Investigation of satellite precipitation product driven rainfall-runoff model using deep learning approaches in two different catchments of India

Pavan Kumar Yeditha [iD][a], Maheswaran Rathinasamy [iD][b,*], Sai Sumanth Neelamsetty[a], Biswa Bhattacharya[c] and Ankit Agarwal[d,e]

[a] Department of Civil Engineering, MVGR College of Engineering, Vizianagaram, Andhra Pradesh 535005, India
[b] Department of Civil Engineering, Indian Institute of Technology Hyderabad, Kandi 502285, India
[c] Water Science and Engineering, IHE-Delft, Delft, Netherlands
[d] Department of Hydrology, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand 247667, India
[e] GFZ German Research Centre for Geosciences, Section Hydrology, Telegrafenberg, Potsdam 14473, Germany
*Corresponding author. E-mail: maheswaran27@yahoo.co.in

[iD] PKY, 0000-0002-1657-8507; MR, 0000-0002-8124-3092

ABSTRACT

Rainfall–runoff models are valuable tools for flood forecasting, management of water resources, and drought warning. With the advancement in space technology, a plethora of satellite precipitation products (SPPs) are available publicly. However, the application of the satellite data for the data-driven rainfall–runoff model is emerging and requires careful investigation. In this work, two satellite rainfall data sets, namely Global Precipitation Measurement-Integrated Multi-Satellite Retrieval Product V6 (GPM-IMERG) and Climate Hazards Group Infrared Precipitation with Station (CHIRPS), are evaluated for the development of rainfall–runoff models and the prediction of 1-day ahead streamflow. The accuracy of the data from the SPPs is compared to the India Meteorological Department (IMD)-gridded precipitation data set. Detection metrics showed that for light rainfall (1–10 mm), the probability of detection (POD) value ranges between 0.67 and 0.75 and with an increasing rainfall range, i.e., medium and heavy rainfall (10–50 mm and >50 mm), the POD values ranged from 0.24 to 0.45. These results indicate that the satellite precipitation performs satisfactorily with reference to the IMD-gridded data set. Using the daily precipitation data of nearly two decades (2000–2018) over two river basins in India's eastern part, artificial neural network, extreme learning machine (ELM), and long short-time memory (LSTM) models are developed for rainfall–runoff modelling. One-day ahead runoff prediction using the developed rainfall–runoff modelling confirmed that both the SPPs are sufficient to drive the rainfall–runoff models with a reasonable accuracy estimated using the Nash–Sutcliffe Efficiency coefficient, correlation coefficient, and the root-mean-squared error. In particular, the 1-day streamflow forecasts for the Vamsadhara river basin (VRB) using LSTM with GPM-IMERG inputs resulted in Nash-Sutcliffe Efficiency Coefficient (NSC) values of 0.68 and 0.67, while ELM models for Mahanadhi river basin (MRB) with the same input resulted in NSC values of 0.86 and 0.87, respectively, during training and validation stages. At the same time, the LSTM model with CHIRPS inputs for the VRB resulted in NSC values of 0.68 and 0.65, and the ELM model with CHIRPS inputs for the MRB resulted in NSC values of 0.89 and 0.88, respectively, in training and validation stages. These results indicated that both the SPPs could reliably be used with LSTM and ELM models for rainfall–runoff modelling and streamflow prediction. This paper highlights that deep learning models, such as ELM and LSTM, with the GPM-IMERG products can lead to a new horizon to provide flood forecasting in flood-prone catchments.

Key words: artificial neural networks (ANN), deep learning, extreme learning machines (ELM), long short time memory (LSTM), rainfall-runoff modelling, satellite rainfall products

HIGHLIGHTS

- Evaluated applicability of satellite rainfall products for rainfall - runoff modelling in Vamsadhara and Mahanadi river basins.
- GPM-IMERG and CHIRPS data could be reliably used with LSTM and ELM models for rainfall -runoff modelling and streamflow forecasting.
- Deep learning models with the CHIRPS products can lead to a new horizon to provide flood forecasting in flood-prone catchments.

1. INTRODUCTION

Rainfall–runoff modelling is vital in reservoir management, hydropower, environmental flow, water management, flood assessment, and so on. Owing to its immense importance, this topic has gathered attention worldwide, and numerous

