Journal of Hydroinformatics

Journal of Hydroinformatics Vol 24 No 5, 1066 doi: 10.2166/hydro.2022.009

Estimating rainfall–runoff modeling using the rainfall prognostic model-based artificial framework with a well-ordered selective genetic algorithm

Shailesh Kumar[a,*], K. K. Pandey[a], Sunil Kumar[b] and Sunidhi Supriya[c]

[a] Indian Institute of Technology (BHU), IIT-BHU, Banaras Hindu University Campus, Varanasi, Uttar Pradesh 221005, India
[b] Indian Institute of Technology, MAIN CAMPUS, Roorkee, Uttarakhand 247667, India
[c] National Institute of Technology, Hazratbal, Srinagar, Jammu and Kashmir 190006, India
*Corresponding author. E-mail: scholar.shaileshkumar213@gmail.com

ABSTRACT

Rainfall–runoff modeling is one of the most well-known applications of hydrology. The goal of rainfall–runoff modeling is to simulate the peak river flow caused by an actual or hypothetical rainfall force. In existing methods, the rainfall–runoff relationships are quantified to predict the daily streamflow of each catchment from its landscape attributes to measure the daily rainfall. However, the structural model error, infiltration rate, and the steep slopes of the hill affect the prediction process. To tackle these issues, this paper proposed a novel rainfall prognostic model-based artificial framework, which predicts day-to-day rainfall to prevent environmental disasters. The day-to-day predictions minimize the risks to life and property and also manage the agricultural farms in a better way because the possibility of rainfall has been estimated earlier. Furthermore, the posterior fire-breathing network is utilized to estimate model errors in the computational runoff by using time-dependent and random noise to the model's internal storage to solve the uncertainty problem. Since the model errors are estimated, there are limits to the infiltration rate and thus a prophetic multilayer network is utilized which relies on the soil runoff levels. Moreover, the network measures the dynamics of soil moisture to regulate the infiltration rate according to the rural or urban section. Moreover, to measure the surface water from the steep slopes, the system offered a well-ordered selective genetic algorithm to calculate the velocity of runoff in different bend areas to overcome the numerical problem. Thus, the model results showed that the work effectively predicts the rainfall from the investigation of model errors, infiltration rates, and velocity to achieve a better prediction range in the rainfall.

Key words: posterior fire-breathing network, prophetic multilayer network, rainfall prognostic artificial model framework, rainfall–runoff model, well-ordered selective genetic algorithm

HIGHLIGHTS

- Despite the progress made in recent years, modeling hydrological reactions to rainfall prediction remains a complex task in runoff modeling. Thus the presented paper effectively introduced a rainfall prognostic artificial model framework for the prediction of rainfall.
- The framework applied a posterior fire breathing network to estimate model errors with random noise to reduce the uncertainty. Further to regulating the infiltration rate, the system suggested a prophetic multilayer network which analyses the runoff levels with the soil moisture in urban and rural areas.
- In addition, to evaluate the velocity at low water depths on steep slopes, the model incorporates a well-ordered selective genetic algorithm to forecast different bend areas to overcome the numerical problem. Thus, from the rainfall model, the daily rainfall is efficiently predicted to prevent the environmental glitches.
- Experimental results show that the framework exhibits the highest prediction range of 12-20 mm with the subjective results, and outperforms the existing runoff models with a better infiltration rate of 1.5 cm/hr, runoff level 0.05 cm, and the obtained velocity of 0.45 mm.

1. INTRODUCTION

More recently, the unit hydrograph concept has been developed to conceptualize the reaction of a catchment to a storm event based on the theory of superposition. The unit hydrograph helps to isolate the runoff of the base flow and storm events from the streamflow (Fowler *et al.* 2016). Runoff models have become more complex with increased computational capacity and a deeper understanding of hydrological processes (Guo *et al.* 2017). Modeling runoff helps to understand hydrological processes better and how changes impact the hydrological process (Dakhlaoui *et al.* 2017). A runoff model is represented as

a sequence of equations that help to estimate how much rainfall functions runoff of the different parameters used to characterize the watershed. It can be difficult to model the surface runoff since the measurement is complex and includes several interconnected variables (Sitterson *et al.* 2018).

A model's overall modules include inputs, control equations, boundary or parameter conditions, model processes, and outputs (Garcia *et al.* 2017). It requires surface runoff simulation to consider yields of floodplain and reactions, predict water quality, change over time, and forecast. Hydrological Simulation Program-Fortran (HSPF1) is a hydrological model which as part of its management utilizes runoff functions for the prediction of sediment charges, nutrients, pollutants, harmful chemicals, and other concentrations of water quality (Fraga *et al.* 2019). While there are many ways to classify models, not all models fall into one category as they are designed for a variety of purposes. The water demand has increased as a result of fast population growth, industrialization, and urbanization (Kratzert *et al.* 2018). It is also one of the prime criteria for farming, domestic, and industrial activities. Nevertheless, the careful preparation and control of water supplies to mitigate property damage and lifelessness due to flooding is becoming very important (Nourani 2017).

According to the Asian Development Bank (ADB), from 1950 to 2011 Pakistan witnessed 21 floods which are mostly attributed to rainfall. Unexpected and strong rainfall contributes to overflowing dams, which are the leading cause of severe flooding (Ledesma & Futter 2017). Streamflow forecasting for a catchment area is one of the well-known and main hydrological variables since hydrological engineering, design, and actions for development and evaluation such as water supply, irrigation, flood control, and hydropower production require information on hydrological catchment stream flows (Song *et al.* 2019). The modeling of the rainfall-streamflow is considered to be complicated, nonlinear, and time-variable because the streamflow produced from a watershed depends not only on the hydro-meteorological parameters but also on the spatio-temporal irregularities in the features of the watershed and the patterns of rainfall (Xiang *et al.* 2020). Precisely forecasting streamflow helps to manage different water uses such as irrigation, hydropower development, preparing water delivery policies, and flood risk management (Jaiswal *et al.* 2020). Hydrologists have developed and used various types of models for modeling hydrological processes such as streamflow forecasting, prediction of rainfall, flood forecasting, and modeling of rainfall–runoffs, etc. (Nourani *et al.* 2019).

Additionally, the calculation of suspended sediment during a flood caused by heavy rain and cyclone is very difficult. Direct sediment yield measurement is complicated because it takes enough time and energy. Computational hydrology has benchmark work criteria in the sediment modeling area (Knoben *et al.* 2019). Many variables influence the yield of suspended sediments, and their partnerships with each other are extremely nonlinear and complex. Suspended sediment yields are therefore very difficult to estimate using traditional mathematical models such as multiple linear regression (MLR) and sediment rating curve (SRC) which do not provide satisfactory performance (Weeser *et al.* 2019). A typical feed-forward neural network has its limitations when it comes to time-series data. Feed-forward neural networks are built on the basis that the examples of training and testing (data points) are independent (Del Giudice & Padulano 2016).

For the efficiency of the spatial-temporal resolution, the entire network state is removed after each example has been processed. When data points are naturally connected, a conclusion is not needed. Besides, a typical artificial neural network (ANN) model (feed-forward) will allow a fixed-sized sliding window over the dataset to handle time-series data (Chlumecký *et al.* 2017). Tuning the scale of these sliding windows for the highest predictive precision adds additional work to the range of models. In finer time resolution flood assessments, the constraint is increasingly noticeable. Despite some examples of using a recurrent neural network (RNN) for hydrological modeling (Reshma *et al.* 2018), there is no long short-term memory (LSTM) model that forecasts future river flow solely based on precipitation inputs to address the issue of accumulative uncertainty in the literature. Therefore, to overcome the model errors, the infiltration rate with the soil moisture and steep slopes in the hill is a vital part to calculate in the runoff model for the prediction of daily rainfall. Thus, the research work proposed a novel framework for better prediction of rainfall with a runoff model. The main contribution of the paper is as follows:

- To estimate model errors in the computational runoff, the time-dependent and posterior fire-breathing network is utilized.
- The prophetic multilayer network is directly considering the relationship between infiltration and soil moisture through a collection of parsimonious computational parameters to estimate the infiltration rate.
- By utilizing a well-ordered selective genetic algorithm the velocity of runoff on steep slopes is calculated.

The remainder of the article is structured as follows. In section 2, the literature survey is presented in detail. The proposed rainfall prognostic model-based artificial framework is characterized in Subsection 3. Section 4 follows the result and the output of the proposed work. Finally, the overall work is resolved in Section 5.

2. LITERATURE SURVEY

Vidyarthi & Chourasiya (2020) introduced the use of the ANN in estimating the runoff of a river. The classical gradient descent (GD) algorithm is the most commonly used algorithm for training the ANN runoff models so far. The achievement of the GD algorithm, however, is affected by changes to get stuck at the local minimum. In the paper, one of the popular evolutionary optimization algorithms, known as particle swarm optimization (PSO), has been explored to train the ANN rainfall–runoff model. However, the accuracy of the models using evolutionary algorithms is low.

Asadi *et al.* (2019) Proposed Inputs from the hydrogeomorphic and biophysical period series, including the Normalized Vegetation Difference Index (NDVI) and Connectivity Index (IC), were evaluated in addition to climatic and hydrological inputs. They used selected inputs for the creation of ANNs in the catchment of the Haughton River and the catchment of the Calliope River, Queensland. Results show that IC is incorporated as a hydrogeomorphic parameter and NDVI. Besides, the availability of remote sensing land use datasets (i.e., NDVI) and high-resolution feature maps to reflect geomorphologic catchment features may also be a limiting factor.

Kumar *et al.* (2019) proposed the hydrological implementation of the ANN and the emotional neural network (ENN) for the simulation of rainfall–runoff in some order, Bihar as the area experiences flooding due to heavy rainfall. The ENN is a revised version of the ANN since it contains neural parameters that improve the process of network learning. The selection of inputs for the rainfall–runoff model is a key task. The paper makes use of a cross-correlation study to recognize possible predictors. The principal component analysis (PCA) was then conducted for the collection of data sets showing key patterns on the chosen data sets. The data sets collected after PCA were then used in the creation of the models ENN and ANN. Moreover, tremendous work is needed to estimate the reliability and performance of the models.

Üneş *et al.*'s (2019) implementation of the rainfall and runoff partnership is very important for the effective use of water supplies and flood prevention. Nowadays various methods of artificial intelligence (AI) techniques are used to assess the relationship between rainfall and runoff. The present research utilizes ANNs. Popular approaches like MLR are also used. The data collected from the US have been integrated into the report. The multi-linear regression and feed-forward back-propagation ANN models used 731 daily rainfalls, runoff, and temperature data to produce input data. However, the lack of evaluation of watershed characteristics creates issues in deciding the relationship between rainfall–runoff results.

Chadalawada *et al.* (2020) proposed a novel model-building algorithm using the full potential of flexible modeling frameworks by searching for model space and using machine-learning (ML)-based model configurations. The proposed algorithm for ML is based upon an evolutionary approach to computation using genetic programming (GP). Up until now, state-of-the-art GP modeling implementations have used the algorithm as a short-term forecasting tool that produces an expected future time-series very similar to neural network implementations. However, no sufficient observations were provided in the catchment area.

Chen & Chau (2019) conducted uncertainty analysis on the hybrid double feed-forward neural network (HDFNN) model for generating the sediment load prediction interval (PI). They used the LUBE technique, in which the lower and upper bounds are created directly as outputs of neural network-based models. Partitioning analysis demonstrates that the HDFNN model consistently performs well in creating PI for low, medium, and high loads. However, the LUBE approach does not solve highly nonlinear, discontinuous, and non-differentiable optimization problems.

