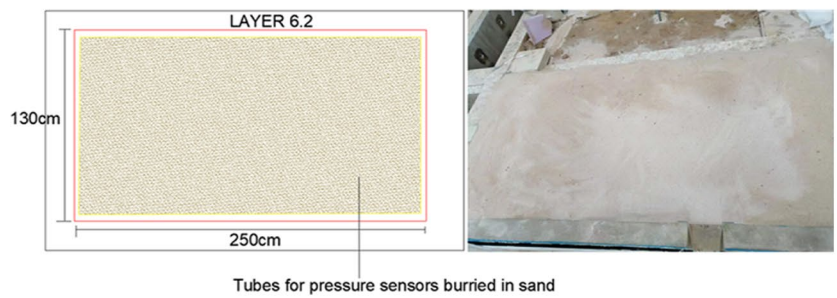
(e)



(f)

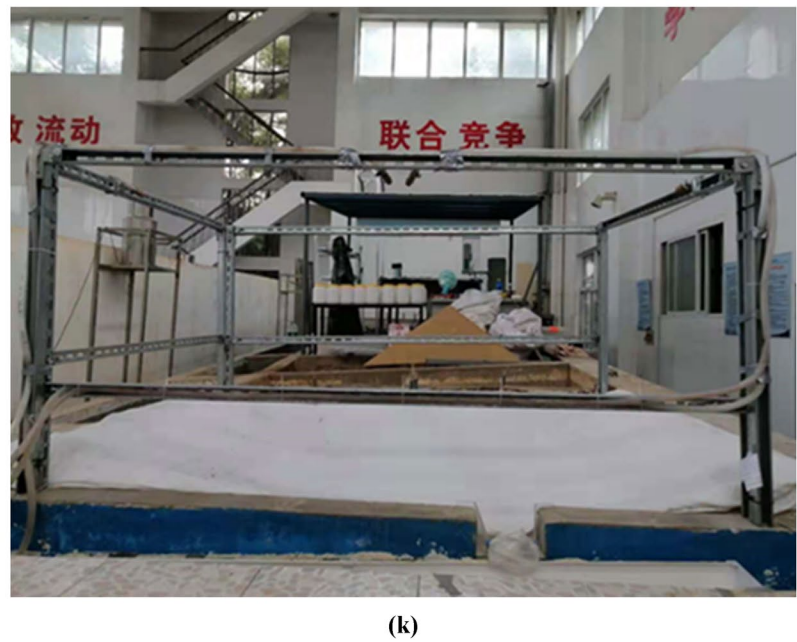
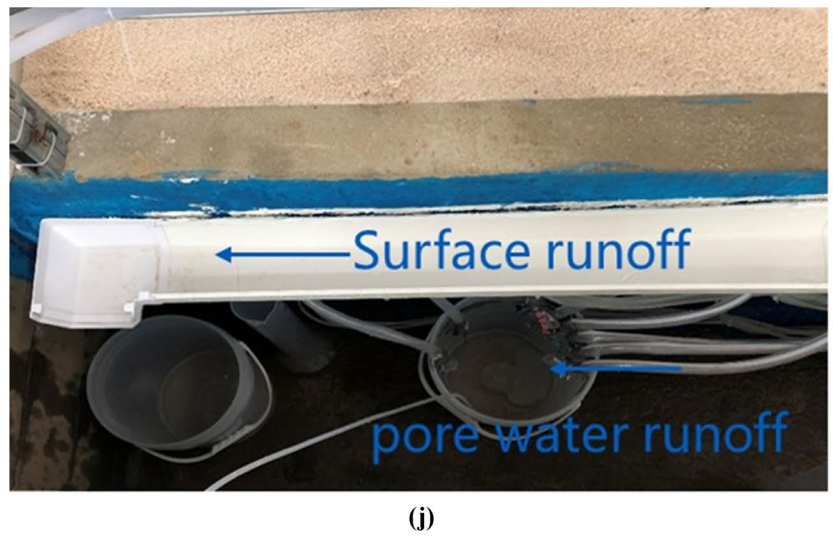
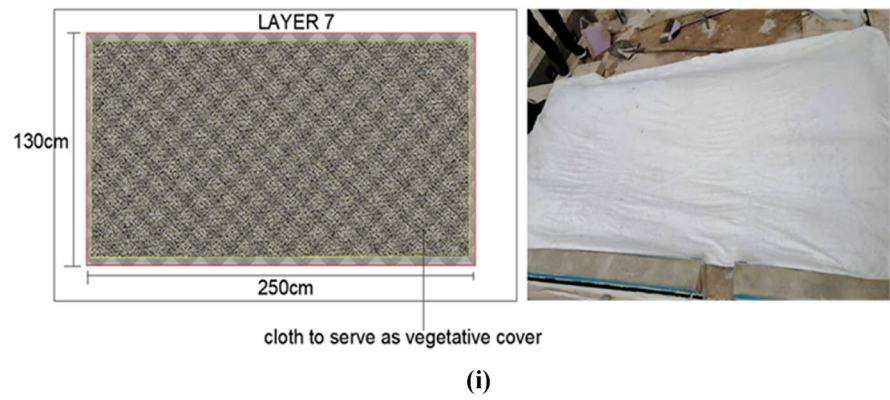


(g)



(h)

Fig. 4 (continued)



spraying apparatus (mimicking rainfall). Using the pump, the intensity of the rain varied at different rates. Surface runoff and fissure karst water runoff (spring discharge) were measured using measuring cups, stop clocks, and a flux flow metre.

The descriptive statistics of the research laboratory simulation are shown in Table 2. 3968 measurement points were obtained in total. The discharge rate, which was used as a response variable in this study, had a mean value of about 37.59 mL/s and ranged between (0.42, 153.51) mL/s. To understand the fluctuations in recorded discharge levels, hydrological parameters such as precipitation intensity, temperature, and humidity were investigated. The

expected values for these stated variables were 0.024 mm/s (precipitation intensity), 31.21 °C (temperature), and 0.61% (humidity).

Also, as shown in Fig. 5, at a precipitation intensity of 0.01755 mm/s, the maximum spring discharge rate was about 78 mL/s, with a median of about 36. Also, at 0.02017 mm/s, the maximum spring discharge rate was about 96 mL/s with a median value of about 9. Furthermore, at a precipitation intensity of 0.02279 mm/s, the maximum spring discharge rate was about 115 mL/s with a median of about 51, while 0.02632 mm/s gave a maximum spring discharge rate of about 152 and a median of about 36. The highest precipitation intensity of 0.02925 mm/s had a maximum spring

Table 2 Descriptive statistics of spring discharge and independent variables (i.e., precipitation intensity, temperature, and humidity)

	Discharge rate (mL/s)	Precipitation intensity (mm/s)	Temperature (°C)	Humidity (%)
Count	3968.00	3968.00	3968.00	3968.00
Mean	37.5925	0.0235	31.2064	0.6092
Std	35.7408	0.0039	1.7453	0.0478
Min	0.416667	0.017550	27.000000	0.510000
25%	2.239985	0.020169	30.400000	0.570750
50%	31.449450	0.022788	31.500000	0.610000
75%	66.162571	0.026320	32.500000	0.645000
Max	153.508772	0.029250	34.500000	0.710000

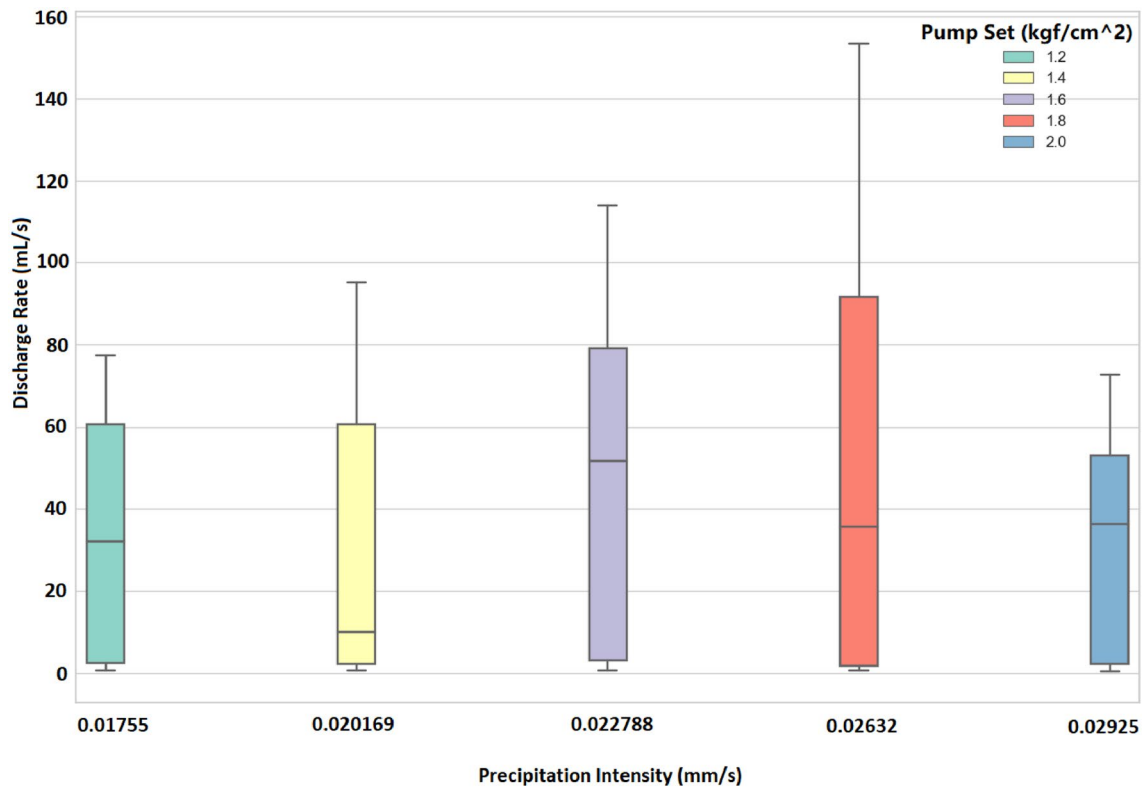


Fig. 5 A univariate box and whisker plot to show the data distribution in a single graph

discharge of about 74 with a median of about 37, less than the values recorded by the least precipitation intensity rate. This was due to the material composition of the porous layer (loam and sand), which showed that not all Rainfall could infiltrate the system and further percolate into the aquifer, owing to the high rainfall intensity. The porous layer got fully saturated quickly, and therefore, overland flow (i.e., surface runoff) was higher because the porous layer could not absorb all the water coming to it instantly for this particular test.

The horizontal axis is the precipitation intensity in mm/s, and the vertical axis is the spring discharge rate in mL/s. The different colours represent different pump set values [1.2, 1.4, 1.6, 1.8 and 2.0 (kgf/cm²)] which gave rise to the respective precipitation intensities, which can be read on the vertical axis.

Due to this phenomenon, data collected on spring discharge with the highest precipitation intensity of 0.02925 mm/s was minimal and had the lowest peak, as shown in the spring hydrograph, as compared to the other precipitation intensity values of 0.01755, 0.02017, 0.02279, and 0.02632 mm/s; because less Rainfall infiltrated into the system due to rate of infiltration as a result of material composition. Hence, a threshold value must be set to manage the spring's inflow and outflow to have a perennial flow, if possible.

The correlation matrix presented in Fig. 6 assesses the strength and direction of the linearity between the two explanatory variables. Strong positive linearity values show that the explanatory variables measure the same characteristics. Figure 6 shows a moderate to low correlation—indicating that the selected variables may measure different characteristics.