approaches were developed for rainfall–runoff modelling. Broadly, rainfall–runoff models can be classified into the following: (i) data-driven-based methods (Khac-Tien Nguyen & Hock-Chye Chua 2012; Elsafi 2014; Nguyen *et al.* 2014; Yaghoubi *et al.* 2019; Hadid *et al.* 2020; Xiang *et al.* 2020), (ii) conceptual model-based approaches (Nash & Sutcliffe 1970; Brath & Rosso 1993; Kan *et al.* 2017; Unduche *et al.* 2018), and (iii) physical model-based methods (Vieux *et al.* 2003; Chen *et al.* 2016; Setti *et al.* 2020). Of these methods, data-driven-based models were found to be of a great value due to their accuracy and simplicity (Teegavarapu & Chandramouli 2005; Wu *et al.* 2009; Wu & Chau 2010; Orouji *et al.* 2013; Nguyen *et al.* 2014; Steyn *et al.* 2017; Ahani *et al.* 2018; Mazrooei & Sankarasubramanian 2019). Precipitation is a significant input variable of the different input parameters required for rainfall–runoff models (Mazrooei & Sankarasubramanian 2019; Guntu *et al.* 2020). For reliable modelling, precipitation plays a crucial role; hence, the quality of the precipitation used must be reliable. For many decades, ground-based data sets were used to develop forecasting models, but it is also well-known that ground-based gauged data are scarce and are limited to few locations.

In recent decades, there is a rise in remote sensing data to analyze various meteorological parameters like temperature, precipitation, humidity, and several others. Some of the recent successful applications of remote sensing data included the works of Yeditha *et al.* (2020) and Valipour & Bateni (2021). In the case of precipitation, due to an increase in various satellite-derived precipitation products and having advantages of easy availability in real-time and high temporal and spatial coverages, they are being used as alternatives for ground-based gauged data sets. Some of the commonly used satellite precipitation products (SPPs) include Tropical Rainfall Measuring Mission (TRMM); Precipitation Estimation from Remote Sensed Information using Artificial Neural Networks (PERSIANN) (Hiep *et al.* 2018); Global Precipitation Measurements (GPM); Global Precipitation Measurement-Integrated Multi-Satellite Retrievals (GPM-IMERG) (Huffman *et al.* 2020; Tang *et al.* 2020); and Climate Hazards Group Infrared Precipitation with Station data (CHIRPS) (Funk *et al.* 2015).

A multitude of studies have been reported on the utility of these SPPs for applications such as hydrological modelling, water balance (Su *et al.* 2008; Zhang *et al.* 2015; Xu *et al.* 2017; Vu *et al.* 2018; Gadhawe *et al.* 2021), drought monitoring (Lai *et al.* 2019), streamflow modelling (Munzimi *et al.* 2019; Sulugodu & Deka 2019; Yeditha *et al.* 2020), and water quality modelling (Ali & Shahbaz 2020). Indeed, a large number of studies have been carried out to understand the application of satellite precipitation and their evaluation (Janowiak *et al.* 2004; Nair *et al.* 2009; Rahman *et al.* 2009; Bitew & Gebremichael 2011; Meng *et al.* 2014; Prakash & Gairola 2014; Prakash *et al.* 2015; Tang *et al.* 2016; Li *et al.* 2018; Wei *et al.* 2018). A very few studies showed the evaluation of different SPPs to apply data-driven rainfall–runoff models.

There are a significant number of studies (Adamowski *et al.* 2012; Kişi 2008; Islam 2010; Yaseen *et al.* 2015; Couta *et al.* 2019; Ding *et al.* 2019; Ali & Shahbaz 2020; Kasi *et al.* 2020; Wagena *et al.* 2020; Yadav *et al.* 2020; Yeditha *et al.* 2020) reporting the efficiency of machine learning methods such as artificial neural networks (ANN), extreme learning machines (ELM) (Huang *et al.* 2006), and long short-term memory (LSTM) (Kratzert *et al.* 2019) models based on the rain gauge data for rainfall–runoff applications. Several studies (Huang *et al.* 2006; Atiquzzaman & Kandasamy 2016; Yaseen *et al.* 2018; Sulugodu & Deka 2019) showed that ELM models outperformed ANN, Support Vector Regression (SVR), and General Regression Neural Network (GRNN) in terms of modelling accuracy and computation time. This increase in performance is because ELM does not use gradient-based techniques, which makes it run faster and also it is not subjected to an issue like local minima, improper learning as observed in the ANN models. Furthermore, all the parameters are tuned once, and ELM does not need iterative training and also produces the same result at all times, unlike ANN (Alizamir *et al.* 2018).

Similarly, in recent years, the LSTM variant of recurrent neural networks has higher capabilities than the traditional neural network in producing robust results when dealing with long-time series information (Sit & Demir 2019; Xiang *et al.* 2020). On comparing the performance of LSTM with other traditional models like convolutional neural networks and multilayer perceptron for different hydrological applications, LSTM showed better understanding (Kratzert *et al.* 2018; Tiwari & Vaibhaw 2020). Along similar lines, Couta *et al.* (2019), Ding *et al.* (2020), and Hu *et al.* (2020) used LSTM models for streamflow forecasting. Li *et al.* (2020) and Ni *et al.* (2020) employed LSTM for rainfall–runoff modelling and streamflow and rainfall forecasting. Similarly, Zhou & Ushiama (2019) have developed LSTM models for forecasting streamflow in the Tunxi catchment in Southeast China and compared the performance of the LSTM model with SVM. The results showed superior performance of LSTM compared to SVM. Similarly, Li *et al.* (2020) developed a high-resolution rainfall–runoff model using the LSTM approach and found that the models efficiently predicted discharge and achieved robust modelling performance.