Fotovatikhah *et al.* (2018) used flood management systems (FMSs) in their research to aid in the decision-making process in crucial scenarios. They looked at AI and computational intelligence (CI) approaches for detecting early flood events with a low false alarm rate. The city administration was able to respond quickly and effectively in post-disaster scenarios using this strategy. They came to the conclusion that using an ensemble CI technique to predict floods was quite effective. However, using an ensemble CI technique is not always the best option. Furthermore, the ensemble method is incapable of addressing unknown discrepancies between the sample and the population.

Amir Mosavi *et al.* (2018) analyzed the state of the art of ML models in flood prediction. They investigated providing an extensive overview on the various ML algorithms used in the field through a qualitative analysis of robustness, accuracy, effectiveness, and speed.

They found that the effective strategies are hybridization, data decomposition, algorithm ensemble, and model optimization. However, the hybridization method is expensive, costing up to five times the value of the normal process. The data decomposition method, ensemble algorithm, and model optimization techniques suffer if not provided with the normal requirements.

Taormina & Chau (2015) tested the suitability of the LUBE approach in producing PIs at different confidence levels (CLs) for the 6 h ahead streamflow discharges of the Susquehanna and Nehalem Rivers, US. The PSO in LUBE applications, variants of this algorithm had been employed for CWC minimization and they were found to vary substantially depending on the chosen PSO paradigm. The algorithm has a low convergence rate in the iterative process so that advanced algorithm should be utilized for prediction.

Wu & Chau (2013) employed several soft computing approaches for rainfall prediction. They utilized moving average (MA) and singular spectrum analysis (SSA) techniques for preprocessing. The modular model was composed of local support vector regression (SVR) models or/and local ANN models. Modular models involved preprocessing the training data into three crisp subsets (low, medium, and high levels) according to the magnitudes of the training data, and finally two SVRs were performed in the medium- and high-level subsets. Also, ANN or SVR was involved in training and predicting the low-level subset. The ANN-MA displayed considerable accuracy in rainfall forecasts compared with the benchmark, although, the neural network should include back-propagation to increase the speed of the training process.

In Vidyarthi & Chourasiya (2020) the accuracy of the models is low, moreover, in the study by Asadi *et al.* (2019), availability of remotely sensed datasets has a limiting factor. In Mosavi *et al.* (2018), the hybridization method costs up to five times the value of the normal process. Data decomposition, ensemble algorithm, and model optimization techniques suffer if not provided with the normal requirements. From the aforementioned issues, it is essential to develop a new framework for the rainfall–runoff model for enhanced prediction.

3. A RAINFALL PROGNOSTIC MODEL-BASED ARTIFICIAL FRAMEWORK

For modeling the rainfall–runoff process, models have been developed that are based on conceptual representations of the physical processes of the water flow lumped over the entire catchment area (lumped conceptual type of models). The parameters of such models cannot, in general, be obtained directly from measurable quantities of catchment characteristics, and hence model calibration is needed. To calibrate a model, values of the model parameters are selected so that the model simulates the hydrological behavior of the catchment as closely as possible. The process of model calibration is normally done either manually or by using computer-based automatic procedures. In manual calibration, a trial-and-error parameter adjustment is made. In this case, the goodness-of-fit of the calibrated model is based on a visual judgment by comparing the simulated and the observed hydrographs. However, since there is no generally accepted objective measure of comparison, and because of the subjective judgment involved, it is difficult to assess explicitly the confidence of the model simulations. The development of such relationships is an example of a regionalization methodology and is a useful goal for a number of reasons. For example, the construction of a hydrologic structure such as a bridge or dam may require a prediction to be made of the hydrologic response of a catchment at an ungauged point.

Rainfall prediction is the integration of science and technology to estimate the atmosphere's environment. In order to effectively utilize water supplies, crop production, and pre-planning water systems, the rainfalls should be calculated in advance. Various strategies for runoff can avoid precipitation, and accurate input of rainfall in time and space is crucial to the response of rainfall–runoff modeling. Recently conceptual rainfall–runoff models use basic spatially lumped storage-flow relationships to compensate for the spatial complexities of hydrological processes within a basin. These rough approximations introduce model errors that are hard to describe. Also, it is imperative to estimate uncertainties in measured data (e.g., discharge), model inputs (e.g., rainfall, parameters), and the model itself. Failure to consider any sources of uncertainty would result in an underestimation of the prediction ability. While formulating an effective model error and parameterization, the parameters (e.g., variance of random noise) must be calculated from the entire available data. The hybrid AI framework is designed by utilizing the novel rainfall prognostic model-based artificial framework which uses the posterior fire-breathing network to estimate model errors. Also, the model errors are estimated by a prophetic multilayer network that relies on soil runoff levels. Furthermore, to tackle the numerical challenge of measuring surface water from steep slopes, the system used a well-ordered selective genetic algorithm to determine the velocity of runoff in distinct bend areas. Further, to determine the runoff level in the state-of-the-art methods, many limitations lead to a high infiltration rate. In such modeling methods, runoff rates from metropolitan centers are believed to be a flat rate value that is essentially disconnected from the underlying soil moisture. Since soils vary from undeveloped areas to urban areas due to the mixture of compaction and synthetic materials in the soil, resulting in different infiltration rates than uncompacted soils. As a result, runoff in urban areas is growing due to impermeable surfaces; soils also impact runoff rates in urban green areas (such as parks and sports fields). Furthermore, to protect the surface water from

steep slopes face velocity measurements on different scales and a watershed application evaluates the velocity field. Many works are involved in describing the fields of velocity in a particular case and some works that tackle numerical problems. The low water depths run off steep slopes which contributes to the physical representation of the friction at the bottom for specific water depth to height ratios of roughness. Nonetheless, the capability can be very difficult in mountain environments like heterogeneous and high low slope values as well as intermittent rain inputs. Therefore, the paper proposed a novel rainfall prognostic artificial model framework, which relies on the models of ANNs. The architecture of the proposed framework is shown in Figure 1.

Initially, the framework introduced and applied a novel posterior fire-breathing network to estimate model errors of a conceptual rainfall–runoff model and to perform uncertainty analysis. The solution accounts for model errors by adding time-dependent, random noise into the internal storage of the hydrologic model. The model also effectively quantifies variability in the performance and the parameters arising from variance in the input data used to calibrate. Furthermore, the level of runoff is simulated using the Prophetic Multilayer Network technique which is employed to model infiltration. The technique is also found to be related to the input as the exactness of the ANN model was higher for soil types with higher proportions in the training data. The fraction of precipitation converted into either runoff or infiltration depends on the amount of soil moisture, and the dynamics of soil moisture vary according to the model's rural or urban portion. In addition to overcoming runoff at low water depths on steep slopes, the framework proposed a well-ordered selective genetic algorithm (WSGA) for velocity measurements and predictions in different bend sections and comparing their test results with each other and with actual data. Four experimental test cases have been used to explore and evaluate the numerical model and to understand its behavior better. The overall rainfall modeling is given in the following sections.

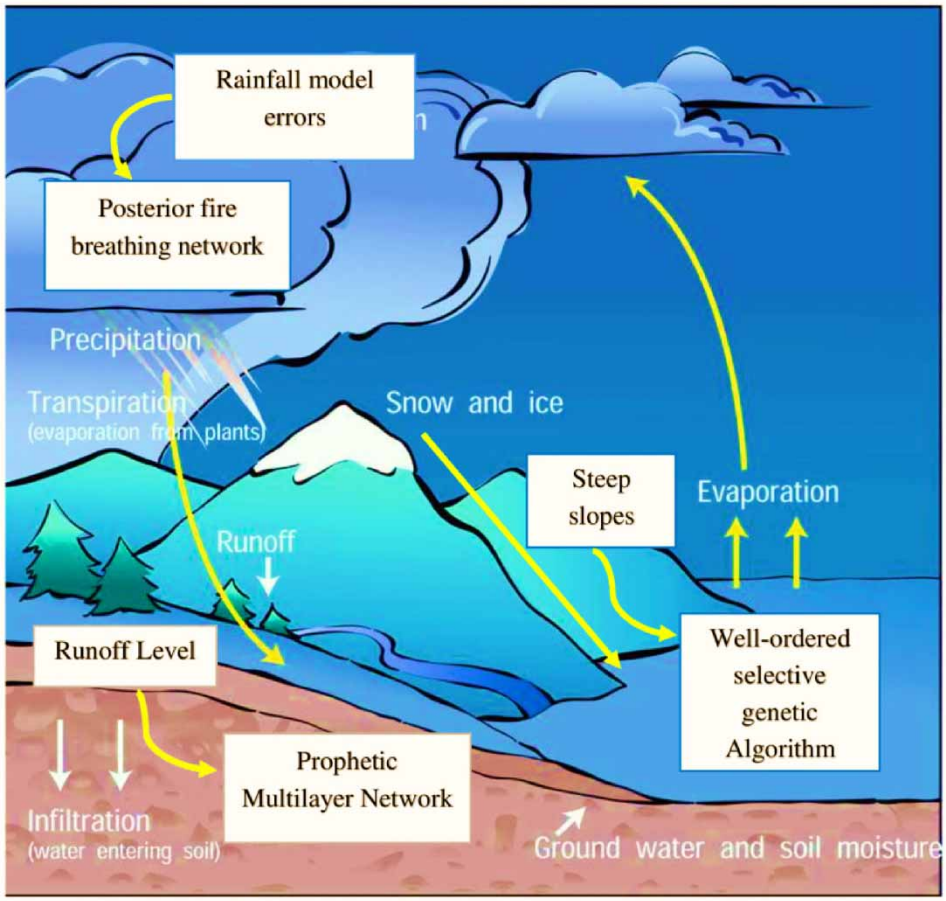

Figure 1 | Architecture of the rainfall prognostic model-based artificial framework.

3.1. Posterior fire-breathing network

The network calculates a conceptual rainfall–runoff process's model errors and analyzes uncertainty. The model error assessment assumes that the results of the discharge reflect exactly the hydrological behavior of the catch. However, time-series are successively extracted from stage measurements and conversions of stage ratings and therefore may contain highly uncertain values. Thus, by incorporating time-dependent, spontaneous noise into the internal storage of the hydrological model, the solution accounts for model errors and thus efficiently quantifies uncertainty in model output and parameters resulting from variation in the input data used for calibration. Adoption of the posterior fire-breathing network in driving a rainfall–runoff model reduces the error by estimating the runoff at the catchment outlet with different rainfall inputs from different parts of the catchment. Initially, the noise parameter of the model errors was calculated and revised sequentially through the approximation of the Gamma density to its posterior. With the approach suggested in the article, Gamma approximations were obtained in these graphs, while Gaussian approximations have the same mean and variance as the Gamma approximations. These observations suggest that the posterior is originally asymmetric and non-Gaussian, but as more data is assimilated, it tends to become more Gaussian.

The objective is to recursively estimate posteriors for the model error precision including for the Hymod states and parameters using the posterior fire-breathing network. Each time a major discharge discovery appears accessible it is incorporated by

- updating the posterior of τ (precision of model errors) and
- updating the states and parameters of Hymod.

It is realized that the model error accuracy posterior is modified separately from the state and parameter posterior. The alternative of introducing the precision as an external unknown to the state vector, i.e., the state-augmentation method, is never considered here. Information on mathematics to update posteriors is given in the following two sections.

3.1.1. Sequential updating of the posterior of τ

The posterior precision or inverse variance of the model errors is calculated sequentially in the system. It is accomplished by preserving a Gamma density approximation of the posterior: as new data become available the approximation is modified in real-time to represent new knowledge on model errors. Thus, the Gamma approximations are used for their corresponding calculations of the accuracy (or variance) of model errors a positive sum. In Bayesian analysis, the Gamma density is a popular approach for modeling precision variables, partially even with its mathematical properties; e.g., a Gamma prior is conjugated to a Linear-Gaussian probability, leading to posterior closed-form expressions. In the article, work with differential equations, so miss conjugacy and the posterior is not accessible analytically, but they can be effectively estimated, as mentioned below. It factors in an approximate but effective recursive approach for determining the posterior of τ that is completely Bayesian and extends point (posterior mode) estimation of the model noise.