Feature importance using extreme gradient boosting

The concept behind a decision-tree-based algorithm is to build a sequential-like tree diagram, where each inner node depicts a test on a feature, each subdivision denotes a test score, and each terminal node holds the regressor target value. The main idea is to generate a regressor using input features ($f_1, f_2, f_3, \dots, f_i$) and at least one target feature (Y). In this study, the input features were precipitation intensity, temperature, and humidity, and the output feature was the spring discharge rate. The learning pattern of a decision tree was developed by splitting the data set into multiple sets based on the feature value test. The splitting method was done recursively on each multiple set. The recursive partitioning process was completed when all of the different sets at a node recorded the exact score of the output feature. When recursive splitting on many sets

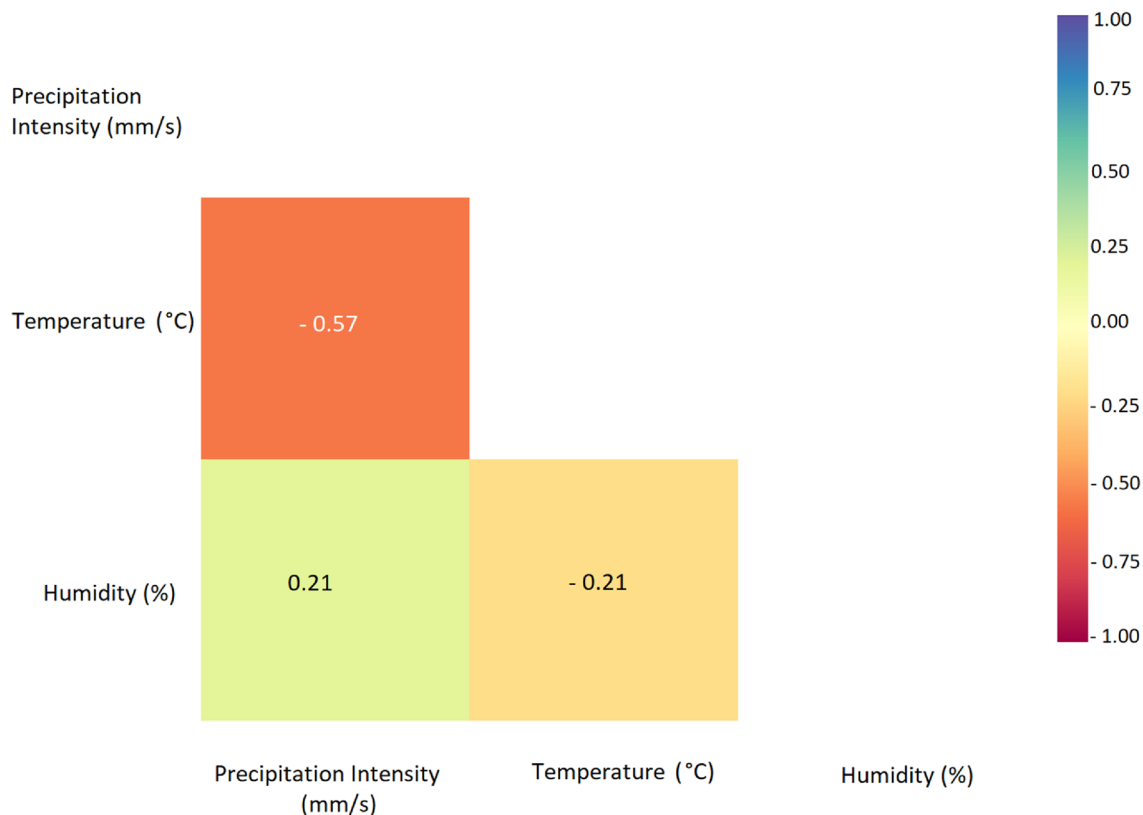


Fig. 6 The correlation matrix of all independent variables is depicted as a heat map

does not improve the prediction score, the learning process may be discontinued.

The algorithms mostly adopted to create ensemble decision trees included bagging and boosting methods (Galar et al. 2012; Quan et al. 2022; Zhou et al. 2010). The former method uses a uniform weak learner by splitting the training set into multiple sets using the bootstrap sampling technique and combining the individual sub-sample based on a deterministic averaging process. The latter method integrates many weak learners to produce a robust learner iteratively. Each weak learner after that is taught using a new sub-sample of data that includes measurements that were not effectively addressed by the previous learners. The ultimate goal is to reduce the regressor’s bias by encouraging the weak learner to rely on more difficult data.

The XGBoost is a supervised learning technique based on the boosting algorithm (Chen and Guestrin 2016), which approximates a function through optimisation and regularisation. The unique nature of this method is how well it can distinguish the relevant features between selected variables and the ability to perform hyper-threading of a computer processor to improve the computational cost. It has been utilised successfully in a variety of practical applications, including stock selection (Li and Zhang 2018), disease diagnosis (Bao 2020), and time series forecasting (Zhai et al. 2020). The decision-tree-based model used in the XGBoost method adopts an iterative approach until the stopping criteria are satisfied. The following equations are used to produce the predicted test values:

$$\hat{Y} = \sum_{\omega=1}^{\kappa} g_{\omega}(f_i), \quad g_{\omega} \in S \tag{1}$$

where \hat{Y} denotes the predicted test value, g_{ω} represents an individual regression tree, f_i is the input feature, κ is the number of regression functions, and S is the space of all possible cases of $g_{\omega}s$.

Figure 7 shows the basic architecture of the XGBoost supervised learning method. The regression model can be optimised by identifying the optimal node for subdivision to minimise the objective function described in Eq. (2). It is imperative to determine the most favourable node to subdivide to minimise the objective function defined in Eq. (2) to fine-tune the regression model as;

$$Q^t = \sum_{j=1}^n l(Y_j, \hat{Y}_j^{t-1} + g_t(f_j)) + \Theta^*(g_t) \tag{2}$$

where $l(\cdot)$ represents the loss expression depicting how well the estimated test value approximates the actual test value, $g_t(\cdot)$ is the predicted outcome of the decision tree at the t th iteration and $\Theta^*(\cdot)$ denotes the model regularisation function. The regularisation function consists of the number of leaf nodes T and the corresponding weights w such that:

$$\Theta^*(g_t) = \beta T + 0.5\varphi \|w\|^2; \tag{3}$$

The parameter β represents the complexity cost of adding more leaves to the decision tree, whilst that of φ is the regularisation coefficient. The coefficients are modulated using the L2 norm.

Concerning this work, the specific hydrological features that influence spring discharge was determined using the XGBoost. Therefore, the bar chart in Fig. 8 indicates that precipitation intensity has a higher impact on the model in

Fig. 7 The principal decision-tree design of the XGBoost algorithm

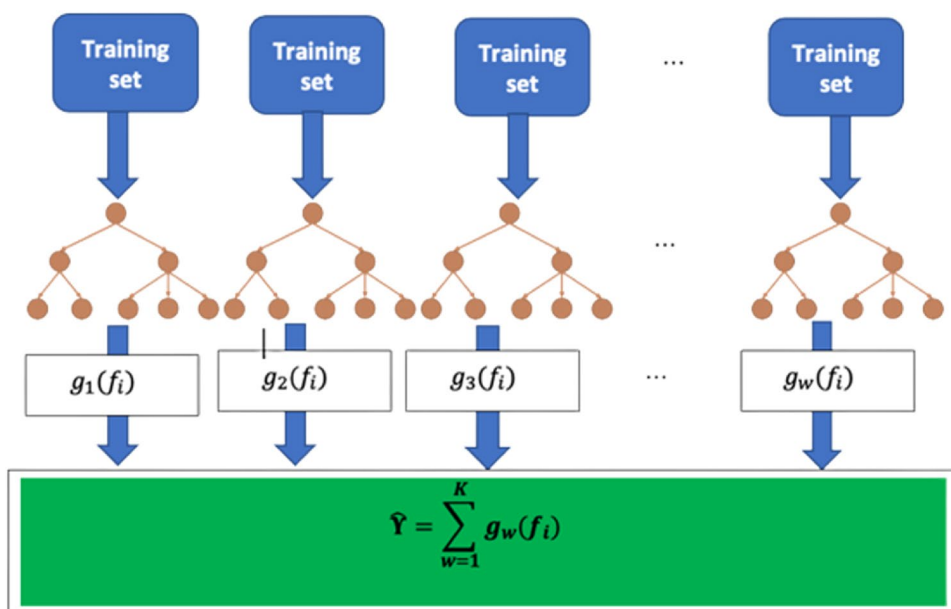
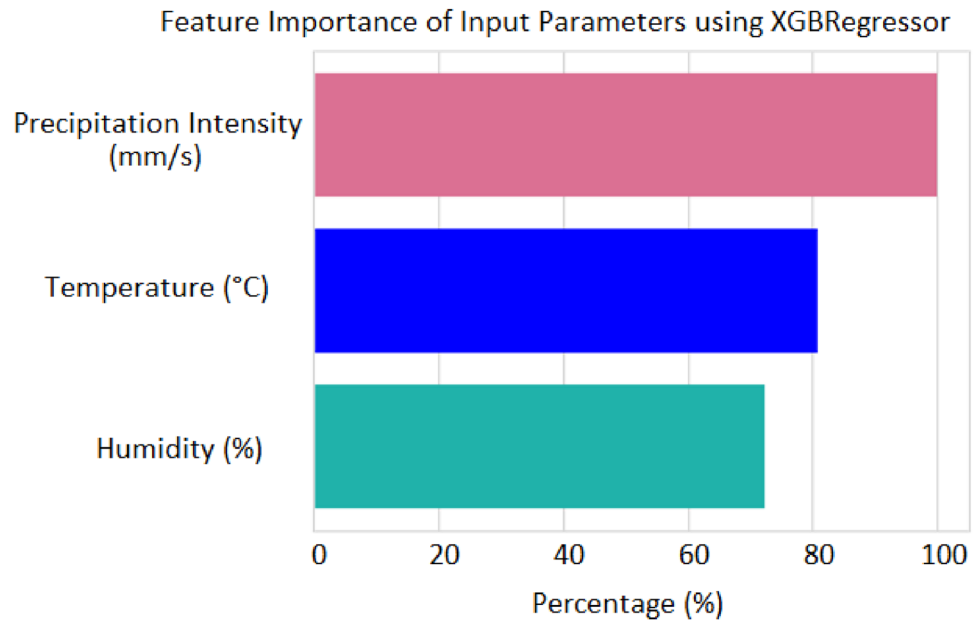


Fig. 8 A plot of the feature importance of input variables using the XGBoost Regressor



predicting spring discharge than temperature and humidity. Thus, in this work, we looked at three significant scenarios:

1. Estimation of the spring discharge rate using all three hydrological features as input variables;
2. Estimation of the spring discharge rate using only the top two features as input variables; and
3. Estimation of the discharge rate using only the highest feature as an input variable.

LSTM with respective gates

The LSTM is an improved model of the conventional RNN. It was initially introduced by Hochreiter and Schmid Huber

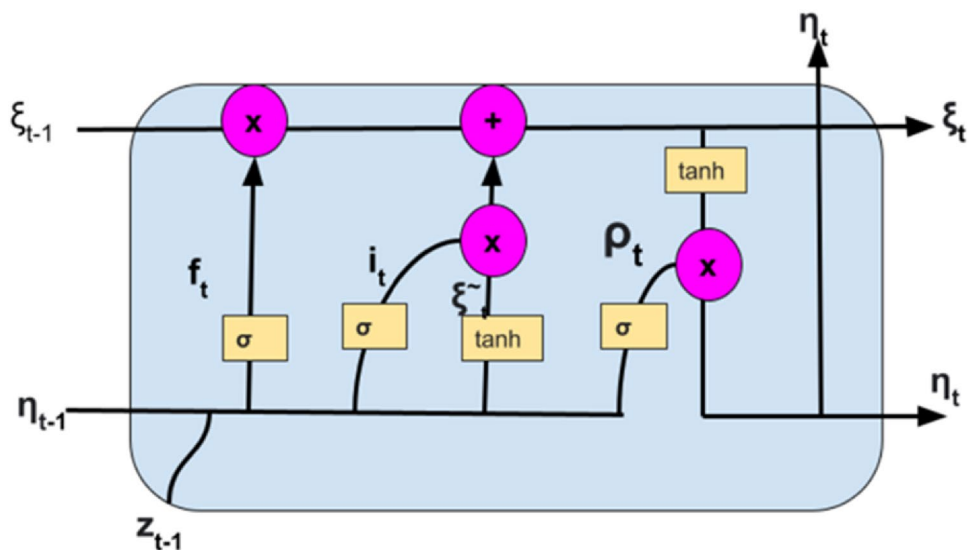
(1997) and operates using a cell and three gates (the input, the output, and the forget gates). These memory components regulate the message inflow inside the network unit. Figure 9 illustrates how a typical LSTM unit processes information using the input gate (i_t), the forget gate (f_t), the output gate (ρ_t), and the cell state (ξ_t). The mathematical formulations of the LSTM unit of these memory parts are expressed as:

$$i_t = \sigma(z_t \Lambda^i + \eta_{t-1} \Omega^i) \tag{4}$$

$$f_t = \sigma(z_t \Lambda^f + \eta_{t-1} \Omega^f) \tag{5}$$

$$\rho_t = \sigma(z_t \Lambda^\rho + \eta_{t-1} \Omega^\rho) \tag{6}$$

Fig. 9 The general architecture of the LSTM model



$$\tilde{\xi}_t = \tanh(z_t \Lambda^g + \eta_{t-1} \Omega^g) \quad (7)$$

$$\xi_t = \sigma(f_t * \xi_{t-1} + i_t * \tilde{\xi}_t) \quad (8)$$

$$\eta_t = \tanh(\xi_t) * \rho_t \quad (9)$$

where, σ is the standard deviation, Λ and Ω are weight vectors for f , ξ , i , and ρ for a given time-step t . The hidden unit (η) studies which info to store and which to discard.