Mathematically, the posterior of τ at time t is written recursively in terms of the posterior at the time $t-1$ and the likelihood at the time t. For model noise on a selected flow or state variable x, the formula is:

$$p(\tau|D_t)\alpha p(\tau|D_{t-1})s\int_{x_t}\int_{\mu_t} p(\mu_t|D_{t-1})N(x_t|\mu_t,\tau^{-1})p(D_t|x_t)d\mu_t dx_t \quad (1)$$

where D_t is the current discharge measurement, and $D_t = (D_{t-1}, D_t)$ which holds all up to and including the date t of discharge. The prior of τ at the time t is inscribed as s. The double integral represents the likelihood τ at the time t: $p(\mu_t|D_{t-1})$ is the predictive distribution of the chosen flow or state variable before adding noise, $N(x_t|\mu_t, \tau^{-1})$ is the model noise term $x_t \sim N(\mu_t, \tau^{-1})$, and $p(D_t|x_t)$ is a likelihood at time t of the true flow or state (after adding noise).

Rectangles in Figure 2 denote reservoirs for which a water balance is determined. Model error is applied to the slow-reservoir state as an example. The changes in areas that make up the posterior τ are seen in blue. Equation (1) can be understood by reference to the model error depicted in Figure 2, which shows the case of model error added to the slow-reservoir state S_s. The case μ_t corresponds to the modeled state $S_{s,m}$ ('m' indicates modeled) before adding model error and x_t corresponds to the true state S_s after adding model error. To update the posterior for τ at time t, three sources of information are combined (arrows in Figure 2):

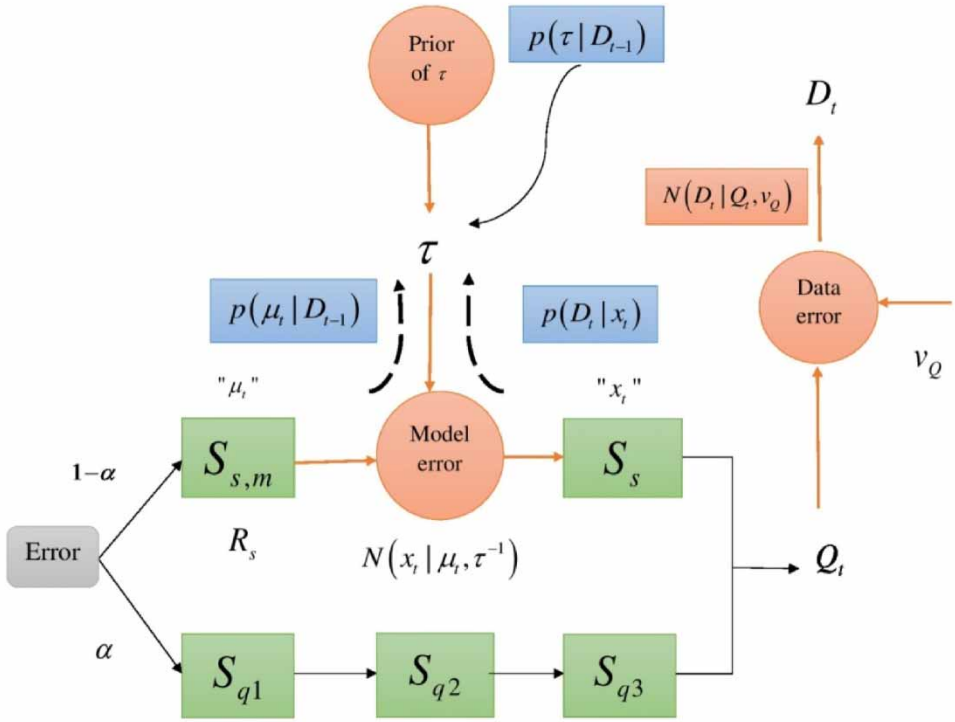

Figure 2 | Schematic diagram of the deterministic (black color) and stochastic (red color) components of the modified Hymod model. Please refer to the online version of this paper to see this figure in color: http://dx.doi.org/10.2166/hydro.2022.009.

- the prior for τ (i.e., the posterior for τ at the time $t-1$),
- the predictive distribution of S_s, $p(\mu_t|D_{t-1})$ and
- probabilistic information on S_s provided by the current discharge observation $p(D_t|x_t)$

As Figure 2 shows, the latter two are combined $(x_t|\mu_t, \tau - 1)$ to provide a probabilistic information stream τ, which constitutes the likelihood term for τ and it corresponds to the double integral in Equation (1). To make the double integral in Equation (1) tractable, both the predictive distribution $p(\mu_t|D_{t-1})$ and likelihood $p(D_t|x_t)$ are approximated by Gaussian distributions. The predictive distribution $p(\mu_t|D_{t-1})$ is approximated by $N(\mu_t|\mu_\mu, v_\mu)$ where μ_μ and v_μ are mean and variance, respectively, of an ensemble of predicted values μ_t, which is the selected flow or state variable before adding model noise. Likewise, likelihood $p(D_t|x_t)$ is approximated by $N(x_t|\mu_x, v_x)$, where μ_x and v_x are estimated by linearizing via linear regression the relation between predicted Q_t and x_t

$$Q_t = \mu_{Q,pred} + \psi(x_t - \mu_{x,pred}) \text{ or } x_t = \frac{Q_t - \mu_{Q,pred}}{\psi} + \mu_{x,pred}.$$

The last equation is then used to translate mean D_t and variance v_Q for the likelihood of Q_t into mean and variance of x_t:

$$\mu_x = \frac{D_t - \mu_{Q,pred}}{\psi} + \mu_{x,pred} \tag{2}$$

$$V_x = \frac{v_Q}{\psi^2} \tag{3}$$

where $\mu_{Q,pred}$ and $\mu_{x,pred}$ (ensemble) means of are predicted Q_t and x_t, respectively, and ψ is the linear regression coefficient

between the two ensembles. With these Gaussian approximations, we can rewrite Equation (1) in the following form:

$$p(\tau|D_t)\alpha p(\tau|D_{t-1})\int_{x_t}\int_{\mu_t} N(\mu_t|\mu_\mu, v_\mu)N(x_t|\mu_t, \tau^{-1})N(x_t|\mu_x, v_x)d\mu_t dx_t \tag{4}$$

As all terms under the integral are now Gaussian, the double integral can be easily computed (see Appendix A). While this yields a closed-form formula for the posterior, the resulting density does not correspond to a known parametric density. To have an interactive recursive estimation algorithm, i.e., one that makes fast posterior sampling and eliminates the need to repurpose data and likelihoods from specified time measures, it recommends estimating the posterior by a new Gamma density. Under such estimation, the posterior will increase in complexity every time step, for each time steps an additional possibility term is introduced that transforms the posterior into a non-standard dimension. Therefore, the mathematical expression for the posterior increases in complexity even as the posterior is one-dimensional. The Gamma approximation avoids and preserves a reliable estimation of the posterior, and therefore enables recursive updating rather than going back to different results, while still making it simple to analyze the form.

Hence, at each time step, the posterior τ is represented by a Gamma distribution: $p(\tau|D_{t-1}) \approx Ga(\tau|\alpha_{t-1}, \beta_{t-1})$ *and* $p(\tau|D_t) \approx Ga(\tau|\alpha_t, \beta_t)$. The distributor allows us to derive closed-form recursive update equations for α_t and β_t

$$\beta_t = \alpha_{t-1} - \tau^2 \frac{d^2 \log f(\tau)}{d\tau^2} \tag{5}$$

$$\beta_t = \beta_{t-1} - \frac{d \log f(\tau)}{d\tau} - (\alpha_t - \alpha_{t-1})\tau^{-1} \tag{6}$$

$$\tau = \frac{\alpha_t - 0.5}{\beta_t} \tag{7}$$

where expressions for $\log f$ and its derivatives are given in the appendix. Since α_t and β_t in Equations (5) and (6) depend on τ, update τ in Equation (7), and recomputed α_t and β_t again with Equations (5) and (6). The update in Equation (7) sets τ to a value that lies in between the mean (α_t/β_t) and the mode $((\alpha_t - 1)/\beta_t)$ of the new Gamma density approximation. At each time t, these three equations are iterated 10 times to yield a Gamma approximation of the posterior of τ for that time step. Iterate 10 times but typically observe convergence after 2–5 iterations. The iteration could be made more sophisticated by monitoring convergence of the iterations and using a stopping criterion, but the computational cost of 10 iterations is small. As new data come in, these update equations result in an automatic tuning of the amount of model noise τ so that the resulting probabilistic prediction is neither too wide nor too narrow. The use of gamma approximation calculates the time between independent events that occur at a constant average rate. Using this approximation, analysts specify the number of events, such as modeling the time until the nth accident occurs.

3.1.2. Sequential updating of the Hymod state and parameter posteriors

The states and parameter posteriors are expressed by ensembles, which are produced in the prediction phase illustrated above. The ensemble Kalman Filter (EnKF) is used to modify the actual flow observation of an ensemble participant which is ideal for systems that are continuously changing. The advantages of EnKF are it is light on memory, thus there is no need to keep any history other than the previous state; and very fast, making it well suited for real-time problems and embedded systems than the particle filter.

$$(S_t, \theta|D_t)_i = (S_t, \theta|D_{t-1})_i + K_t(\widetilde{D}_t - Q_t)_i \;\; i = 1, 2, \ldots..N \tag{8}$$

where $\widetilde{D}_t$ represents randomly perturbed observed discharge (with perturbation variance v_Q). On the right-hand side of Equation (8), predicted values for states S_t and discharge Q_t are generated by the sampling process described in the predict step. Finally, K_t Equation (8) is the Kalman gain vector and is written as:

$$K_t = V_{(S_t,\theta),Q_t}(v_{Q,pred} + v_Q)^{-1} \tag{9}$$

where $V_{(S_t,\theta),Q_t}$ is a vector of ensemble cross-covariance between states-parameters (S_t, θ) and predicted discharge Q_t and $v_{Q,pred}$ is the ensemble variance of predicted discharge Q_t.

$$[S_t] > 0 \tag{10}$$

$$Ib < [\theta] < ub \tag{11}$$

$$SMc < \frac{C_{\max}}{(\beta + 1)} \tag{12}$$

An upgrade to EnKF values for states and parameters outside their physical range may result in Equation (8). With the purpose of preceding the EnKF update, the accompanying restrictions are extended to states and the model parameters with better optimization in the gamma distribution. Furthermore, to approximate the penetration and runoff level the system integrates a prophetic multilayer network which is concise in the following section.

3.2. Prophetic multilayer network

The paper explores whether linking soil moisture and urban runoff will facilitate model simulations that overtake fixed percentage runoff. A modern approach to modeling infiltration and runoff levels across urban surfaces is implemented for the purpose, which directly considers the relationship between infiltration and soil moisture through a collection of parsimonious computational parameters. The models of infiltration are incorporated within a framework of uniform rainfall–runoff. The prophetic multilayer network allows for probability-based predictions and the classification of items into multiple labels. The advantages of a prophetic multilayer network are capability to learn nonlinear models and models in real-time.

The prophetic multilayer network is trained by a back-propagation algorithm, which is fast and easy to program and it does not require additional parameters to tune apart from the numbers of input, nor does it require prior knowledge about the network. Back-propagation ANN is used to model the relationship between experimentally obtained scattering coefficient data and the degree of ground unevenness, and the soil moisture content of the field. As a system of linked layers of artificial neurons, the BP network is structured, input, hidden, and output layers are shown in Figure 3 as a simple representation of such a network. When data are provided to a neural group through the input layers, the input layer neurons spread the weighted data and randomly draw bias via the hidden layers. The output at each node of the hidden layer is then calculated consuming a transmission purpose. Afterward, the generated signals are transferred either to the next hidden layer or the output layer.

The prophetic multilayer network incorporates a deterministic, continuous-time, lumped, computational rainfall–runoff model intended to simulate catchment runoff like the urban land-cover model (URMOD) for sustainable development. A

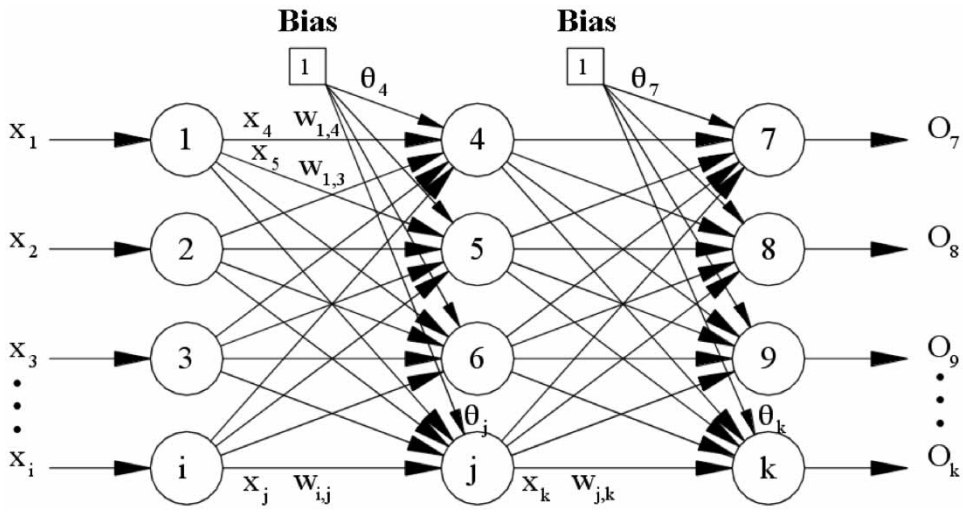

Figure 3 | The network layer.

lumped model doesn't break down catchments into different sub-catchments but takes into consideration a single catchment with only one outlet. The hydrological implications of urbanization are directly accounted for by dividing a catchment into a rural and an urban portion, with specific infiltration (and therefore runoff) and routing features allocated to each segment. URMOD has nine parameters requiring calibration of the observed precipitation, possible evaporation, and river flow.

Figure 4 shows the processes of infiltration and runoff with the conceptual soil column and three sectors, sector 2 with field capability but exceeding the rooting depth ($R < m < F$), and sector 3 with soil moisture below rooting depth ($m < R$) where soil moisture regulates infiltration and runoff generation.

3.2.1. Infiltration and runoff line of attack

A soil column-based approach is used to model the surface runoff and infiltration. The water which does not enter into the column of the soil is transformed into a direct runoff. The proportion of precipitation that turns into either runoff or absorption depends on the amount of soil moisture, and soil moisture characteristics differ according to the rural or urban component of the model. In the rural region, the temporal increase in soil moisture is induced by three processes: (i) infiltration, (ii) runoff, and (iii) evaporation. When the urban land-cover is present then two inputs will influence the infiltration through the catchment; infiltration from rural areas and industrial areas, respectively. The cumulative infiltration from the two land-cover classifications is expressed as a weighted average infiltration (f) of:

$$f = i(1-u)f_{rur} + iuf_{urb} \tag{13}$$

U is the fraction of the total catchment area covered by urban land-cover where there rainfall f_{rur} represents rural infiltration and is defined in Equation (14), and they signify infiltration in urban areas f_{urb}. The rural infiltration indicated in Equation (14) is based on the PDM model.

$$f_{rur} = \left(1 - \frac{m}{s}\right)^{1/2} \tag{14}$$

where m is the soil moisture content (mm), and S is the soil column capacity (mm), thus $0 \leq m/S \leq 1$. When the column of the soil reaches saturation ($m/S \approx 1$), the infiltration is poor and the rest of the rain is turned into a direct runoff. It is assumed that the conceptual soil column has three separate zones reflecting soil moisture levels, all of which are balanced parameters, controlled by field ability (F) and rooting depth (R). Drainage and soil column evaporation depend on the degree of soil moisture as shown in Figure 4.

Section 1, near the soil surface, is defined as if soil moisture is above the capacity of the field ($m > F$). In this case, the possible rate of evaporation (E_p) is assumed, and drainage to deeper storage depends on moisture content (m) and a calibrated drainage coefficient (k) so drainage from the column occurs at a rate of k $(m - F)$. Then, the equation for the region is

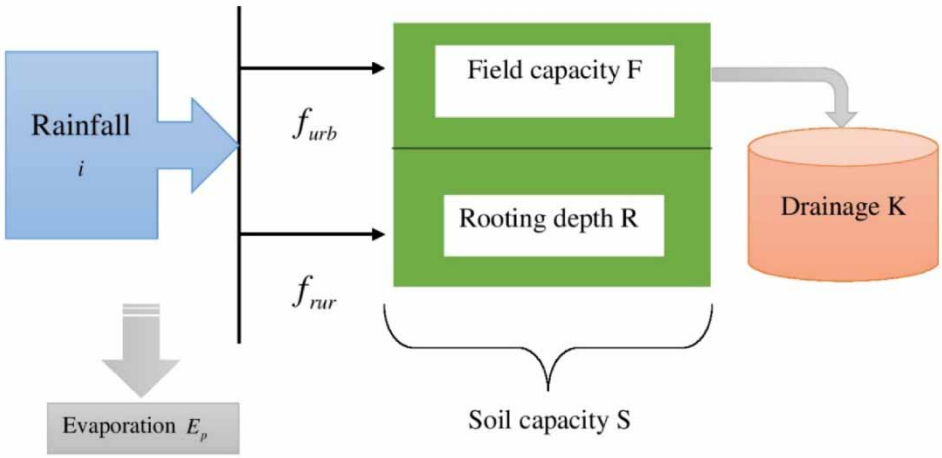

Figure 4 | Visual representation of the URMOD model.

shown in Equation (15).

$$\frac{dm}{dt} = \underbrace{i(1-u)\overbrace{\left(1-\frac{m}{S}\right)^{\frac{1}{2}}}^{f_{rur}} + iuf_{urb}}_{Infiltration} - \underbrace{k(m-F)}_{Drainage} - \underbrace{E_p}_{Evaporation} \tag{15}$$

Sector 2 is where the soil moisture does not exceed the potential of the field but exceeds the rooting depth ($R < m < F$), the evaporation is again at the highest rate (E_p) but no drainage takes place. Then, the equation for the region is shown in Equation (16).

$$\frac{dm}{dt} = \underbrace{i(1-u)\overbrace{\left(1-\frac{m}{S}\right)^{\frac{1}{2}}}^{f_{rur}} + iuf_{urb}}_{Infiltration} - \underbrace{E_p}_{Evaporation} \tag{16}$$

In sector 3, when soil moisture is below the rooting depth ($m < R$), there is again no drainage, and evaporation reduces linearly with depth $E = E_p(m/R)$ until it reaches $E = 0$ for $m = 0$. Three different differential equations explain the distribution of soil moisture in each of the three different regions. The definition of infiltration does not alter in any of these equations and is defined by Equation (14). Therefore, the zone equation is shown in Equation (17).

$$\frac{dm}{dt} = \underbrace{i(1-u)\overbrace{\left(1-\frac{m}{S}\right)^{\frac{1}{2}}}^{f_{rur}} + iuf_{urb}}_{Infiltration} - \underbrace{E_p\frac{m}{R}}_{Evaporation} \tag{17}$$

A leveling concept for the amount of evaporation occurring in urban areas is applied. Although there is no agreement on the importance of evaporation in urban areas in modeling studies, importance is agreed to be lower than the total in rural areas and greater than no evaporation. The first urban extension assumes a fixed percentage of infiltrated rainfall, thus de-coupling the hydrological cycle on urban surfaces with soil moisture. The second extension presumes the generation of runoff and infiltration in urban areas are dependent on a scaling term denoted γ, the urban surface infiltration depends on the soil moisture content of the rural areas but is decreased by a factor $(1-\gamma)$. It thus links urban infiltration directly to soil humidity levels.

3.2.2. Urban enlargement 1: fixed percentage runoff

Urban enlargement 1 assumes a fixed proportion of runoff from the metropolitan region, and typically runoff is independent of soil moisture. The fixed percentage runoff will be denoted ω. Comparing ω the runoff generated for the rural areas it is obvious that at ω a threshold would be defined since there are high levels of soil moisture that the amount of runoff generated from rural areas exceeds urban runoff rates, which is considered counter-intuitive. The threshold for soil moisture m/s at which the shift occurs is derived as:

$$\underbrace{1-\left(1-\frac{m}{S}\right)^{\frac{1}{2}}}_{rural} > \underbrace{\omega}_{urban} \Rightarrow \frac{m}{S} > 1-(1-\omega)^2 \tag{18}$$

For the soil moisture levels above and below the point, therefore, penetration from urban areas must be addressed. When the amount of soil moisture reaches this threshold, urban area infiltration can return to actions as in rural areas. Therefore, the infiltration into urban areas is described as:

$$f_{urb} = \begin{cases} 1-\omega: & 0 \le m/S \le 1-(1-\omega)^2 \\ \left(1-\frac{m}{S}\right)^{\frac{1}{2}}: & 1-(1-\omega)^2 < m/S \le 1 \end{cases} \tag{19}$$

By replacing Equations (14) and (19) with Equation (13), total infiltration can be described as urban and rural.

$$f = \begin{cases} (1-u)\left(1-\frac{m}{S}\right)^{\frac{1}{2}} + iu(1-\omega): & m/S \le 1-(1-\omega)^2 \\ \left(1-\frac{m}{S}\right)^{\frac{1}{2}}: & m/S > 1-(1-\omega)^2 \end{cases} \tag{20}$$

While ω would traditionally be set as a fixed value such as 70 or 100%, evaluating the specific amount of valuation that depends entirely on urbanization is categorized as a calibrated value for the research.

3.2.3. Extension 2 to urban infiltration: multiplicative impacts on urban areas

Urban extension 2 believes that urban infiltration relies on soil moisture, comparable to rural infiltration but reduced by a multiplicative factor $(1-\gamma)$. Extension 2 thus prevents the explicit implementation of a threshold as provided in extension 1. The functional form is defined as:

$$f_{urb} = (1-\gamma)\left(1-\frac{m}{S}\right)^{1/2} \tag{21}$$

The calibration parameter $\gamma \in [0, 1]$ is implemented to account for the variability of infiltration across various urban catchments so that a large one indicates that the urban area is largely impervious and more runoff is generated, while a smaller value indicates that the urban area has more perennial surfaces and less runoff is produced. If γ is zero then penetration is the same for the impervious area as the rural area, whereas if γ is one area then the area will be fully sealed and no intrusion will occur. In substitution of Equations (14) and (21) into Equation (13), the complete infiltration is derived as:

$$f = \underbrace{i(1-u)\left(1-\frac{m}{S}\right)^{\frac{1}{2}}}_{rural} + \underbrace{iu(1-\gamma)\left(1-\frac{m}{S}\right)^{\frac{1}{2}}}_{urban} = i(1-u\gamma)\left(1-\frac{m}{S}\right)^{\frac{1}{2}} \tag{22}$$

Equation (22) defined for infiltration can then be replaced by the three soil moisture equations. Thus, the infiltration rate in the soil moisture and runoff level is determined successfully in the urban and rural areas. Then the effect of steep slopes in the low depth of hills on the surface water is estimated by the velocity with a novel genetic algorithm, discussed below.

3.3. Well-ordered selective genetic algorithm

The well-ordered selective algorithm evaluates the importance of a mathematical model to determine the velocity profile in a watershed with rain-induced runoff, using only a displacement rule customized to every disturbance and upwelling regime. More broadly, for the representation of the velocity field, we describe the speed field in each situation and trust the numerical model.