The input gate process is described in Eq. (4). Given new input data (z_t) and the preceding hidden state (η_{t-1}), the process here is to search for new information to add to the model's long-term memory (also known as the cell state). The forget gate presented in Eq. (5) is used to decide which part of the information in the cell state is useful when you are provided with both the preceding hidden state and the new input data. To achieve this, the preceding hidden state and the current input data are supplied into the model network to generate a series of elements using the sigmoid activation function.

The network considers the elements that are close to 0 as insignificant whilst, those close to 1 as relevant. The activated values are then multiplied element-wise with the preceding cell state such that the numbers close to 0 will have less effect on the following processes. The output gate is presented in Eq. (6), and it decides on the new hidden state. The expressions above show that the LSTM structure uses a set of gates to modulate how a sequence of instructions flows into the system, how they are gathered and eventually, how they leave the network.

Hyperparameter selection based on Bayesian optimisation (BO)

One of the key challenges in designing a neural network is knowing the right number of network sizes to use for a specific task. The network size includes the number of layers in the network architecture, the number of nodes in every layer, and even the number of connections to be adopted in the design. A neural network is mostly designed to map a target variable x onto a non-linear function $\psi(v)$. The non-linear function $\psi(v)$ is gleaned during the training phase where the network tries to study the appropriate way to link the input variables v to the output variables x .

However, building a proper network entail fine-tuning the hyperparameters. The common hyperparameter tuning techniques are the grid search (Ghawi and Pfeffer 2019; Quijano et al. 2021; Shekar and Dagnew 2019) and the random search (Bergstra and Bengio 2012; Mantovani et al. 2015) algorithms. The former search algorithm works by setting up a framework of hyperparameter values such that for every combination in the framework, the network

is trained and evaluated on the test data. This makes the searching approach in selecting the best network to be very exhaustive. By contrast, the framework set in the latter search works by choosing random combinations to train the network and evaluate. Nevertheless, even though the latter approach allows for explicit control in selecting the number of parameters to use, there is the possibility of not finding an accurate model.

One state-of-the-art hyperparameter optimisation algorithm introduced by Li et al. (2018) is the hyperband. This optimisation algorithm is based on a successive halving technique known as the bandit method (Li et al. 2018), where more credits are assigned to hyperparameter configurations with very high performance. Here, the configurations with low performance are less likely to predict the best model and are removed in the process to allocate more weights to hyperparameter configurations with good performance. Furthermore, this algorithm ensures early stopping in a principled way—making it more robust than the regular grid and random search schemes.

One major drawback in the previous optimisation techniques is that the hyperparameter settings are treated independently. Meanwhile, the space for these settings is mostly even; therefore, prior knowledge of the previous performance might help improve the performance of the next space. The Bayesian optimisation (BO) method addressed this issue and ensured homogeneity in the hyperparameter space. BO is a probability distribution model that tries to understand complex objective functions based on prior observations. As a result, in this work, an LSTM model was generated using BO to set the network size. Recent studies (Bergstra et al. 2011, 2013; Greenhill et al. 2020; Shahriari et al. 2016; Turner et al. 2021) show the potential and robustness of the BO method compared to the previously mentioned algorithms.

Bayesian optimisation for a classic LSTM

In the prediction of spring discharge values, a simple classic LSTM neural network might be adequate, but this same network can perform poorly if parameters such as the learning rate, the momentum term, the number of layers, the number of hidden units for different layers, and even the batch size are not set correctly. As a result, the units for the LSTM and dropout, the activation type and the learning rate of the model optimiser were chosen as hyperparameters for optimisation in this work. Fine-tuning these parameters was essential because the LSTM units handle the dimensionality of the output vector, whereas the activation function adds complexity to the model, making it non-linear. Also, to improve the performance of the LSTM model and avoid over-fitting, the learning rate must be optimised and some

network dropouts must be performed. Table 3 shows the parameter sets for fine-tuning.

By fine-tuning, the aim was to discover the appropriate hyperparameter for the proposed LSTM algorithm to produce the best performance when measured on the test data. Additionally, during the tuning process, the objective function usually becomes very complex to compute (Doke et al. 2020). This is because, to obtain a new set of hyperparameters, the neural network must be trained from scratch to determine its performance. Furthermore, calculating the weights of the hyperparameters may be difficult, leaving retraining the model as the only option. As a result, the BO algorithm was used to solve the problem and ensure that the model network learns from the hyperparameter settings. Based on the hyperparameter settings provided, this algorithm generates a probability distribution function (also known as a surrogate model) to compute the probability of obtaining the best score for the objective function.

D was considered to be the independent input data in predicting the spring water discharge, f to be the performance model, ξ to be the hyperparameter sets displayed in Table 3, and $\hat{\xi}_\theta$ to be the optimal parameter, then from Bayesian inference (Matsubara et al. 2021) (Dempster 1968), the likelihood function of the sample data is given as;

$$L(\xi) = f(D|\xi) = \prod_{j=1}^n f(D_j|\xi), \tag{10}$$

Note that Eq. (10) holds based on the assumption that D obeys a certain probability density distribution with parameter say θ . Therefore, the maximum likelihood estimation in Eq. (10) becomes;

$$\hat{\xi}_\theta = \operatorname{argmax}_{\theta \in \xi} f(D|\theta) \tag{11}$$

where *argmax* is an operation that finds the argument that gives the maximum value from a target function. Assuming the surrogate model obeys a Gaussian process (GP) prior, then the operation of Bayesian optimisation is depicted below:

- (a) Define an objective using a GP prior model: The GP is normally an acceptable model to estimate a non-linear

Table 3 Tabular description of hyperparameters (LSTM units, dropout units, activation function and learning rate) used in the learning algorithm with their respective min, max, step size and type functions

Hyperparameter	Min	Max	Step size	Type
LSTM units	32	512	32	–
Dropout	0	0.9	0.1	["True", "False"]
Activation function	–	–	–	["relu", "tanh"]
Learning rate	e^{-4}	e^{-2}	–	"log"

regression output which is exactly the purpose of this work. The model assumes that input data sampled from a distribution with unknown parameters produces an output similar to that distribution, resulting in a continuous objective function. As a result, because the hyperparameter settings in Table 3 are non-parametric, using the GP as a prior model for this analysis was the ideal choice. Following that, a brief review of GP for non-linear regression was provided. For an in-depth discussion on GP theory, refer to these studies (Rasmussen 2004; Rasmussen and Williams 2005; Snelson 2007).

Typically, a GP regression uses a joint Gaussian density function to map an input observation onto its corresponding target variable. It is characterised by an expectation function $m(v)$ and covariance function $\Sigma(v, v')$ such that $f(v) \sim GP(m(v), \Sigma(v, v'))$. Let define $\Lambda = \{v_j, \Phi_j\}_{j=1}^K$ to be a set of measurements, where, $v_j \in R^n$ are n measurement points in n -dimensions and $\Phi_j \in R$ are the matching targets. Mathematically, we can express this measurement as a supervised learning model to have;

$$\Phi_j = \Phi(v_j) = f(v_j) + \epsilon, \epsilon \sim N(0, \sigma_k^2) \tag{12}$$

Taking $f(v_j)$ as a GP regression, then from (Boyle 2007) we can represent f as;

$$f_{gp} = \begin{bmatrix} f(v_1) \\ \vdots \\ f(v_K) \end{bmatrix} \sim \mathcal{N} \left(0, \begin{bmatrix} \Sigma(v_1, v_1) & \dots & \Sigma(v_1, v_K) \\ \vdots & \ddots & \vdots \\ \Sigma(v_K, v_1) & \dots & \Sigma(v_K, v_K) \end{bmatrix} \right) \tag{13}$$

Note that in Eq. (12), $\Phi(\cdot)$ is also GP, since $\Phi = f + \epsilon$ and adding two random variables that are separate but exactly distributed is also a Gaussian. Therefore, to build a GP model using a training dataset then, Eq. (12) becomes;

$$\Phi_y \sim N(0, \Sigma(v, v) + \sigma_{\tau < k}^2 I). \tag{14}$$

Equation (14) does not hold a definite parameter and is depicted as non-parametric.

Finally, to perform a prediction, we identify the joint distribution of the original target label Φ_y of the train points (v) and the latent target label $\hat{\Phi}_y$ of the test observation (\hat{v}). Under the GP, the joint distribution of Φ_y and $\hat{\Phi}_y$ is defined as

$$\begin{bmatrix} \Phi_y \\ \hat{\Phi}_y \end{bmatrix} \sim N \left(0, \begin{bmatrix} \Sigma(\mathbf{v}, \mathbf{v}) + \sigma_{\tau < k}^2 \mathbf{I} & \Sigma(\mathbf{v}, \hat{\mathbf{v}}) \\ \Sigma(\hat{\mathbf{v}}, \mathbf{v}) & \Sigma(\hat{\mathbf{v}}, \hat{\mathbf{v}}) + \hat{\sigma}_{k-\tau}^2 \mathbf{I} \end{bmatrix} \right). \tag{15}$$

By the conditional probability property of Gaussian (Roberts et al. 2013), we have

$$p(\hat{\phi}_y | \hat{v}, v, \Phi_y) = N(\hat{\mu}, \hat{\Sigma}), \tag{16}$$

where the predicted mean is

$$\hat{\mu} = \Sigma(\hat{v}, v)(\Sigma(v, v) + \hat{\sigma}^2 I)^{-1} \Phi_y, \tag{17}$$

and variance is

$$\hat{\Sigma} = \Sigma(\hat{v}, \hat{v}) + \hat{\sigma}^2 I - \Sigma(\hat{v}, v)(\Sigma(v, v) + \hat{\sigma}^2 I)^{-1} \Sigma(v, \hat{v}). \tag{18}$$

Therefore, we can estimate the posterior on the testing set given the expected and the covariance kernel functions of the GP. With this, we can generate a confidence level of 95% to obtain a Bayesian stochastic model on the objective function.

- (b) Compute the hyperparameter values that boost the acquisition function: The surrogate function described in Step (1) provides an assay of the objective function, making it useful for direct sampling. The sampling uses the posterior in Step (1) to choose the subsequent sample from the search space. This technique helps us to account for the areas in the domain space that are worth exploiting and the ones that are worth exploring. Analytically, we can utilise this acquisition function by considering a posterior function Ψ models the loss such that after k evaluations, the lower bound of this posterior becomes Ψ^* at some measured point, ω^* . Performing more evaluations will alter the posterior values and the objective function values at, let us say $\Psi(\omega)$. The result of this optimisation problem becomes $\min(\Psi^*, \Psi(\omega))$ and the expected difference distinguishing Ψ^* and the latter produces Eq. (19), (Doke et al. 2020);

$$\mathcal{E}_I = E_k[\Psi^* - \min(\Psi^*, \Psi(\omega))]. \tag{19}$$

- (c) Evaluate the optimal function and insert the outputs into the Gaussian process posterior: In this Step, we determine the objective function using the optimal parameters gathered from the acquisition function. This approach improves the performance of the objective we aim to model by taking the Gaussian process posterior from k measurements to $k + 1$.
- (d) Reproduce Steps 2 and 3 until the maximum number of iterations is attained.

This work used the BO method to perform a hyperparameter search by taking the set of hyperparameter values presented in Table 3 as inputs and producing the evaluation accuracy for the Bayesian optimiser. New models are constructed with specific hyperparameters during the search process and then trained for several epochs. Note that these new models are evaluated against a set of metrics. Table 4 lists the top 5 new models (described in this work as LSTM-1, LSTM-2,

Table 4 A summary report displaying the top 5 hyperparameters used to build the LSTM model via BO

Parameter Types	LSTM-1	LSTM-2	LSTM-3	LSTM-4	LSTM-5
LSTM	512	512	512	160	32
Activation	relu	tanh	relu	relu	tanh
Dropout	True	False	False	True	False
Learning rate	0.01	0.01	0.01	0.01	0.01

LSTM-3, LSTM-4, and LSTM-5) adopted to measure the spring discharge.

To ensure that a robust LSTM is produced for this work, the BO approach was used to fine-tune the hyperparameters of the neural network. This was achieved by creating a training pipeline for the LSTM with the simulated measurement set and then customising the hyperparameters into an objective function as depicted in Table 4. The hyperparameter optimisation was used to develop and train 5 new models, that is, LSTM-1 to LSTM-5. A flowchart illustrating the procedure for implementing the LSTM model developed in conjunction with Bayesian Optimisation hyperparameter tuning for spring discharge estimation is shown in Fig. 10.