The two-dimensional deep ocean models are transformed from the Naiver–Stokes equations with hydrostatic pressure and standardized vertical velocity statements where the flow is deep and the slope fluctuations are not increasing, certain assumptions are accessible which is written as:

$$\begin{cases} \dfrac{\partial h}{\partial t} + \dfrac{\partial hu}{\partial x} + \dfrac{\partial hv}{\partial y} = R - I \\ \dfrac{\partial hu}{\partial t} + \dfrac{\partial (hu^2 + (gh^2/2))}{\partial y} = gh\left(-\dfrac{\partial z}{\partial y} - S_{fx}\right) \\ \dfrac{\partial hv}{\partial t} + \dfrac{\partial huv}{\partial x} + \dfrac{\partial (hv^2 + (gh^2/2))}{\partial y} = gh\left(-\dfrac{\partial z}{\partial y} - S_{fy}\right) \end{cases} \tag{23}$$

where t is the time, h is the height of the water, v the velocity of flow in the y-direction, u the velocity of flow in the x-direction R the strength of the rain, g the gravity constant, z the bottom elevation, and S_{fx} *and* S_{fy} the friction slope respectively in the x and y directions.

3.3.1. Numerical resolution

An unstructured mesh is used to discretize the domain in space. The model enables better adaptation than structured mesh to a basin's complex topography by better representing its topography, especially the river and gully shape in the river system. Also, these meshes lie in the ability to deal with arbitrary complex geometries and the ease with which local refinements of the grid are to be implemented. At each node of the mesh, a control volume is constructed creating a boundary between each neighboring node of the mesh, as shown in Figure 5.

The Shallow Water Equation (23) can be solved with a finite volume scheme written as:

$$U_i^{t+1} = U_i^t - \sum_{j \in V_i} \left(\frac{\Delta t}{\Delta x_{ij}} F_{ij}^t + \frac{\Delta t}{\Delta x_{ij}} S_{ij}^t \right) \tag{24}$$

where $U_i^t = (h_i^t, h_i^t u_i^t, v_i^t)$ the states at a time t in the node P_i, V_i all the nodes surrounding P_i, Δ_t the time step, $\Delta x_{ij} = C_i / L_{ij}$ the space step with C_i the area of the cell constructed around the node P_i and L_{ij} the length of the boundary edge between the cells C_i and C_j,F_{ij}^t is the numerical value at the interface between the cells C_i and C_j along n_{ij}, the normal vector to the boundary edge and $S_{ij}^t = (R_i - I_i, s_{ijx}^t, s_{ijy}^t)$ are the source terms of mass and momentum along the x and y directions. The key challenge in solving these equations is how to treat the terms slope source s_{ijx}^t and s_{ijy}^t to respect the water depth and hydrostatic balance positivity and, more specifically, for wet/dry transitions and steep slopes. The Shallow Water Equations are solved by scheme (24).

A fully wet state when the water depth is higher than the bottom difference between two cells, and a partially wet regime otherwise, are distinguished at the interface of two cells. Consider the non-dimensional number $\beta = h/\Delta x \partial_x z$; if β is inferior to 1 the regime is partially wet and fully wet otherwise. The exact error made on the source term can be estimated. Then the experimental test cases are to be carried out.

3.3.2. Experimental test cases

Four test cases are selected at different scales and degrees of complexity. The first two laboratory test cases are selected with simple geometry. The first laboratory case is used to evaluate the effect of the resolution method and the other one to assess the effect of Friction Law on local velocities. Then on a small plot, the model is applied to a real case to confirm the results of the two previous test cases. Finally, the model is tested over a watershed of around 1 km^2 on which the discharge at the outlet can be repeated.

The first experiment is a flume with a constant slope on which a constant raindrop is used to measure the error due to the numerical scheme. The second test case is a furrow-shaped sinusoidal mold reflecting the agricultural land. In comparison with measurements, the conclusions of these simulated results are then represented using two laboratory experiments

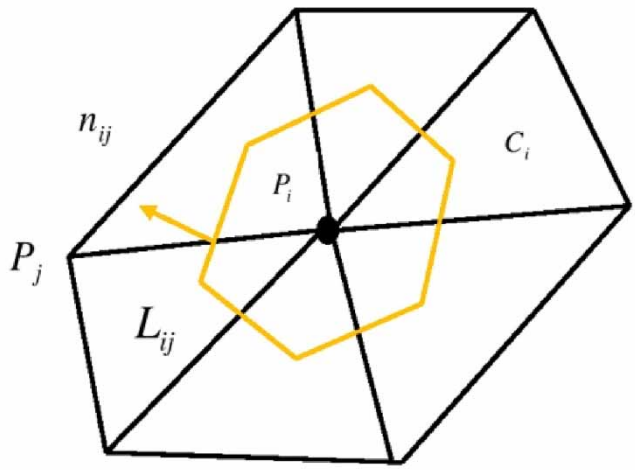

Figure 5 | Representation of the formalism used for the 2D unstructured mesh resolution.

obtained on this test case. Eventually, the model was applied to a specific catchment in the Laval where only outlet discharge measurements are available.

3.3.2.1. Flume experiment. A steady stream, with an intensity of 25 or 50 mm/hour depending on the chosen scenario, is applied with a stream simulator over the domain over 600 s. With a time phase of 0.1 s the temporal evolution of the discharge at the outlet of the channel is observed. The temporal evolution of the discharge is achieved by measuring the water that flows from the canal. Once the steady-state is achieved, the water depth and the velocities are also determined along the canal every 60 cm. These are calculated using a system of salt-tracing. For the three slope values, the dataset is available: 2, 5, and 25%. All combinations of slope and rain intensity are used except for the case with a slope of 5% and a rain intensity of 50 mm/hour has no data on the event. This result implies that there is convergence in mesh size because whatever the slope and the positive water height defined, a mesh size is adapted in the case where depth is large enough to reduce the error.

3.3.2.2. Sinusoidal molds experiment. The second test case is a laboratory test case describing a real area with several flood-ratio values. Runoff of water depths from 1 to 3 mm at high speeds is observed at the inlet. The flow is then led to furrows where the depths of the water can exceed 2 cm. The test case is representative of the hydraulic change that can be seen on an actual plot with a hydraulic leap between the runoff and the flow in the furrows. The number of furrows ranges from one to three depending on the slope option and the discharge from the inlet. The data on the test case are calculated in a steady-state so that they do not change in time. The water depths and the bottom elevation are determined using a laser scanner with a spatial resolution of 0.5 mm and the experimental water depth uncertainty is 1 mm. Concerning the velocities, they are measured using the LSPIV process, with a 5 mm spatial resolution and a 0.1 m/s experimental uncertainty. The location of the hydraulic jump and the flow in the furrows are in good agreement with the experiments. Indeed, the maximum water depths, as well as the location of the hydraulic jump entrance in the furrows, are very well reproduced. However, particularly in the experiments with three furrows, the water level is a little overestimated in the furrows that are the furthest apart.

3.3.2.3. Plot experiment. The plot is 10 m long and 4 m wide. For the domain, a uniform rain is applied using a rain simulator. Following many rain events that eroded the soil, the mobility of sediments in the soil was considered marginal and the soil is assumed to be fixed for the time of interest. The soil is composed mainly of sand with a ruggedness scale that will be called spatially uniform. Measurements of the velocities were taken at 62 separate locations on the map. As in the first experimental scenario, the measuring technique is salt-tracing. For a plot experiment case, no water level calculation is required as compared to the previous test case. Some anomalies were found on the topography at the edges, which can be attributed primarily to measurement errors.

3.3.2.4. Watershed application. The model is then put to the test by measuring the effects of the experiment to observed discharges on a particular catchment from field observations. The climate is a system of mountains influenced by the Mediterranean climate, reflected in particular by heavy summer storms and an annual average of 900 mm of rainfall. Its total area is 0.86 km^2 and 58% of the mean slope. The soil consists primarily of black marls, and 68% of the surface is bare earth. The data are valid for several rainfall events at the catchment source, with a time stage of 60 s. Only the rainfall is measured every 60 s for each case. The flow rate is measured over a specified fixed section by measuring the water depth by laser. The basin size and steep slopes make the hydrological response fast. The choice of these events is motivated by the fact that the initial state of the soil is very different from one to another, but also because they are the two events that have had the greatest impact on the transport of sediments in one year.

The model highlights well-known difficulties in solving the Shallow Water Equations for steep slopes and low depths of water. This offers an empirical criterion for determining, in a given situation, the discretization of the term slope source to prevent the public from erosion. The workflow of the proposed method is given in Figure 6.

Thus, a greater prediction of precipitation is given by the proposed rainfall model-based artificial framework. The method analyses the errors of the model and decreases the uncertainty for the fine discharge of rainfall. The infiltration regulation also comes with the runoff level estimation in the soil moisture. Formerly the model analyzes the velocity from the steep slopes to inhibit the surface water. Thus, the model estimates the predicted total precipitation.

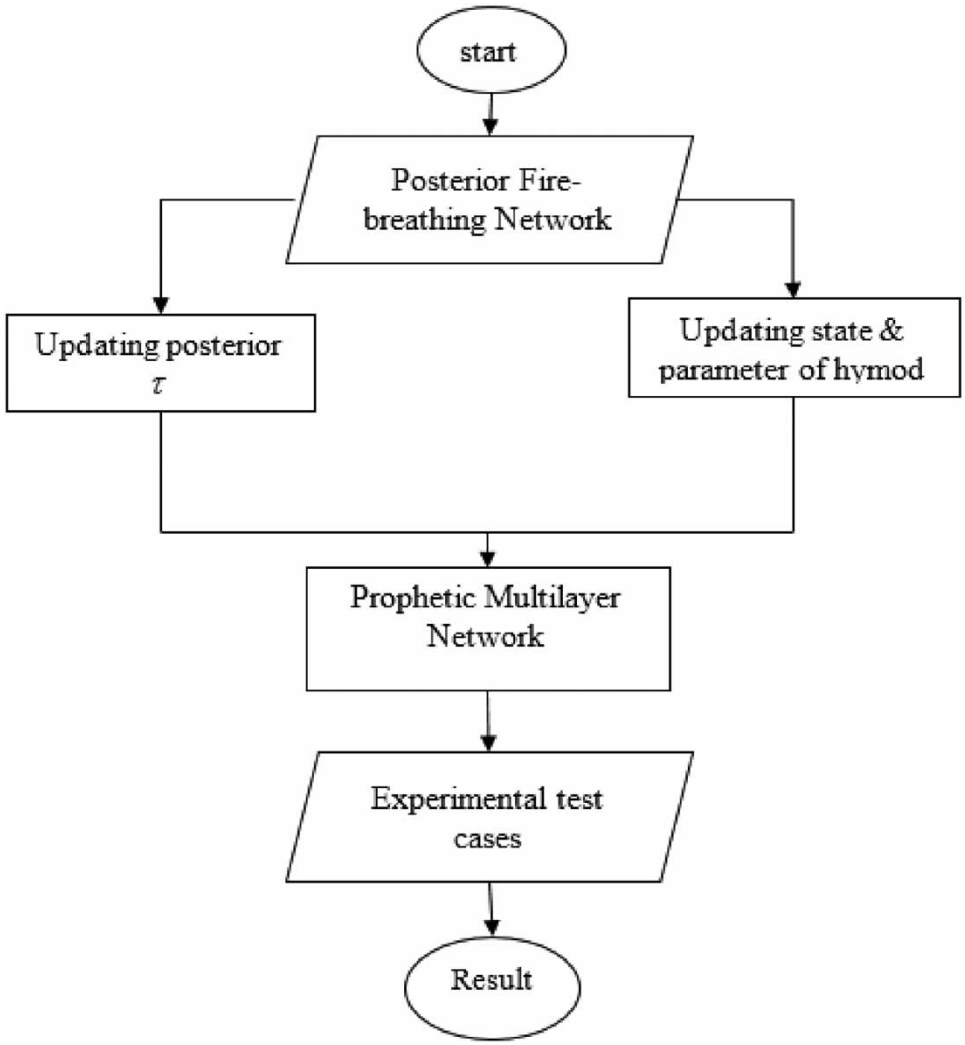

Figure 6 | Flow chart of the proposed framework.

4. RESULTS AND DISCUSSION

The section clearly explains the feasibility of the proposed method by evaluating and contrasting the experimental results obtained with traditional methods. Specification tools and the dataset description for implementation are given in the following.

4.1. Dataset description

The rainfall in India dataset is used for analyzing the rainfall–runoff. It determines the period from 1901 to 2020. Granularity is represented month-wise as January–February, March–May, June–September, October–December and its location is from the north-east. The annual rainfall record for each month is also described. The rainfall unit is expressed as millimeters in the dataset.

4.2. System specification

The proposed technique is described in Section 3, where a detailed explanation and its performance are analyzed. The proposed method is implemented in the working platform of MATLAB with the following system specification.

Platform	MATLAB 2019a
OS	Windows 8
Processor	Intel core i5
RAM	8 GB RAM

4.3. Simulation results

Simulation analysis of the proposed rainfall prognostic artificial model framework is presented in the section. Simulated results show the model errors, infiltration rate with runoff level, and steep slopes of the velocity field using the posterior fire-breathing network, prophetic multilayer network, and well-ordered selective genetic algorithm. The several parameters are taken to estimate the performance of the proposed framework such as precision probabilistic discharge with and without error, infiltration rate of soil, velocity of water in different bend areas. Those parameters are considered to estimate accurate day-by-day prediction as well as month-wise prediction. The infiltration rate affects the runoff prediction so that the soil infiltration rate and discharge is estimated.

Figure 7 states the rainfall flow region between the years 1900 and 2020. The proposed rainfall prognostic model-based artificial framework predicts the rainfall in the year range from 1900 to 2020.

From the year 1900 to 1920, the average predicted rainfall level gradually increased from 1,025 to 1,500 mm. This is the highest rainfall level marked to date in the dataset. Then the rainfall value is 1,400 mm predicted and gets constant for several years including 1935, 1960, and 1990. The marked average rainfall value for a couple of years is less compared to the early stage. The predicted rainfall value in the prognostic network of the proposed model is approximately 915–930 mm.

Figure 8 states the month-wise rainfall–runoff, i.e., January, February, March, April, May, June, July, August, September, October, November, and December. Each month's rainfall level in India between the years 1900 and 2020 is represented in Figure 8. The predicted rainfall and measured rainfall are given in Table 1.

Figure 9 compares the actual posterior to a specific time step, together with the Gamma approximation obtained by iteration: the iteration starts from the prior for that time step and iteratively moves the posterior approximation closer to the actual posterior. Thus, it obtains a good posterior approximation with the posterior fire-breathing network because the precision of the prior density moves towards posterior approximation. Also, the gamma density from the prior network reaches a maximum of 15% probability precision. Due the presence of gamma approximation, the network is tuned for new values.

Figure 10 shows the evolution of its estimated posterior, with the posterior fire-breathing network which is updated recursively with each new measurement. The posterior τ fluctuates through time in response to the mismatch between predicted and observed discharge: the value τ increases during periods of small error and decreases during periods of large error. The

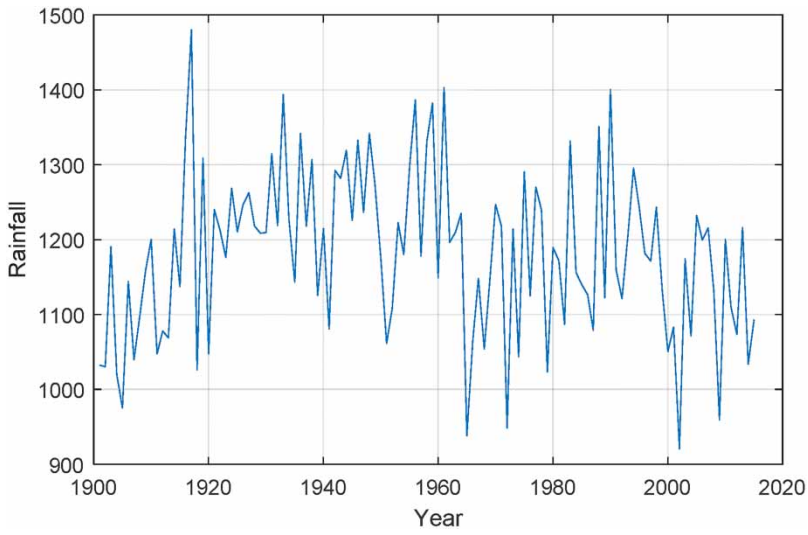

Figure 7 | The rainfall region.

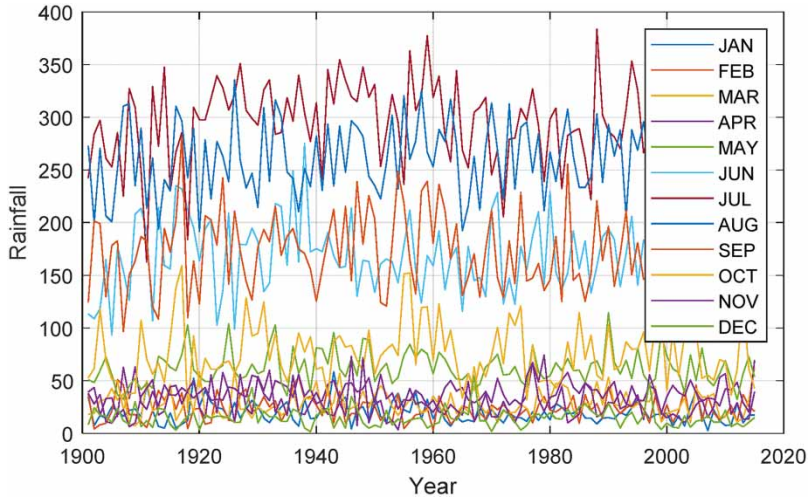

Figure 8 | Month-wise rainfall–runoff in India.

Table 1 | Predicted rainfall vs. measured rainfall

Year	Predicted rainfall (mm)	Measured rainfall (mm)
1988	375	376.3
2002	950.2	936.4
2005	215	212.2
2009	975.9	974.9
2019	1,285	1,284
2020	1,170.5	1,177

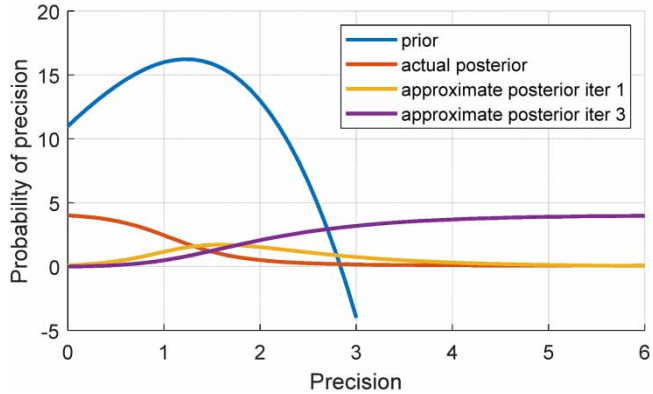

Figure 9 | Approximating the actual posterior of τ, a Gamma density by iterating, starting from the prior.

behavior corresponds to an automatic EnKF deflation to ensure the discharge of the rainfall without the model errors in predictive and observation distributions. Figure 11 shows without model error the predictive distribution is already quite good, and the predictive distribution with model error becomes too wide when compared to the likelihood. The results in a shift of the posterior τ find smaller model noise. In the model, without error, the probability is one.

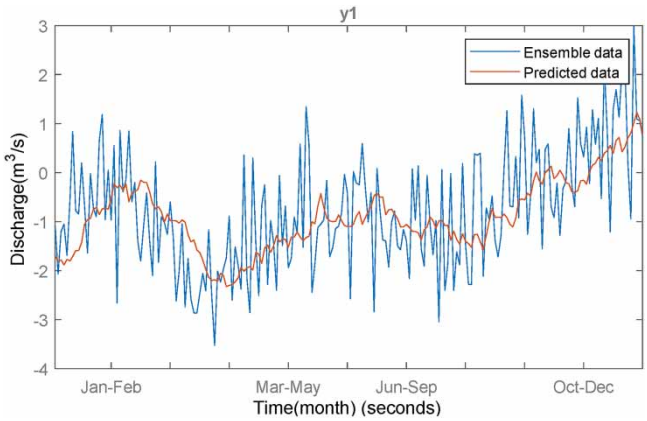

Figure 10 | One-day-ahead probabilistic discharge prediction.

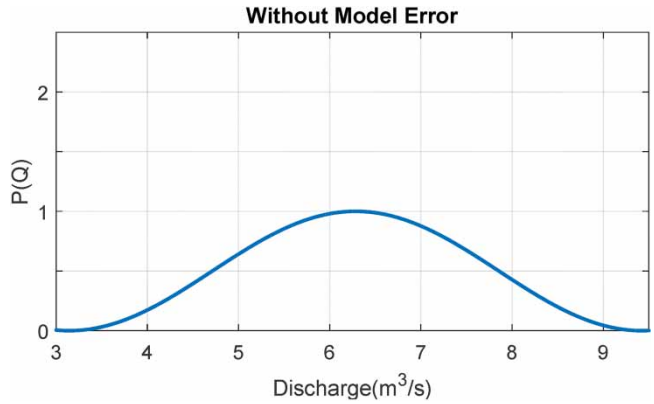

Figure 11 | Without model error.

Figure 12 shows with model error in the rainfall–runoff region. Adding model errors widens the predictive distribution, and the posterior of τ shifts further, increasing model errors and overlapping with the likelihood function. The posterior of shifts in the proposed method maintains the formal wave form from discharge 0 to 10 m^3/s. The posterior shift increases to maximum at 6.5 m^3/s. In the model with error, the obtained probability is 2 because of the EnKF that filters the history of states rather than the previous state.

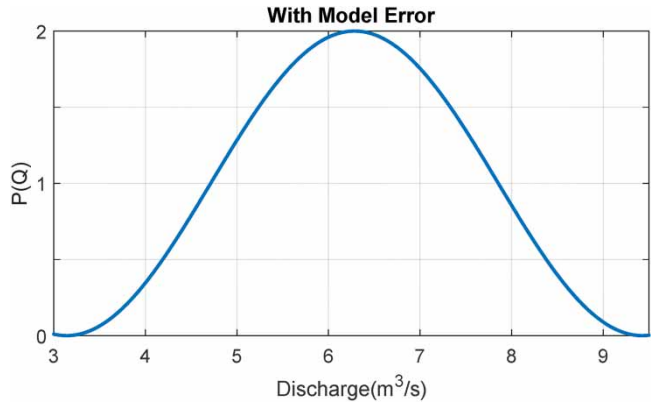

Figure 12 | With model error.

Figure 13 shows the infiltration rate and runoff level of the proposed network; the larger the rainfall intensity of 1.5 cm/hour, the higher the initial and steady infiltration rates from 6 to 8, and the cumulative infiltration increased faster with time and the runoff level is from 0 to 0.5 cm. Infiltration rate increased due to the layers present in the proposed prophetic multilayer network which estimates infiltration rate using the URMOD which does not separate the catchment for estimation.

Figure 14 shows the measured velocity in terms of position, velocity, and acceleration by the well-ordered selective genetic algorithm. The genetic algorithm measures the velocity with bend areas thus avoiding the low depth steep slopes in the hill. The highest velocities are better represented by the model by 0.45 m/s. The highest velocity was obtained in the proposed method by using well order selective genetic algorithm for unstructured mesh which deals with arbitrary complex geometries.