Performance evaluation

For training the top 5 LSTMs, the mean squared error (MSE) of the training dataset was selected as a benchmark to calibrate the recurrent networks, and their corresponding optimal values of the MSEs were acquired using the Adam optimisation scheme. This optimiser is an improved function of the stochastic gradient descent, and it was designed to update network weights in the training pipeline.

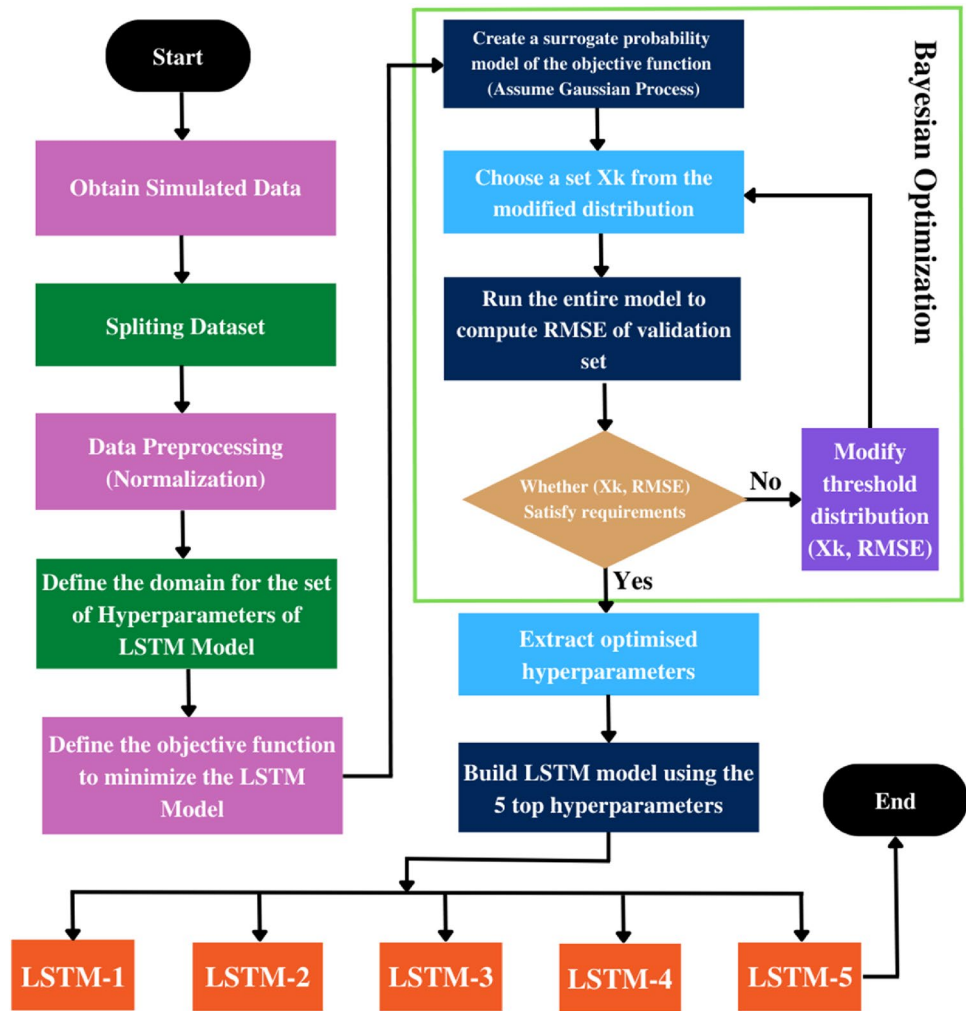
In addition to the above, the prediction performances of the five LSTMs were evaluated using the root mean square error (RMSE), the mean absolute error (MAE), the mean absolute percentage error (MAPE), Pearson product–moment correlation coefficients (ρ), and the coefficient of determination (R^2). The statistical metrics are defined as follows:

$$MSE = \frac{1}{N} \sum_{j=1}^N (d_j - d_j^*)^2 \tag{20}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (d_j - d_j^*)^2} \tag{21}$$

$$MAE = \frac{1}{N} \sum_{j=1}^N |d_j - d_j^*| \tag{22}$$

Fig. 10 Flowchart showing the implementation procedure of the LSTM model developed coupled with Bayesian Optimisation hyperparameter tuning for spring discharge estimation



$$MAPE = \frac{1}{N} \sum_{j=1}^N \left| \frac{d_j - d_j^*}{d_j} \right| \times 100\% \quad (23)$$

$$\rho = \frac{\sum_{j=1}^N (d_j - \bar{d})(d_j^* - \bar{d}^*)}{\sqrt{\sum_{j=1}^N (d_j - \bar{d})^2 (d_j^* - \bar{d}^*)^2}} \quad (24)$$

$$R^2 = \rho^2 \quad (25)$$

where N is the sample size of the measured data (that is, the real spring discharge data), (d_j, d_j^*) are the actual and predicted data, respectively, \bar{d} is the mean of the values of the d_j , and \bar{d}^* is the mean of the values of the d_j^* . The RMSE which is also referred to as the standard error, reports on the dispersion of the prediction errors. MAE denotes the average of absolute errors, and it is used to compute how far the predicted values deviate from the test data. MAPE measures

model accuracy as a percentage, and it works best if there are no extremes in the data. As such, MAPE is primarily used to determine any irregularity in the test data (Swanson et al. 2011). The computed values of RMSE, MAE, and MAPE are all within the range of $[0, \infty]$. Usually, a small value of RMSE indicates a small deviation between the predicted and the observed test data. Therefore, the smaller the value (that is, $RMSE < 1$), the better the model design. On the other hand, the correlation coefficient measures the strength of the relationship between the predicted values and the original test data. The metric measure ranges from $[-1, 1]$. The extreme values of -1 and 1 indicate a perfectly linear relationship between the predicted values and the original test data. Meanwhile, as the coefficient approaches -1 or 1 , the strength of the relationship increases and the data points tend to draw closer to a reference line. Lastly, the coefficient of determination measures the extent to which the predicted values explain the original test data.

Results and discussions

Generally, three primary methods are used in groundwater research to analyse hydrological behaviours; mathematical models, numerical simulations, and the use of laboratory-scaled experiments (Ding et al. 2020; Suresh Kumar 2014). Mathematical models and numerical simulations have gained much attention recently (Chang et al. 2015). Initially, laboratory-scaled experiments had some limitations when complex geological conditions needed to be studied due to the limited use of materials (Ding et al. 2020). However, with the advancement of technology, the use of laboratory-scaled experiments in conjunction with other methods, such as machine learning approaches, has made tremendous progress in simulating complex geological processes such as karst formations.

The hydrological cycle process in karst environments is so complex that a single linear input–output model cannot properly explain it (Song et al. 2022). This complexity was also evident in research conducted by Labat et al. (2000), who created a linear stochastic rainfall-runoff model and performed a Fourier analysis on three karst springs in France. They concluded that linear input–output models could not accurately represent the hydraulic behaviours of karst springs. Additionally, research by Tao et al. (2022) revealed the importance of humidity in the hydrological cycle processes and further use in water resources planning and management.

Therefore, this work built a 3D physical model based on the conceptual model of a karst aquifer and predicted spring discharge by combining three factors, precipitation intensity, temperature, and humidity. A novel hybrid LSTM-RNN model coupled with BO was employed to model the multivariate time series behaviour of the input parameters. The study's findings demonstrated the capability of a deep learning model for simulating spring discharge with various input parameters in a karst environment, which validates the findings of previous studies (Hu et al. 2018; Kratzert et al. 2018; Song et al. 2022) that demonstrated the ability of neural networks in solving non-linear relationships in karst terrains.

Influence of rainfall intensity on spring discharge

Examining the influence of rainfall intensity on spring discharge is critical for assessing its impacts and the non-linearity trends in karst environments. The spring hydrograph was used to investigate the effect of varied precipitation intensities on spring flow. The findings revealed that the lower the intensity of the rainfall, the lower the peak flow of the spring hydrograph, and vice versa. Figures 11a–c

and 12a–c demonstrate this. However, an intriguing result was found in 'd' of both Figs. 11 and 12, which deviated from the previous trend. At the high precipitation intensity of 0.02925 mm/s for 30 min duration of simulated rainfall, the peak flow for spring discharge was very low compared to the other precipitation intensities which had the highest peak flow for the longest duration rainfall of 30 min. The deviation shown in 'd' of both Figs. 11 and 12, could be explained by the heterogeneous nature of the soil (sand and loam), which comprised the porous layer.

The porous layer is a layer of soil that has interconnected pore spaces through which water can flow. Porosity refers to the total amount of pore space available in the soil or rock, while permeability refers to how easily water can flow through that pore space. When the rainfall intensity is high, such as during a heavy rainstorm, the amount of water that falls onto the porous layer can exceed the capacity of the soil to absorb it. If the porosity and permeability of the porous layer are low, the excess rainwater will not be able to infiltrate into the system and will instead accumulate on the surface as overland flow. The infiltration rate is affected by the nature of the soil, i.e., its physico-chemical properties (Djukem Fenguia and Nkouathio 2023).

The degree of permeability and porosity of the porous layer showed that not all the simulated Rainfall could infiltrate into the system and further percolate into the aquifer, owing to the high rainfall intensity. The porous layer became fully saturated quickly, and overland flow (i.e., surface runoff) was higher. This led to a shorter peak of the spring hydrograph, as shown in Figs. 11d and 12d. This deviation is also analogous to findings by Djukem Fenguia and Nkouathio (2023), whose results identified the nature of the soil as one major cause of flooding.

With respect to this study, the porous layer quickly reaches its maximum capacity to hold water and becomes saturated at a precipitation intensity of 0.02925 mm/s for 30 min duration of simulated rainfall. This means that any additional rainwater will simply run off the surface of the soil and contribute to surface runoff. This can lead to increased soil erosion, flooding, and decreased groundwater recharge. Therefore, it is important to consider the permeability and porosity of the porous layer when designing and managing water systems for effective water management and sustainability. It is also imperative to pay attention to other physico-chemical properties of the soil, such as grain size, moisture, pH, compactness, and organic matter in water resources planning and management (Ansar and Naima 2021; Djukem Fenguia and Nkouathio 2023).

Predictive models

Furthermore, to determine the variance of the proposed model, the regression models, i.e., LSTM-1 to LSTM-5,

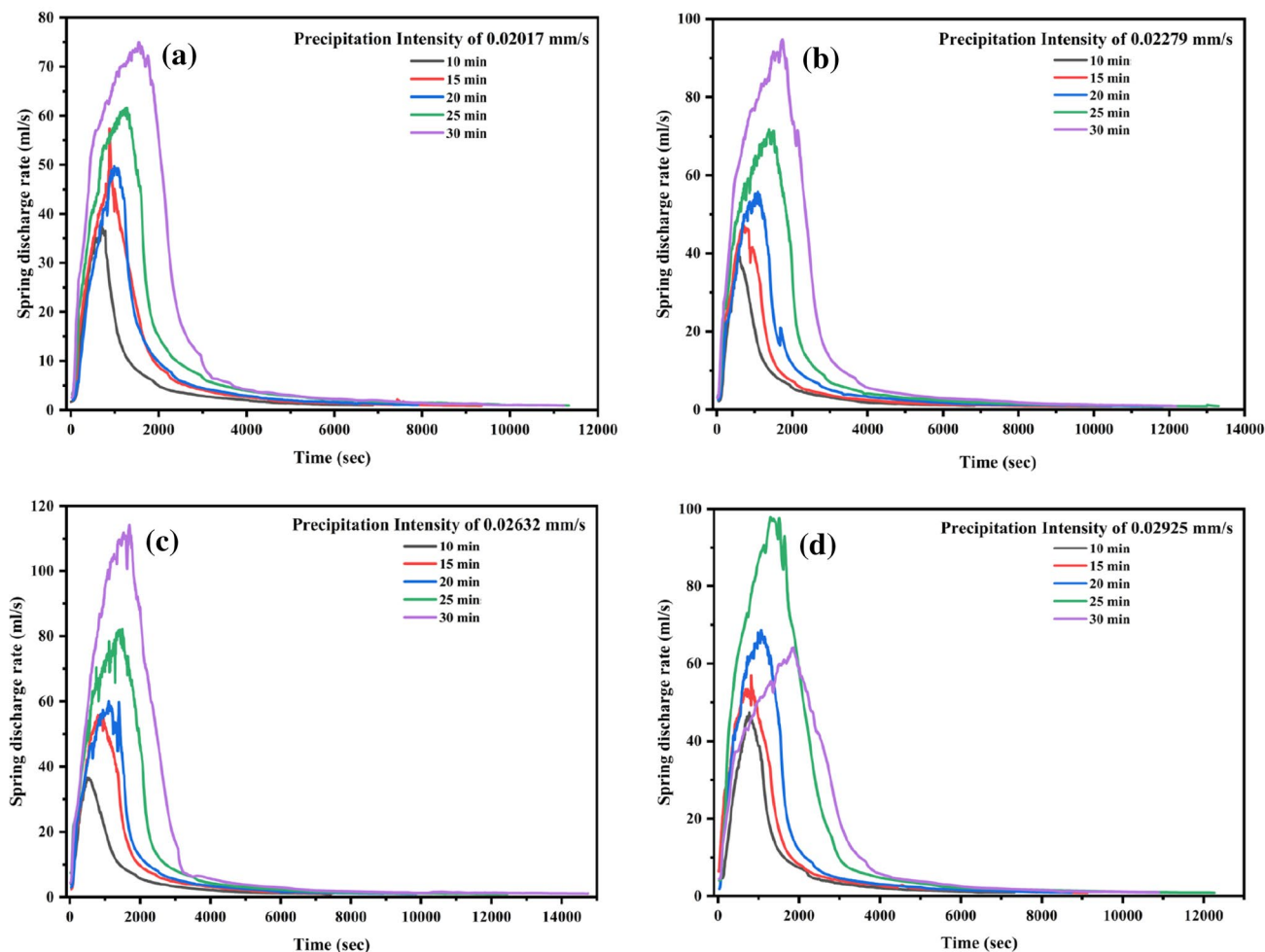


Fig. 11 Spring hydrographs showing the peak flows for different precipitation intensities and different duration of 15, 20, 25, and 30 min. The precipitation intensities are 0.02017, 0.02279, 0.02632, and 0.02925 mm/s

were diagnosed and the results are presented in Fig. 13a–e. Figure 13a–c shows the prediction plot generated from LSTM-1, whilst Fig. 13d, and Fig. 13e were produced from LSTM-5. A detailed description of the performance of these 5 models using 3, 2, or 1 input parameter(s) is presented in Table 5. The plots in Fig. 13a–c were predicted using three input parameters (precipitation intensity, temperature, and humidity), two input parameters (precipitation intensity and temperature), and only one input parameter (precipitation intensity), respectively.

The left side of the prediction plots presents the original targets from the test data versus the predicted values generated from the trained LSTM models. It can be observed from the residual plots on the right-hand side that plots with 3 and 2 input parameters produced symmetric distributions with mean errors of 0.003 ± 0.036 and 0.001 ± 0.037 , respectively.

The results indicate an acceptable prediction performance of LSTM-1, LSTM-2, LSTM-3, and LSTM-4. It

also validated the established models' ability to construct a robust and dependable learning process. However, based on the results, it was discovered that including more climate parameters in the prediction, matrix improved prediction performance. This can be seen from the results tabulated in Table 5 which is also confirmed by studies conducted by Tao et al. (2022) on the integration of extreme gradient boosting features with machine learning models on relative humidity prediction. His results showed that an increase in climatic parameters as input parameters improved the model performance. Furthermore, studies conducted by Fiorillo et al. (2021) revealed that a decrease in spring discharge does not appear to be solely dependent on precipitation changes but that temperature fluctuations play a vital role. This shows that in studying the hydrodynamic characteristics of the karst aquifers, particularly pertaining to spring discharge prediction, it is imperative to consider at least precipitation intensity and temperature in the analysis, as evident in Fig. 8 on the feature importance using the XGB regressor. However,

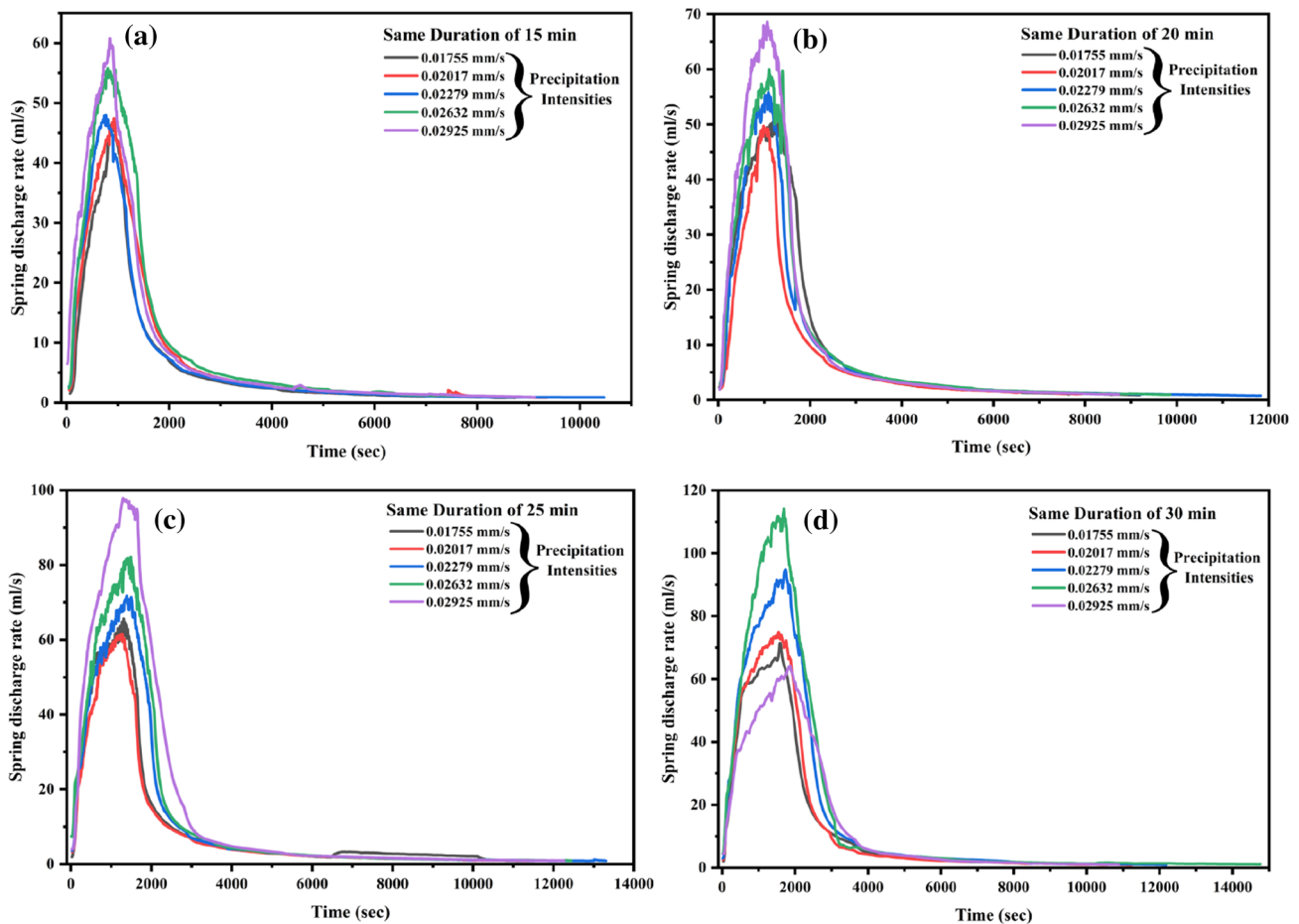


Fig. 12 Spring hydrographs showing the peak flows for different precipitation intensities with the same duration. The precipitation intensities are 0.02017, 0.02279, 0.02632, and 0.02925 mm/s corresponding to times of 15, 20, 25, and 30 min

using only precipitation intensity as an input variable in the LSTM model (refer to Fig. 13c), is not sufficient to predict spring discharge.

To investigate areas within the target that may be prone to more or less error, the residual plots displayed in Fig. 14a–c capture the disparities between errors on the y-axis and the dependent variable on the x-axis.

In Fig. 14a, b, a fairly random, uniform distribution of the residuals against the target in two dimensions—indicating that our non-linear LSTM-1 model is very robust in predicting the spring discharge using at least 2 input variables is seen. The R^2 values for adopting 3 input variables are 97.2% (for the training set) and 97.1% (for the testing set). In addition, the histograms show that the errors obtained are symmetrically distributed around zero, which also generally indicates a well-fitted model when we consider at least 2 input variables. However, the scenario is not the same when a single input variable is used as shown in Fig. 14c. The R^2 values, in this case, are 4.7% (for the training set) and 2.3% (for the testing set).

Additionally, a two-dimensional graphical presentation i.e., “Taylor Diagram (Ansar and Naima 2021)” was generated for the proposed LSTM models 1–5 for the estimation of spring discharge in karst terrains and this is shown in Fig. 15.

Taylor’s Diagram is normally used to justify the degree of agreement between the model developed and the measurement set used (the commonly used precision indexes are the Pearson correlation coefficient, the root-mean-square error (RMSE), and the standard deviation) to aid in evaluating the accuracy of the model. In this study (refer to Fig. 15), the same statistical metrics were adopted to assess the performance of the 5 LSTM models produced from the Bayesian optimisation tuning approach.

The five colours in the plot represent the 5 LSTM models developed. The black star on the horizontal axis is the observed or reference point. When the simulation point is close to the observed point, it means they are similar in terms of standard deviation, with a high correlation and a centred root mean square error (CRMSE) value close to

Fig. 13 Prediction error plots depicting 20% of the actual discharge from the simulated dataset against the predicted values generated by the proposed model: **a** the best-optimised model using precipitation intensity, temperature and humidity as input parameters; **b** the best-optimised LSTM model using precipitation intensity and temperature as input parameters; **c** the best-optimised LSTM model using only precipitation intensity as an input parameter; **d** the 5th-ranked-optimised LSTM model using precipitation intensity, temperature and humidity as input parameters; **e** the 5th-ranked-optimised LSTM model using precipitation intensity and temperature as input parameters