4.4. Comparison graph

There are three statistical metrics (root mean square error (RMSE), performance coefficient, coefficient of determination) taken to evaluate the performance of the proposed method that allows term-by-term comparison to identify the actual difference between the estimated and measured value to provide information on a model's short-term performance.

To determine the superiority of the proposed method a comparison is performed with various recent hybrid AI framework methods for runoff regions, i.e., hydro-climatic, hydro-climatic + NDVI, hydro-climatic + illumination condition (IC) [24],

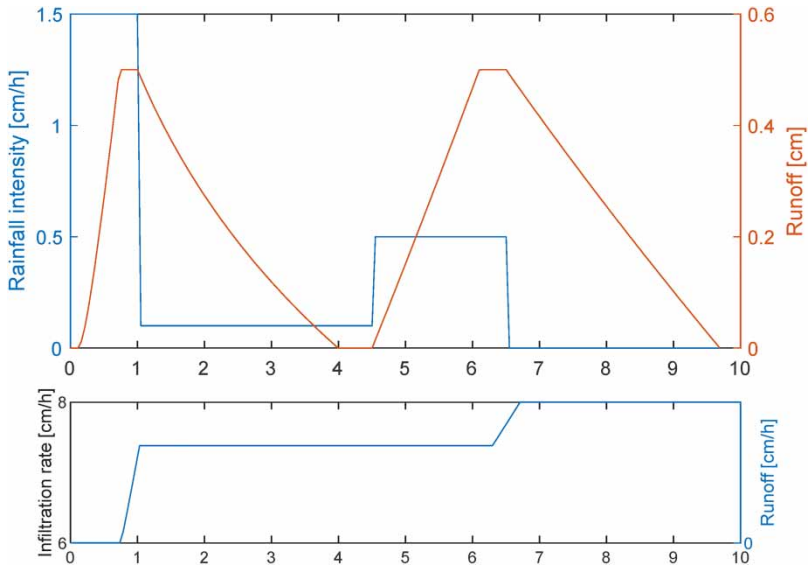

Figure 13 | Infiltration rate in runoff.

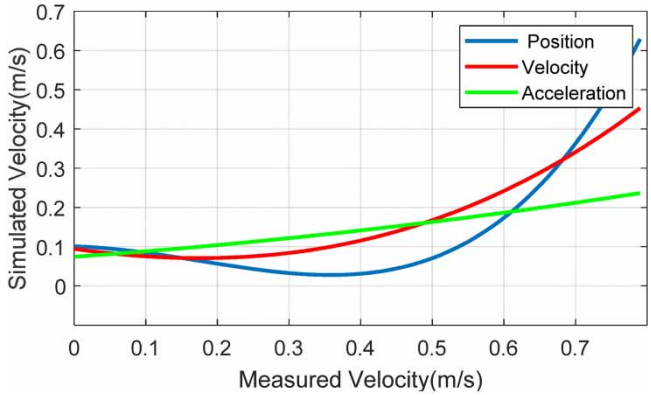

Figure 14 | Simulated velocity.

hydro-climatic + NDVI + IC for the Calliope River catchment [24] and the Haughton River Catchment [24]. The Haughton River Catchment is utilized in this study because it is one of the largest wetland rivers (316 km^2 or 122 square miles).

4.4.1. Root mean square error

RMSE is the standard deviation of the residuals (prediction errors). RMSE measures the data points from the regression line and the residuals which grouped the data in the most relevant axis.

$$RMSE = \sqrt{\overline{(f-o)^2}} \tag{25}$$

Figure 15 shows the comparison graph for RMSE in the Calliope River Catchment with runoff regions, where for the hydro-climatic method, the RMSE value is 28 mm and for the NDVI method, the obtained value is 26 mm, for the IC method, it is slighter than the NDVI + IC method and the measured value of RMSE is 23 mm. From these observations, it is clear that the proposed rainfall prognostic model-based artificial framework achieves less RMSE of 20 mm than the compared methods in the Calliope River Catchment region.

Figure 16 shows the comparison graph for RMSE in the Haughton River Catchment with runoff regions in which for the hydro-climatic method, the RMSE value is 10 mm and for the NDVI method, the obtained value is 9 mm, while for the IC and NDVI + IC method, an identical value of 8 mm was obtained. The suggested artificial system based on the rainfall prognostic model has a root average square of 6 mm lower than the comparable methods in the catchment area in the Haughton River.

4.4.2. Performance coefficient

The coefficient of performance or COP (sometimes CP or CoP) of a heat pump, refrigerator, or air conditioning system is a ratio of useful heating or cooling provided to work required. COP equates to lower operating costs and it exceeds 1 to pump

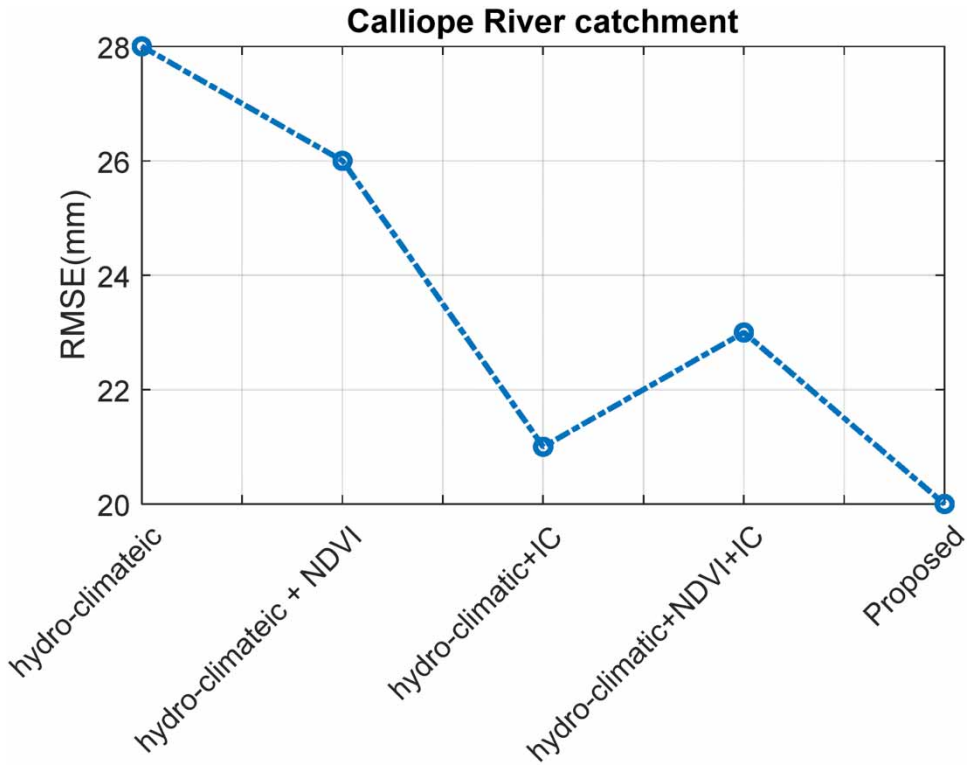

Figure 15 | Comparison graph for RMSE in the Calliope River Catchment.

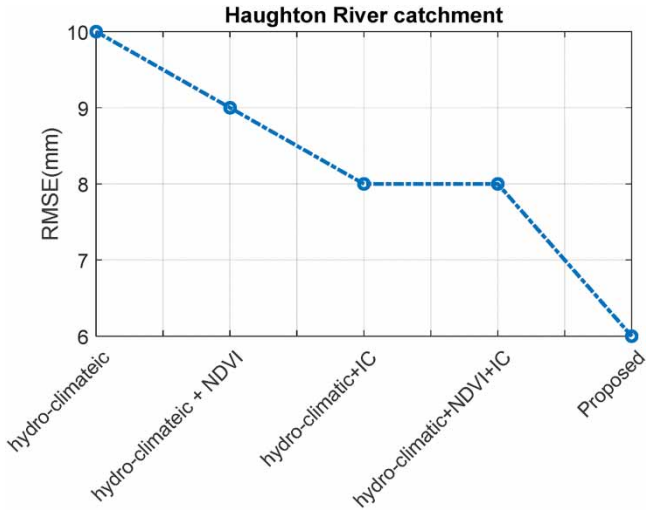

Figure 16 | Comparison graph for RMSE in the Haughton River Catchment.

additional heat from a heat source to the point where heat is required.

$$COP = \frac{Q}{W} \tag{26}$$

where Q is the heat supplied to or removed from the reservoir; W is the work done by the heat pump.

Figure 17 shows the comparison graph for performance coefficient in the Calliope River Catchment with runoff regions, where for the hydro-climatic method, the obtained value of performance coefficient is 0.81 mm and for the NDVI method, the obtained value is 0.85 mm, then for the IC method the coefficient value is 0.91 and for the NDVI + IC method, the obtained value is 0.87 mm. It can be found from these experiments that the suggested artificial system based on the rainfall prognostic model achieves a coefficient of performance greater than the compared approaches in the Haughton River catchment area by 0.93.

Figure 18 shows the comparison graph for performance coefficient in the Haughton River Catchment with runoff regions, where for the hydro-climatic method, the attained value of performance coefficient is 0.77 mm and for the NDVI method, the value obtained is 0.82 mm, then for the IC method the coefficient value is 0.85 and for the NDVI + IC method, the obtained

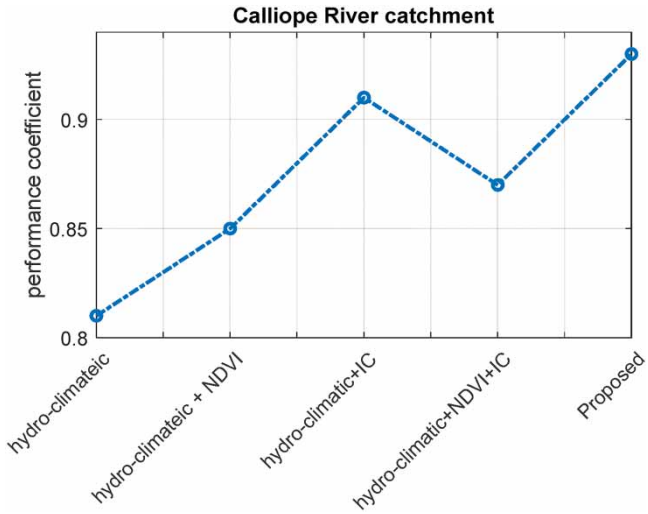

Figure 17 | Comparison graph for performance coefficient in the Calliope River Catchment.

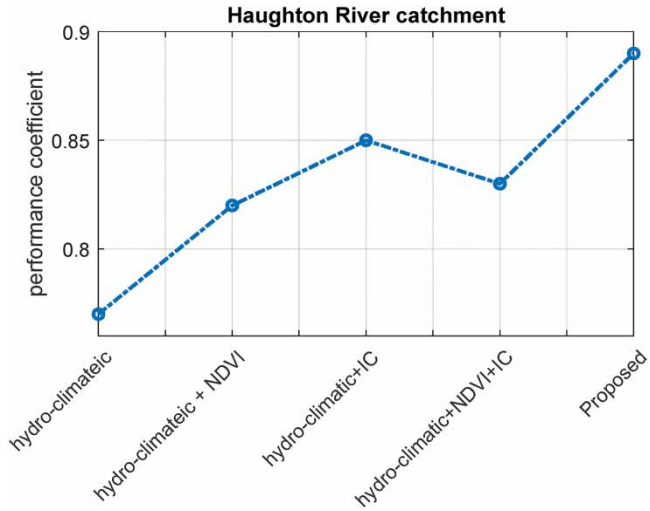

Figure 18 | Comparison graph for performance coefficient in the Haughton River Catchment.

value is 0.83 mm. These results indicate that the suggested artificial system based on rainfall prognostic is 0.89 higher than the comparable approaches in the Haughton River catchment area.