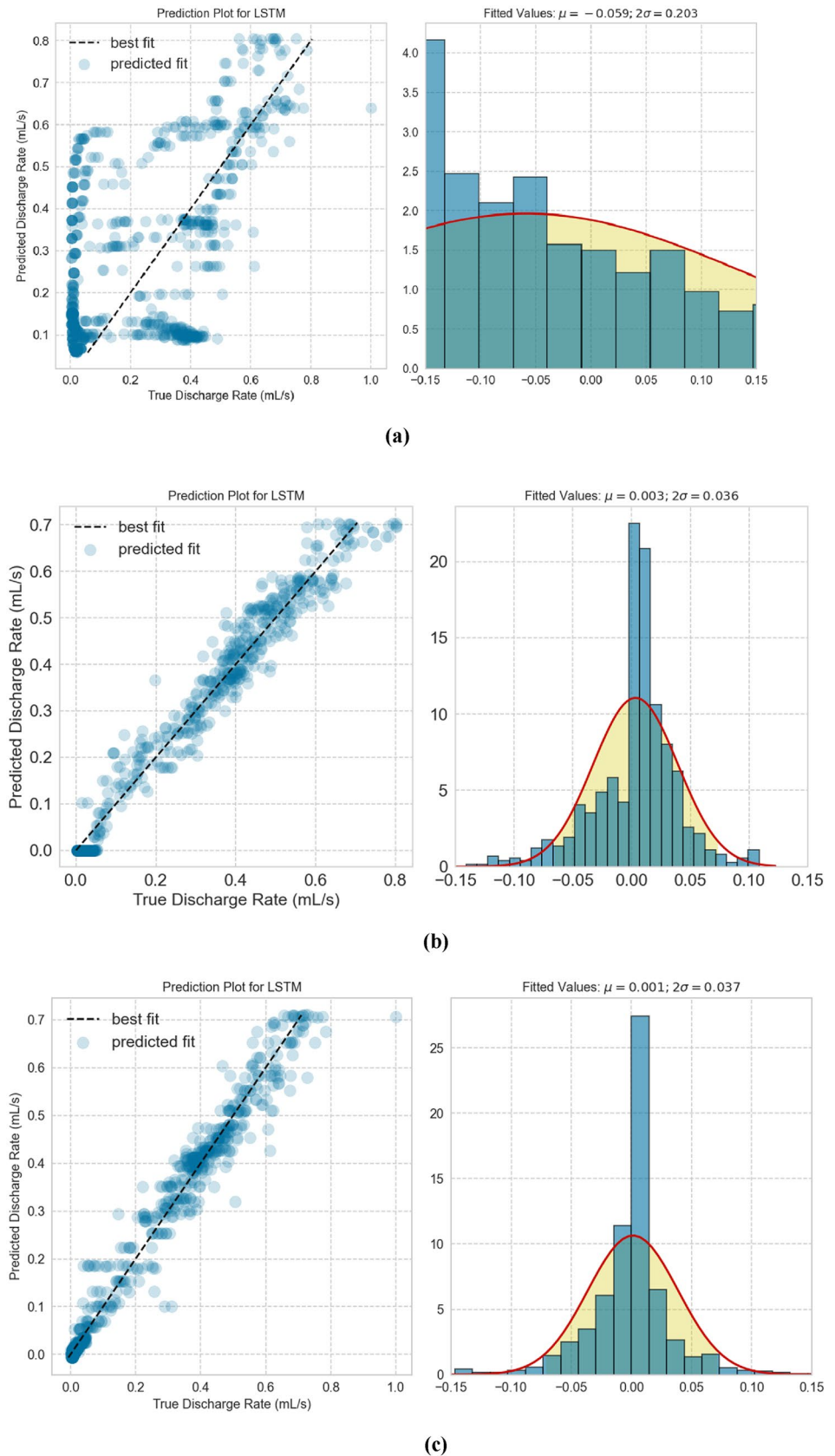
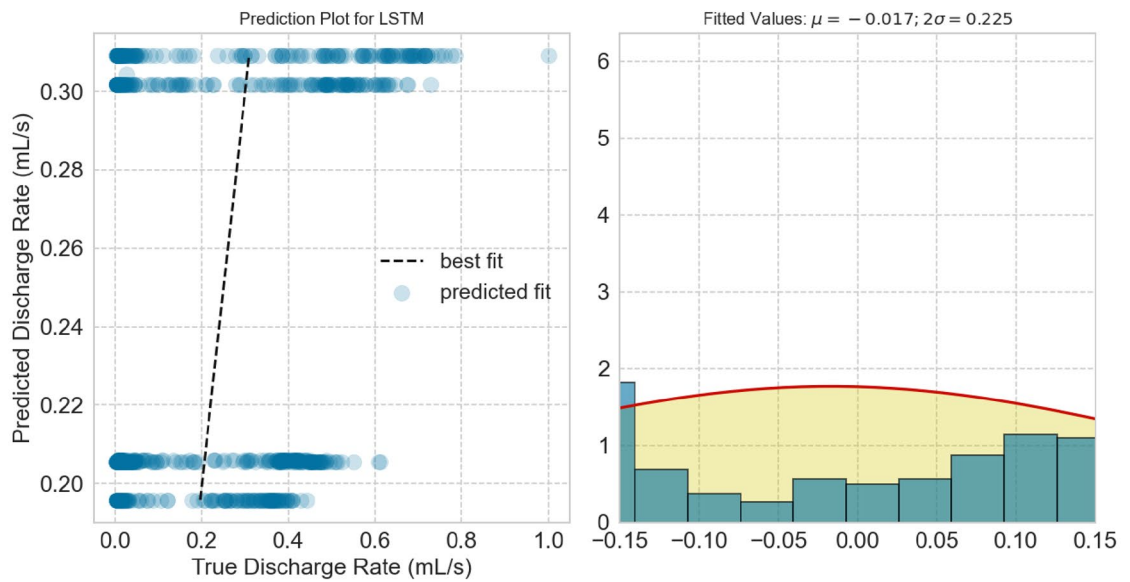
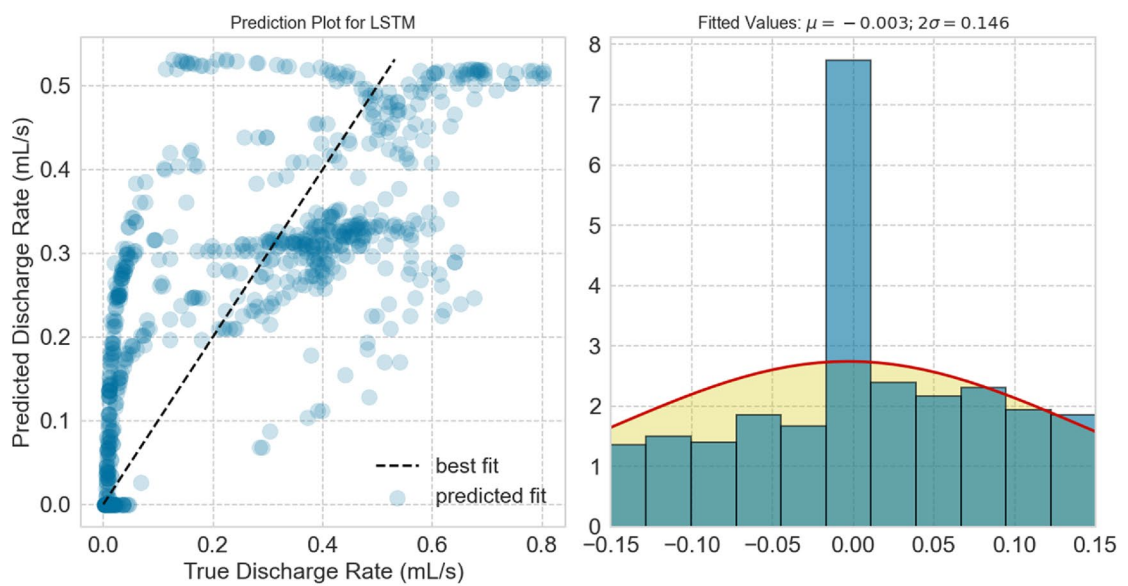


Fig. 13 (continued)



(d)



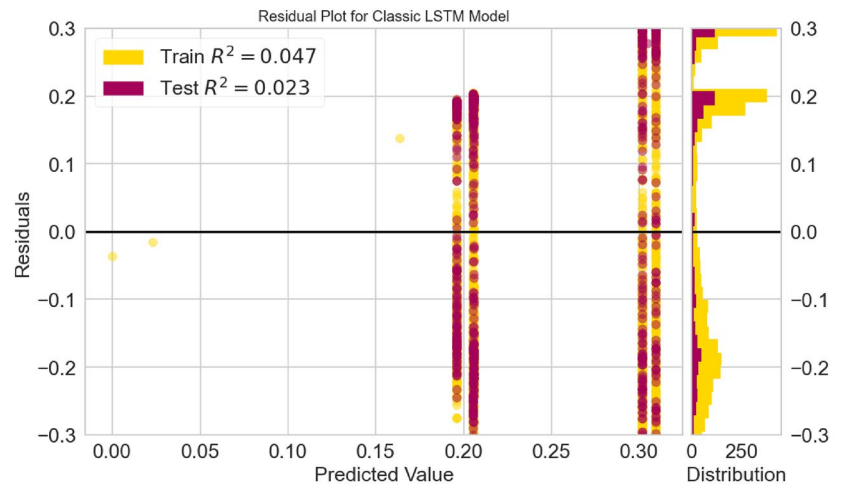
(e)

zero. The black dashed lines represent the standard deviation of the observed time series. Values above the black dashed line mean the simulated data set has a higher variation, and values below mean the data set has a lower variation. The contour lines on the polar plot show the values of CRMSE.

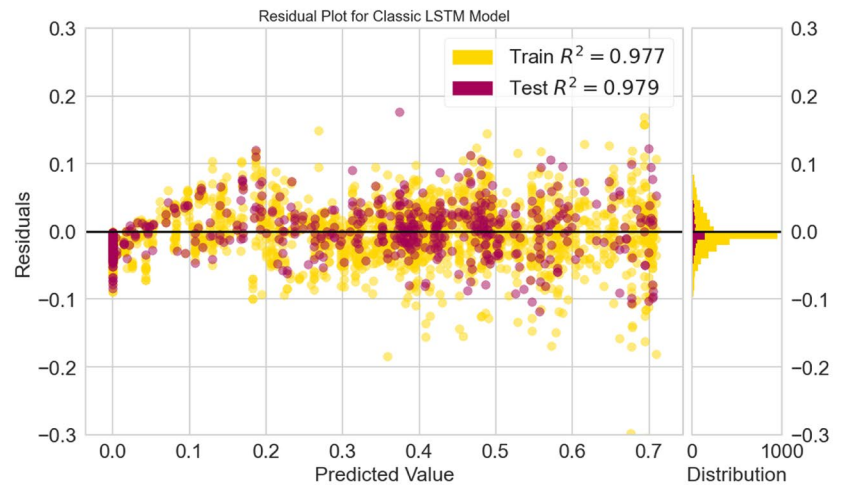
Figure 15 (left panel), models 1, 2, 3, and 4 are similar in terms of their standard deviation, which is close to the reference or observed point and also falls on or close to the black dashed line with a correlation above 99% and a

CRMSE of about 0.04. Model 5, on the other hand, has a lower standard deviation (below the black dashed line) with a correlation of around 91%, and the CRMSE is close to 0.16. On the other hand, Fig. 15 (right panel) shows models 1, 2, and 3 to be slightly below the black dashed line even though their respective standard deviations were closer to the reference or observed point. All three models had a correlation above 99% and a CRMSE of about 0.04. Model 4 had a lower deviation than the referenced black

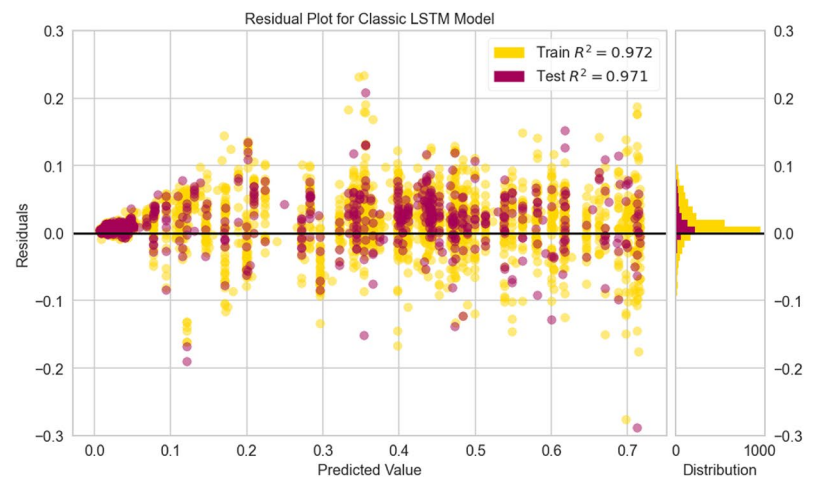
Fig. 14 Residual plots showing irregularities between the simulated and fitted response values



(a) The effect of using three (3) input parameters in the LSTM



(b) The effect of using two (2) input parameters in the LSTM



(c) The effect of using one (1) input parameter in the LSTM

Table 5 Comparison of the true mean value (μ) and the 95% confidence interval (2σ) of the five (5) different LSTM models using three input parameters (precipitation intensity, temperature, and humidity), two input parameters (temperature and humidity), and one input parameter (precipitation intensity)

Model type	Estimated errors: $\mu \pm 2\sigma$		
	3 inputs	2 inputs	1 input
LSTM-1	0.003 ± 0.0362	0.001 ± 0.0370	0.017 ± 0.225
LSTM-2	0.008 ± 0.036	0.001 ± 0.0037	–
LSTM-3	0.001 ± 0.036	0.007 ± 0.039	–
LSTM-4	0.010 ± 0.038	0.003 ± 0.071	–
LSTM-5	-0.003 ± 0.146	-0.059 ± 0.203	–

dashed line (standard deviation of the observed data set). It also had a correlation of about 98%, and the CRMSE was very close to the lowest deviation among all with a correlation of about 78% and a CRMSE of about 0.22. From the performance metrics of ρ and R^2 , LSTM-1 with three input parameters performed better in estimating spring discharge than LSTM-2 and LSTM-3, even though the deviation was not much, in estimating spring discharge. LSTM-4 and LSTM-5 performed relatively poorly in comparison.

Conclusions

Predicting spring flow is very important for management of water demand especially in cities. In this study, the LSTM-RNN model of deep learning coupled with the BO hyperparameter tuning is identified, trained and tested to predict spring flow in karst environment using climatic variables, based on data gathered from a physical laboratory experiment. Different combinations of model input parameters are considered in this study to predict spring flow.

The experimentation revealed in the results section that model performance is primarily determined by input variables and combinations of parameter selection. Model accuracy improves significantly when the number of neurons and epochs are optimised; the Bayesian Optimizer was employed in this study. The Bayesian Optimiser significantly improved the efficiency of the top 5 hybrid versions and gave a stronger prediction ability of the model. The parameter optimisation using the BO was used to determine the top 5 high-precision LSTM models (LSTM-1, LSTM-2, LSTM-3, LSTM-4, and LSTM-5). Also, the use of epochs and hidden layer neurons improves computation speed as well.

This study has proven that the LSTM-RNN model can be used to predict spring flow and can also predict other variables such as rainfall and temperature. It also proved that when the input variables of the model are more, the output of prediction is better. The advantage of this LSTM model is that it may require a smaller training data set to

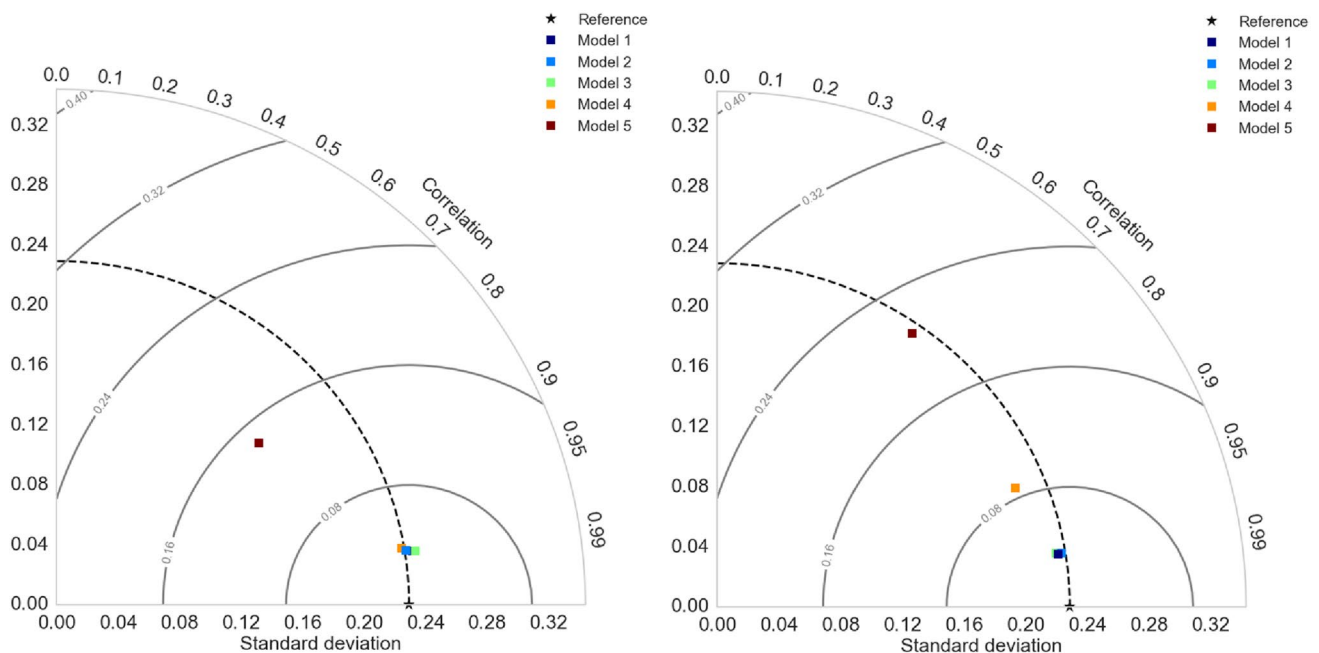


Fig. 15 Taylor’s Diagram of the prediction LSTM model developed using three input variables (left panel) and two input variables (right panel), respectively

train the model to predict input variables. It is also beneficial for water resources managers to understand the relationship between spring discharge and climatic variables.

Overall, the development of the physical model, the experimental results, and the deep learning models will serve as valuable reference for future research and investigations in the Jinan Spring Basin and also as useful guidelines for other areas.

Acknowledgements The author thanks the staff of State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, 210098, China and research team mates for the assistance rendered.

Author contributions Conceptualisation, project administration, data curation, methodology, investigation, software; formal analysis, validation, writing-original draft, visualisation: PAO; conceptualisation, resources, funding acquisition, project administration, supervision, writing-review and editing: LS; methodology, software, validation, writing-original draft: TA-N; data curation, writing-review and editing: PB; data curation KBMOY; writing-review and editing: AKK; data curation: SN.

Funding This research work was funded by the Key Technologies and Application Demonstration of Groundwater Over-Extraction Control and Protection in Huang-Huai-Hai Region (No. 2021YFC3200500).

Data availability The datasets and codes generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

References

- Akano TT, James CC (2022) An assessment of ensemble learning approaches and single-based machine learning algorithms for the characterization of undersaturated oil viscosity. *Beni-Suef Univ J Basic Appl Sci* 11:149. <https://doi.org/10.1186/s43088-022-00327-8>
- Alameer Z, Fathalla A, Li K, Ye H, Jianhua Z (2020) Multistep-ahead forecasting of coal prices using a hybrid deep learning model. *Resour Policy* 65:101588. <https://doi.org/10.1016/j.resourpol.2020.101588>
- Alizadeh B, Ghaderi Bafti A, Kamangir H, Zhang Y, Wright DB, Franz KJ (2021) A novel attention-based LSTM cell post-processor coupled with Bayesian optimisation for streamflow prediction. *J Hydrol* 601:126526. <https://doi.org/10.1016/j.jhydrol.2021.126526>
- An L, Ren X, Hao Y, Yeh TCJ (2019) Utilising precipitation and spring discharge data to identify groundwater quick flow belts in a karst spring catchment. *J Hydrometeorol* 20(10):2057–2068. <https://doi.org/10.1175/JHM-D-18-0261.1>
- An L, Hao Y, Yeh TCJ, Liu Liu Y, W, Zhang, B, (2020) Simulation of karst spring discharge using a combination of time–frequency analysis methods and long short-term memory neural networks. *J Hydrol* 589:125320. <https://doi.org/10.1016/j.jhydrol.2020.125320>
- Anand V, Oinam B (2022) Modeling the potential impact of land use/land cover change on the hydrology of Himalayan River Basin. *Handbook of Himalayan ecosystems and sustainability*, vol 2. CRC Press, Boca Raton, pp 189–204
- Andreo B (2012) Introductory editorial: advances in karst hydrogeology. *Environ Earth Sci* 65(8):2219–2220. <https://doi.org/10.1007/s12665-012-1621-3>
- Ansar A, Naima A (2021) Mapping of flood zones in urban areas through a hydro-climatic approach: the case of the city of Abha. *Earth Sci Res* 10(2):1. <https://doi.org/10.5539/esr.v10n2p1>
- Bailer-Jones CAL, MacKay DJC, Withers PJ (1998) A recurrent neural network for modelling dynamical systems. *Network Comput Syst* 9(4):531–547. https://doi.org/10.1088/0954-898X_9_4_008
- Bao J (2020) Multi-features-based arrhythmia diagnosis algorithm using xgboost. In: *Proceedings—2020 international conference on computing and data science CDS 2020*, pp 454–457. <https://doi.org/10.1109/CDS49703.2020.00095>
- Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimisation. *J Mach Learn Res* 13:281–305
- Bergstra J, Bardenet R, Bengio Y, Kégl B (2011) Algorithms for hyper-parameter optimisation. In: *Advances in neural information processing systems 24: 25th annual conference on neural information processing systems 2011 NIPS 2011*, pp 1–9. <https://doi.org/10.5555/2986459.2986743>
- Bergstra J, Yamins D, Cox DD (2013) Making a science of model search: hyperparameter optimisation in hundreds of dimensions for vision architectures. In: *30th International conference on machine learning ICML 2013 part 1*, pp 115–123
- Biondić B, Biondić R, Dukarić F (1998) Protection of karst aquifers in the Dinarides in Croatia. *Environ Geol* 34(4):309–319. <https://doi.org/10.1007/s002540050283>
- Boyle P (2007) Gaussian processes for regression and optimisation [open access Victoria University of Wellington Te Herenga Waka]. <https://doi.org/10.26686/wgtn.16934869.v1>
- Chang Y, Wu J, Jiang G (2015) Modeling the hydrological behavior of a karst spring using a non-linear reservoir-pipe model. *Hydrogeol J* 23(5):901–914. <https://doi.org/10.1007/s10040-015-1241-6>
- Chen T, Guestrin C (2016) XGBoost. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp 785–794. <https://doi.org/10.1145/2939672.2939785>
- Cheng H, Ding X, Zhou W, Ding R (2019) A hybrid electricity price forecasting model with Bayesian optimisation for German energy exchange. *Int J Electr Power Energy Syst* 110(February):653–666. <https://doi.org/10.1016/j.ijepes.2019.03.056>
- Dempster AP (1968) A generalization of Bayesian inference. *J Roy Stat Soc Ser B Methodol* 30(2):205–232. <https://doi.org/10.1111/j.2517-6161.1968.tb00722.x>
- Ding H, Zhang X, Chu X, Wu Q (2020) Simulation of groundwater dynamic response to hydrological factors in karst aquifer system. *J Hydrol* 587:124995. <https://doi.org/10.1016/j.jhydrol.2020.124995>
- DjukemFengua SN, Nkouathio DG (2023) Contribution of soil physical properties in the assessment of flood risks in tropical areas: case of the Mbo plain (Cameroon). *Nat Hazards*. <https://doi.org/10.1007/s11069-023-05818-0>
- Doke P, Shrivastava D, Pan C, Zhou Q, Zhang YD (2020) Using CNN with Bayesian optimisation to identify cerebral microbleeds. *Mach vis Appl* 31(5):1–14. <https://doi.org/10.1007/s00138-020-01087-0>
- Fiorillo F, Leone G, Pagnozzi M, Esposito L (2021) Long-term trends in karst spring discharge and relation to climate factors and