4.4.3. Coefficient of determination

The determination coefficient, in statistics, is R^2 (or r^2), a measure that evaluates a model's ability to predict or explain an outcome in the linear regression setting. R^2 displays the proportion of variance in the dependent variable and the predictor variable, explained by linear regression.

$$R^2 = \frac{MSS}{TSS} = \frac{(TSS - RSS)}{TSS} \tag{27}$$

where RSS is the sum of squares of residuals and TSS is the total sum of squares.

Figure 19 shows the comparison graph for coefficient determination in the Calliope River Catchment with runoff regions, where for the hydro-climatic method, the determination of coefficient is 0.81 mm and for the NDVI method, the obtained value is 0.89 mm; then, in the IC method, the coefficient value is 0.91 and further for the NDVI + IC, the determination

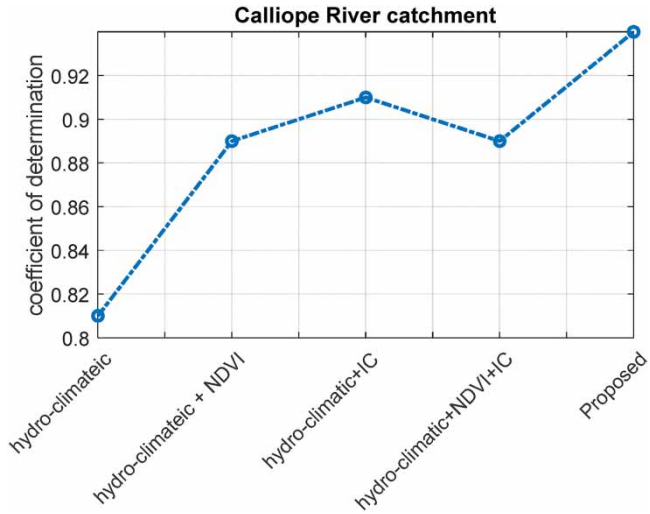

Figure 19 | Comparison graph for coefficient determination in the Calliope River Catchment.

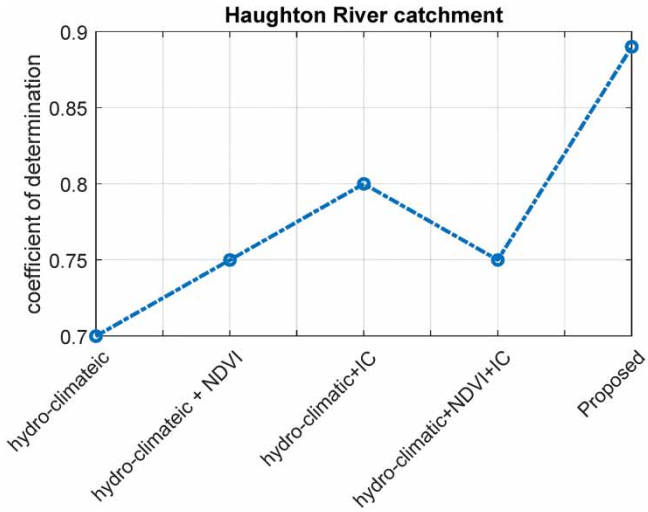

Figure 20 | Comparison graph for coefficient determination in the Haughton River Catchment.

value is 0.89 mm. These observations demonstrated a more obvious performance coefficient of 0.94 mm compared to the comparative methods in the river catchment area in the proposed framework.

Figure 20 shows the comparison graph for coefficient determination in the Haughton River Catchment with runoff regions, where for the hydro-climatic method, the determination of coefficient is 0.7 mm and for the NDVI method, the obtained value is 0.75 mm, then, for the IC method, the coefficient value is 0.8 and further for the NDVI + IC method, the determination value is 0.75 mm. From these interpretations, it is clear that the proposed framework achieves a greater performance coefficient of 0.89 than the compared methods in the Haughton River Catchment region.

Thus it is inferred from the above findings that the proposed approach provides a better prediction in the range of the rainfall–runoff model with the infiltration rate and the runoff level. Accordingly, it concludes that the rainfall prognostic artificial model framework effectively predicts model errors, runoff levels with infiltration rate, and velocity calculation the assistance for the prediction of rainfall in various temperatures.

5. CONCLUSION

Despite the progress made in recent years, modeling hydrological reactions to rainfall prediction remains a complex task in runoff modeling. Thus, the presented paper effectively introduced a rainfall prognostic artificial model framework for the prediction of rainfall. The framework applied a posterior fire-breathing network to estimate model errors with random noise to reduce the uncertainty. Further, to regulate the infiltration rate, the system suggested a prophetic multilayer network that analyses the runoff levels with the soil moisture in urban and rural areas. In addition, to tackle the numerical challenge, the model integrates a well-ordered selective genetic algorithm to forecast different bend zones at low water depths on steep slopes. As a result of the rainfall model, everyday rainfall can be accurately anticipated to avoid environmental problems. Experimental results reveal that the framework surpasses existing runoff models with a better infiltration rate of 1.5 cm/hour, runoff level of 0.05 cm, achieved a velocity of 0.45 mm, and has the highest prediction range of 12–20 mm with subjective results.

6. FUTURE PERSPECTIVES

- In the future, uncertainty of rainfall–runoff model error shall be estimated using a screening method to extract values from stage measurements and conversions of stage ratings. Also, the noise parameter of the model errors may have been calculated by normality and deviation methods to obtain more accurate estimation.
- Unscented EnKF-based measurement of the nonlinear function may be developed to measure the values even from the nonlinear dynamic model.
- Further rainfall–runoff estimation may be developed with some deep learning methods with optimization technique and run off levels of catchment estimated by using the dataset values to evaluate the performance of the method.

- In the future, steep slope run off velocity has to be calculated with various velocity measurement components as well as decision algorithms to find the probability of run off velocity for better prediction results.

DATA AVAILABILITY STATEMENT

All relevant data are available from https://www.kaggle.com/datasets/rajanand/rainfall-in-india.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Asadi, H., Shahedi, K., Jarihani, B. & Sidle, R. C. 2019 Rainfall-runoff modelling using hydrological connectivity index and artificial neural network approach. *Water* **11** (2), 212.

Chadalawada, J., Herath, H. M. V. V. & Babovic, V. 2020 Hydrologically informed machine learning for rainfall-Runoff modeling: a genetic programming-based toolkit for automatic model induction. *Water Resources Research* **56** (4), e2019WR026933.

Chen, X. Y. & Chau, K-W. 2019 Uncertainty analysis on hybrid double feedforward neural network model for sediment load estimation with LUBE method. *Water Resources Management* **33** (10), 3563–3577.

Chlumecký, M., Buchtele, J. & Richta, K. 2017 Application of random number generators in genetic algorithms to improve rainfall-runoff modelling. *Journal of Hydrology* **553**, 350–355.

Dakhlaoui, H., Ruelland, D., Tramblay, Y. & Bargaoui, Z. 2017 Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia. *Journal of Hydrology* **550**, 201–217.

Del Giudice, G. & Padulano, R. 2016 Sensitivity analysis and calibration of a rainfall-runoff model with the combined use of EPA-SWMM and genetic algorithm. *ActaGeophysica* **64** (5), 1755–1778.

Fotovatikhah, F., Herrera, M., Shamshirband, S., Chau, K. W., Faizollahzadeh Ardabili, S. & Piran, M. J. 2018 Survey of computational intelligence as basis to big flood management: challenges, research directions and future work. *Engineering Applications of Computational Fluid Mechanics* **12** (1), 411–437.

Fowler, K. J., Peel, M. C., Western, A. W., Zhang, L. & Peterson, T. J. 2016 Simulating runoff under changing climatic conditions: revisiting an apparent deficiency of conceptual rainfall-runoff models. *Water Resources Research* **52** (3), 1820–1846.

Fraga, I., Cea, L. & Puertas, J. 2019 Effect of rainfall uncertainty on the performance of physically based rainfall–runoff models. *Hydrological Processes* **33** (1), 160–173.

Garcia, F., Folton, N. & Oudin, L. 2017 Which objective function to calibrate rainfall–runoff models for low-flow index simulations? *Hydrological Sciences Journal* **62** (7), 1149–1166.

Guo, D., Westra, S. & Maier, H. R. 2017 Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models. *Water Resources Research* **53** (1), 435–454.

Jaiswal, R. K., Ali, S. & Bharti, B. 2020 Comparative evaluation of conceptual and physical rainfall–runoff models. *Applied Water Science* **10** (1), 1–14.

Knoben, W. J., Freer, J. E., Fowler, K. J., Peel, M. C. & Woods, R. A. 2019 *Modular Assessment of Rainfall-Runoff Models Toolbox (MARRMoT) v1.2: an Open-Source, Extendable Framework Providing Implementations of 46 Conceptual Hydrologic Models as Continuous State-Space Formulations*.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K. & Herrnegger, M. 2018 Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences* **22** (11), 6005–6022.

Kumar, S., Roshni, T. & Himayoun, D. 2019 A comparison of emotional neural network (ENN) and artificial neural network (ANN) approach for rainfall-runoff modelling. *Civil Engineering Journal* **5** (10), 2120–2130.

Ledesma, J. L. & Futter, M. N. 2017 Gridded climate data products are an alternative to instrumental measurements as inputs to rainfall–runoff models. *Hydrological Processes* **31** (18), 3283–3293.

Mosavi, A., Ozturk, P. & Chau, K. W. 2018 Flood prediction using machine learning models: literature review. *Water* **10** (11), 1536.

Nourani, V. 2017 An emotional ANN (EANN) approach to modeling rainfall-runoff process. *Journal of Hydrology* **544**, 267–277.

Nourani, V., DavanlouTajbakhsh, A., Molajou, A. & Gokcekus, H. 2019 Hybrid wavelet-M5 model tree for rainfall-runoff modeling. *Journal of Hydrologic Engineering* **24** (5), 04019012.

Reshma, T., Reddy, K. V., Pratap, D. & Agilan, V. 2018 Parameters optimization using Fuzzy rule based multi-objective genetic algorithm for an event based rainfall-runoff model. *Water Resources Management* **32** (4), 1501–1516.

Sitterson, J., Knightes, C., Parmar, R., Wolfe, K., Avant, B. & Muche, M. 2018 *An Overview of Rainfall-Runoff Model Types*.

Song, J. H., Her, Y., Park, J. & Kang, M. S. 2019 Exploring parsimonious daily rainfall-runoff model structure using the hyperbolic tangent function and Tank model. *Journal of Hydrology* **574**, 574–587.

Taormina, R. & Chau, K-W. 2015 ANN-based interval forecasting of streamflow discharges using the method and MOFIPS. *Engineering Applications of Artificial Intelligence* **45**, 429–440.

Üneş, F., Keskin, L. & Demirci, M. 2019 Artificial neural networks method for prediction of rainfall-runoff relation: regional practice. *Natural and Engineering Sciences* **4** (3), 220–230.

Vidyarthi, V. K. & Chourasiya, S. 2020 Particle swarm optimization for training artificial neural network-based rainfall–runoff model, case study: Jardine River Basin. *Micro-Electronics and Telecommunication Engineering* **8**, 641–647.

Weeser, B., Jacobs, S., Kraft, P., Rufino, M. C. & Breuer, L. 2019 Rainfall-Runoff modeling using crowdsourced water level data. *Water Resources Research* **55** (12), 10856–10871.

Wu, C. L. & Chau, K-W. 2013 Prediction of rainfall time series using modular soft computing methods. *Engineering Applications of Artificial Intelligence* **26** (3), 997–1007.

Xiang, Z., Yan, J. & Demir, I. 2020 A rainfall-runoff model with LSTM-based sequence-to-sequence learning. *Water Resources Research* **56** (1), e2019WR025326.

First received 18 January 2022; accepted in revised form 8 August 2022. Available online 5 September 2022