- changes. *Hydrogeol J* 29(1):347–377. <https://doi.org/10.1007/s10040-020-02265-0>
- Ford D, Williams P (2007) *Karst hydrogeology and geomorphology*. John Wiley & Sons Ltd, New Jersey
- Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F (2012) A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Trans Syst Man Cybern C (Appl Rev)* 42(4):463–484. <https://doi.org/10.1109/TSMCC.2011.2161285>
- Ghawi R, Pfeffer J (2019) Efficient hyperparameter tuning with grid search for text categorization using kNN approach with BM25 similarity. *Open Comp Sci* 9(1):160–180. <https://doi.org/10.1515/comp-2019-0011>
- Greenhill S, Rana S, Gupta S, Vellanki P, Venkatesh S (2020) Bayesian optimisation for adaptive experimental design: a review. *IEEE Access* 8:13937–13948. <https://doi.org/10.1109/ACCESS.2020.2966228>
- Haggerty R, Sun J, Yu H, Li Y (2023) Application of machine learning in groundwater quality modelling—a comprehensive review. *Water Res* 233:119745. <https://doi.org/10.1016/j.watres.2023.119745>
- Hao Y, Zhang J, Wang J, Li R, Hao P, Zhan H (2016) How does the anthropogenic activity affect the spring discharge? *J Hydrol* 540:1053–1065. <https://doi.org/10.1016/j.jhydrol.2016.07.024>
- Hartmann A, Liu Y, Olarinoye T, Berthelin R, Marx V (2021) Integrating field work and large-scale modeling to improve assessment of karst water resources. *Hydrogeol J* 29(1):315–329. <https://doi.org/10.1007/s10040-020-02258-z>
- Hassanzadeh Y, Ghazvinian M, Abdi A, Baharvand S, Jozaghi A (2020) Prediction of short and long-term droughts using artificial neural networks and hydro-meteorological variables. <http://arxiv.org/abs/2006.02581>
- He F, Zhou J, Feng Z, Kai Liu G, Yang Y (2019) A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimisation algorithm. *Appl Energy* 237:103–116. <https://doi.org/10.1016/j.apenergy.2019.01.055>
- He X, Luo J, Li P, Zuo G, Xie J (2020) A hybrid model based on variational mode decomposition and gradient boosting regression tree for monthly runoff forecasting. *Water Resour Manage* 34(2):865–884. <https://doi.org/10.1007/s11269-020-02483-x>
- Hochreiter S (1998) The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Int J Uncertain Fuzziness Knowl Based Syst* 06(02):107–116. <https://doi.org/10.1142/S0218488598000094>
- Hochreiter S, Schmid Huber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hu C, Wu Q, Li H, Jian S, Li N, Lou Z (2018) Deep learning with a long short-term memory networks approach for rainfall-runoff simulation. *Water* 10(11):1543. <https://doi.org/10.3390/w10111543>
- Ikard S, Pease E (2019) Preferential groundwater seepage in karst terrane inferred from geoelectric measurements. In: *Proceedings of the symposium on the application of geophysics to engineering and environmental problems SAGEEP 2019 March*, p 57. <https://doi.org/10.1002/nsg.12023>
- Jeannin PY, Artigue G, Butscher C, Chang Y, Charlier JB, Duran L, Gill L, Hartmann A, Johannet A, Jourde H, Kavousi A, Liesch T, Liu Y, Lüthi M, Malard A, Mazzilli N, Pardo-Igúzquiza E, Thiéry D, Reimann T, Wunsch A (2021) Karst modelling challenge 1: results of hydrological modelling. *J Hydrol* 600:126508. <https://doi.org/10.1016/j.jhydrol.2021.126508>
- Kareem DA, Amen ARM, Mustafa A, Yüce MI, Szydłowski M (2022) Comparative analysis of developed rainfall intensity–duration–frequency curves for Erbil with other Iraqi urban areas. *Water (Switzerland)* 14(3):419. <https://doi.org/10.3390/w14030419>
- Khorram S, Jehbez N (2023) A hybrid CNN-LSTM approach for monthly reservoir inflow forecasting. *Water Resour Manage*. <https://doi.org/10.1007/s11269-023-03541-w>
- Khorrami M, Alizadeh B, Tousi EG, Shakerian M, Maghsoudi Y, Rahgozar P (2019) How groundwater level fluctuations and geotechnical properties lead to asymmetric subsidence: a PSInSAR analysis of land deformation over a transit corridor in the Los Angeles metropolitan area. *Remote Sens* 11(4):377. <https://doi.org/10.3390/rs11040377>
- Kratzert F, Klotz D, Brenner C, Schulz K, Herrnegger M (2018) Rainfall-runoff modelling using long-short-term-memory (LSTM) networks. *Hydrol Earth Syst Sci* 22:6005–6022. <https://doi.org/10.5194/hess-22-6005-2018>
- Labat D, Ababou R, Mangin A (2000) Rainfall–runoff relations for karstic springs. Part I: convolution and spectral analyses. *J Hydrol* 238(3–4):123–148. [https://doi.org/10.1016/S0022-1694\(00\)00321-8](https://doi.org/10.1016/S0022-1694(00)00321-8)
- Li J, Zhang R (2018) Dynamic weighting multi factor stock selection strategy based on xgboost machine learning algorithm. In: *Proceedings of 2018 IEEE international conference of safety produce informatization IICSPI 2018*, pp 868–872. <https://doi.org/10.1109/IICSPI.2018.8690416>
- Li L, Jamieson K, DeSalvo G, Rostamizadeh A, Talwalkar A (2018) Hyperband: a novel bandit-based approach to hyperparameter optimisation. *J Mach Learn Res* 18:1–52
- Lin Q, Zheng J, Zou C, Cheng J, Li J, Xia W, Shi H (2020) An improved 3-pentanone high temperature kinetic model using Bayesian optimisation algorithm based on ignition delay times, flame speeds and species profiles. *Fuel* 279:118540. <https://doi.org/10.1016/j.fuel.2020.118540>
- Ma J, Ding Y, Cheng JCP, Jiang F, Gan VJL, Xu Z (2020) A Lag-FLSTM deep learning network based on Bayesian optimization for multi-sequential-variant PM25 prediction. *Sustain Cities Soc* 60:102237. <https://doi.org/10.1016/j.scs.2020.102237>
- Mantovani RG, Rossi ALD, Vanschoren J, Bischl B, De Carvalho ACPLF (2015) Effectiveness of random search in SVM hyperparameter tuning. In: *Proceedings of the international joint conference on neural networks 2015 September*. <https://doi.org/10.1109/IJCNN.2015.7280664>
- Matsubara T, Knoblauch J, Briol F-X, Oates CJ (2021) *Robust generalised Bayesian inference for intractable likelihoods*. Oxford University Press, Oxford
- Meng E, Huang S, Huang Q, Fang W, Wu L, Wang L (2019) A robust method for non-stationary streamflow prediction based on improved EMD-SVM model. *J Hydrol* 568:462–478. <https://doi.org/10.1016/j.jhydrol.2018.11.015>
- Milly PCD, Betancourt J, Falkenmark M, Hirsch RM, Kundzewicz ZW, Lettenmaier DP, Stouffer RJ (2008) Climate change: stationarity is dead: whither water management? *Science* 319(5863):573–574. <https://doi.org/10.1126/science.1151915>
- Mohammed UD, Legesse SA, Berlie AB, Ehsan MA (2022) Climate change repercussions on meteorological drought frequency and intensity in South Wollo, Ethiopia. *Earth Syst Environ*. <https://doi.org/10.1007/s41748-022-00293-2>
- Niu W, Feng Z (2021) Evaluating the performances of several artificial intelligence methods in forecasting daily streamflow time series for sustainable water resources management. *Sustain Cities Soc* 64:102562. <https://doi.org/10.1016/j.scs.2020.102562>
- Parise M, Qiriazzi P, Sala S (2004) Natural and anthropogenic hazards in karst areas of Albania. *Nat Hazards Earth Syst Sci* 4(4):569–581. <https://doi.org/10.5194/nhess-4-569-2004>
- Quan Q, Hao Z, Xifeng H, Jingchun L (2022) Research on water temperature prediction based on improved support vector regression.

- Neural Comput Appl 34(11):8501–8510. <https://doi.org/10.1007/s00521-020-04836-4>
- Quijano AJ, Nguyen S, Ordonez J (2021) Grid search hyperparameter benchmarking of BERT, ALBERT, and LongFormer on DuoRC. <https://doi.org/10.48550/arXiv.2101.06326>
- Rasmussen CE (2004) Gaussian processes in machine learning. In: Bousquet O, von Luxburg U, Rätsch G (eds) Advanced lectures on machine learning. ML 2003. Lecture notes in computer science, vol 3176. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-28650-9_4
- Rasmussen CE, Williams CKI (2005) Gaussian processes for machine learning. In: Dietterich T (ed) Adaptive computation and machine learning. Massachusetts Institute of Technology, The MIT Press. ISBN 026218253X. www.GaussianProcess.org/gpml. <https://doi.org/10.7551/mitpress/3206.001.0001>
- Roberts S, Osborne M, Ebdon M, Reece S, Gibson N, Aigrain S (2013) Gaussian processes for time-series modelling. Philos Trans Roy Soc A Math Phys Eng Sci 371(1984):20110550. <https://doi.org/10.1098/rsta.2011.0550>
- Roshani SH, Saha TK, Rahaman MH, Masroor M, Sharma Y, Pal S (2022) Analyzing trend and forecast of rainfall and temperature in Valmiki Tiger Reserve, India, using non-parametric test and random forest machine learning algorithm. Acta Geophys 71:531–552. <https://doi.org/10.1007/s11600-022-00978-2>
- Sahar A, Han D (2018) An LSTM-based indoor positioning method using Wi-Fi signals. ACM Int Conf Proc Ser. <https://doi.org/10.1145/3271553.3271566>
- Shahriari B, Swersky K, Wang Z, Adams RP, De Freitas N (2016) Taking the human out of the loop: a review of Bayesian optimisation. Proc IEEE 104(1):148–175. <https://doi.org/10.1109/JPROC.2015.2494218>
- Sharafati A, Asadollah SB, Neshat A (2020) A new artificial intelligence strategy for predicting the groundwater level over the Rafsanjan aquifer in Iran. J Hydrol 591:125468. <https://doi.org/10.1016/j.jhydrol.2020.125468>
- Shekar BH, Dagne G (2019) Grid search-based hyperparameter tuning and classification of microarray cancer data. In: 2019 2nd International conference on advanced computational and communication paradigms ICACCP 2019. <https://doi.org/10.1109/ICACCP.2019.8882943>
- Singh K, Singh B, Sihag P, Kumar V, Sharma KV (2023) Development and application of modeling techniques to estimate the unsaturated hydraulic conductivity. Model Earth Syst Environ. <https://doi.org/10.1007/s40808-023-01744-z>
- Snelson EL (2007) Flexible and efficient Gaussian process models for machine learning. ACM SIGKDD Explor Newsl 7(2001):1–135
- Song X, Hao H, Liu W, Wang Q, An L, Jim Yeh T-C, Hao Y (2022) Spatial-temporal behavior of precipitation driven karst spring discharge in a mountain terrain. J Hydrol 612:128116. <https://doi.org/10.1016/j.jhydrol.2022.128116>
- Suresh Kumar G (2014) Mathematical modeling of groundwater flow and solute transport in saturated fractured rock using a dual-porosity approach. J Hydrol Eng 19(12):1–8. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000986](https://doi.org/10.1061/(asce)he.1943-5584.0000986)
- Swanson DA, Tayman J, Bryan TM (2011) MAPE-R: a rescaled measure of accuracy for cross-sectional subnational population forecasts. J Popul Res 28(2–3):225–243. <https://doi.org/10.1007/s12546-011-9054-5>
- Tao H, Awadh SM, Salih SQ, Shafik SS, Yaseen ZM (2022) Integration of extreme gradient boosting feature selection approach with machine learning models: application of weather relative humidity prediction. Neural Comput Appl 34(1):515–533. <https://doi.org/10.1007/s00521-021-06362-3>
- Tavakol-Davani H, Rahimi R, Burian SJ, Pomeroy CA, McPherson BJ, Apul D (2019) Combining hydrologic analysis and life cycle assessment approaches to evaluate sustainability of water infrastructure: uncertainty analysis. Water (Switzerland) 11(12):2592. <https://doi.org/10.3390/w11122592>
- Telesca L, Lovallo M, Shaban A, Darwich T, Amacha N (2013) Singular spectrum analysis and Fisher–Shannon analysis of spring flow time series: an application to Anjar Spring, Lebanon. Physica A 392(17):3789–3797. <https://doi.org/10.1016/j.physa.2013.04.021>
- Thonglek K, Ichikawa K, Takahashi K, Iida H, Nakasan C (2019) Improving resource utilisation in data centers using an LSTM-based prediction model. In: 2019 IEEE international conference on cluster computing (CLUSTER) 2019 September 1–8. <https://doi.org/10.1109/CLUSTER.2019.8891022>
- Tsung F-S (2010) Modeling dynamical systems with recurrent neural networks. Acad Med J Assoc Am Med Coll 85(9 Suppl):S92–S96
- Turner R, Eriksson D, McCourt M, Kiili J, Laaksonen E, Xu Z, Guyon I (2021) Bayesian optimisation is superior to random search for machine learning hyperparameter tuning: analysis of the black-box optimization challenge 2020, pp 3–26. <http://arxiv.org/abs/2104.10201>
- Williams PW (2009) Book review: Methods in karst hydrogeology Nico Goldscheider and David Drew (eds). Hydrogeol J 17(4):1025–1025. <https://doi.org/10.1007/s10040-008-0393-z>
- Xie T, Zhang G, Hou J, Xie J, Lv M, Liu F (2019) Hybrid forecasting model for non-stationary daily runoff series: a case study in the Han River Basin, China. J Hydrol 577:123915. <https://doi.org/10.1016/j.jhydrol.2019.123915>
- Yan X, Chang Y, Yang Y, Liu X (2021) Monthly runoff prediction using modified CEEMD-based weighted integrated model. J Water Clim Change 12(5):1744–1760. <https://doi.org/10.2166/wcc.2020.274>
- Yang S, Yang D, Chen J, Zhao B (2019) Real-time reservoir operation using recurrent neural networks and inflow forecast from a distributed hydrological model. J Hydrol 579:124229. <https://doi.org/10.1016/j.jhydrol.2019.124229>
- Yin H, Zhang X, Wang F, Zhang Y, Xia R, Jin J (2021) Rainfall-runoff modeling using LSTM-based multi-state-vector sequence-to-sequence model. J Hydrol 598:126378. <https://doi.org/10.1016/j.jhydrol.2021.126378>
- Yu Y, Zhang H, Singh VP (2018) Forward prediction of runoff data in data-scarce basins with an improved ensemble empirical mode decomposition (EEMD) model. Water (Switzerland) 10(4):1–15. <https://doi.org/10.3390/w10040388>
- Zhai N, Yao P, Zhou X (2020) Multivariate time series forecast in industrial process based on XGBoost and GRU, 2020(X), pp 1397–1400. <https://doi.org/10.1109/ITAIC49862.2020.9338878>
- Zhang Z, Wang W, Qu S, Huang Q, Liu S, Xu Q, Ni L (2018) A new perspective to explore the hydraulic connectivity of karst aquifer system in Jinan Spring. <https://doi.org/10.3390/w10101368>
- Zhou ZH, Wu J, Tang W (2010) Erratum: Ensembling neural networks: Many could be better than all [(Artificial Intelligence) (2002) 137:1–2:239–263]. Artif Intell 174(18):1570. <https://doi.org/10.1016/j.artint.2010.10.001>
- Zhu H, Xing L, Meng Q, Xing X, Peng Y, Li C, Li H, Yang L (2020) Water recharge of Jinan Karst Springs, Shandong, China. Water 12(3):694. <https://doi.org/10.3390/w12030694>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.