The critical analysis of the recent literature revealed that despite significant progress in new machine learning algorithms and their applications for hydrological problems, there has been little work reported on the application of SPPs for rainfall–

runoff modelling and their forecasting capabilities. With increasing reliance on SPPs in the modern age due to either poor ground-based data or unavailability, it is of utmost importance to understand their accuracy and capabilities in the domain of rainfall–runoff modelling and forecasting. Therefore, this study aims to build machine learning-based rainfall–runoff models driven by different SPPs. For this purpose, deep learning models were developed for two river basins using state-of-the-art methods such as ELM and LSTM models. We have selected these models since many studies have shown them superior to other contemporary models (Yaseen *et al.* 2015; Le *et al.* 2019; Sahoo *et al.* 2019; Sulugodu & Deka 2019; Li *et al.* 2020; Ni *et al.* 2020). As a part of the study, we also develop models based on the traditional methods such as feedforward neural networks (FFNN) and multiple linear regression (MLR) for benchmarking the results.

2. STUDY AREA AND DATA USED

2.1. Study area description

We considered two major river basins, namely Vamsadhara river basin (VRB) and Mahanadi river basin (MRB), in India. Figure 1 shows the geographical location of the VRB and the MRB in India. Both the river basins are prone to frequent floods leading to loss of lives and property. During monsoon seasons, >150 villages and thousands of acres face flood threats. Flash floods were observed in 1980 and 2013, leading to a severe loss of lives and property in the VRB. The MRB is considered to have recorded one of the most notorious floods in India. Some of the recent flood events recorded are the floods during the monsoon months in 2008, where 16 people lost lives during the river bank breach. Flash floods during the year 2011 in the monsoon season led to the flooding of 25 villages upstream of the river. The Central Water Commission has developed a rainfall–runoff model using MIKE tools for flood forecasting applications; however, the model is data-intensive and requires higher computational effort. Hence, there is an urgent need for a straightforward and dependable model for these river basins with high reliability.

2.1.1. Vamsadhara river basin

Vamsadhara river is an east-flowing river originating from Lanjigarh in the Kalahandi district of Orissa at 18°15′ to 19°55′ North latitudes and 83°20′ to 84°20′ East longitudes that joins the Bay of Bengal at Kalingapatnam in Andhra Pradesh. The total catchment area is 10,830 km^2, out of which 8,015 km^2 is in Odisha and the remaining 2,815 km^2 is in Andhra Pradesh that runs for about 254 km between the Mahanadi and Godavari rivers. The basin receives 1,400 mm of annual average rainfall, and its average discharge is 125 m^3/s. Climate is characterized by hot summers and mild winters. The maximum temperature in plains during May rises to 43 °C and goes down to 12 °C in December–January. The river basin is influenced by the southwest monsoon and occasional cyclones due to depressions in the Bay of Bengal.

2.1.2. Mahanadi river basin

Mahanadi river is an east-flowing river originating from Sihawa in the Dhamtari district of Chhattisgarh at 85°40′ to 86° 45′ East longitudes and 19°40′ to 20°35′ North latitudes. It flows through a distance of 900 km from Chhattisgarh to Odisha. Mahanadi river has a total catchment area of about 139,681 km^2 and is the eighth largest basin in India. The average discharge is 2,119 m^3/s with a maximum discharge of 57,000 m^3/s, almost as the Ganges during monsoon, and has an annual average rainfall of 1,291 mm. Of the tributaries, the upstream catchment of the Basantpur gauging station (having a catchment of 59,000 m^2) is considered in this study as it is flood-prone and contributes to maximum flow. The climate in the region of Mahanadi is sub-tropical. The temperature in summer and winter is around 29 and 21 °C, respectively. The southwest monsoon starts in the middle of June and continues until the end of September.

2.2. Data used

This study has assessed the utility of two SPPs, namely GPM-IMERG V6 and CHIRPS. For comparison, we have used the gauge-based India Meteorological Department (IMD)-gridded data set. A brief description of the data set is provided here.

2.2.1. GPM-IMERG data

The GPM project is developed by the combined effort of the Japan Aerospace Exploration Agency and the National Aeronautics and Space Administration of the United States. GPM is a widely used precipitation product over the globe, and it is also an updated precipitation product of the TRMM. GPM has two primary sensors: GPM Microwave Imager, which means precipitation size, type, and intensity; and Dual-frequency Precipitation Radar, which gauges the storm's internal structure of clouds (Kim *et al.* 2017). GPM Microwave Imager has more frequency channels than
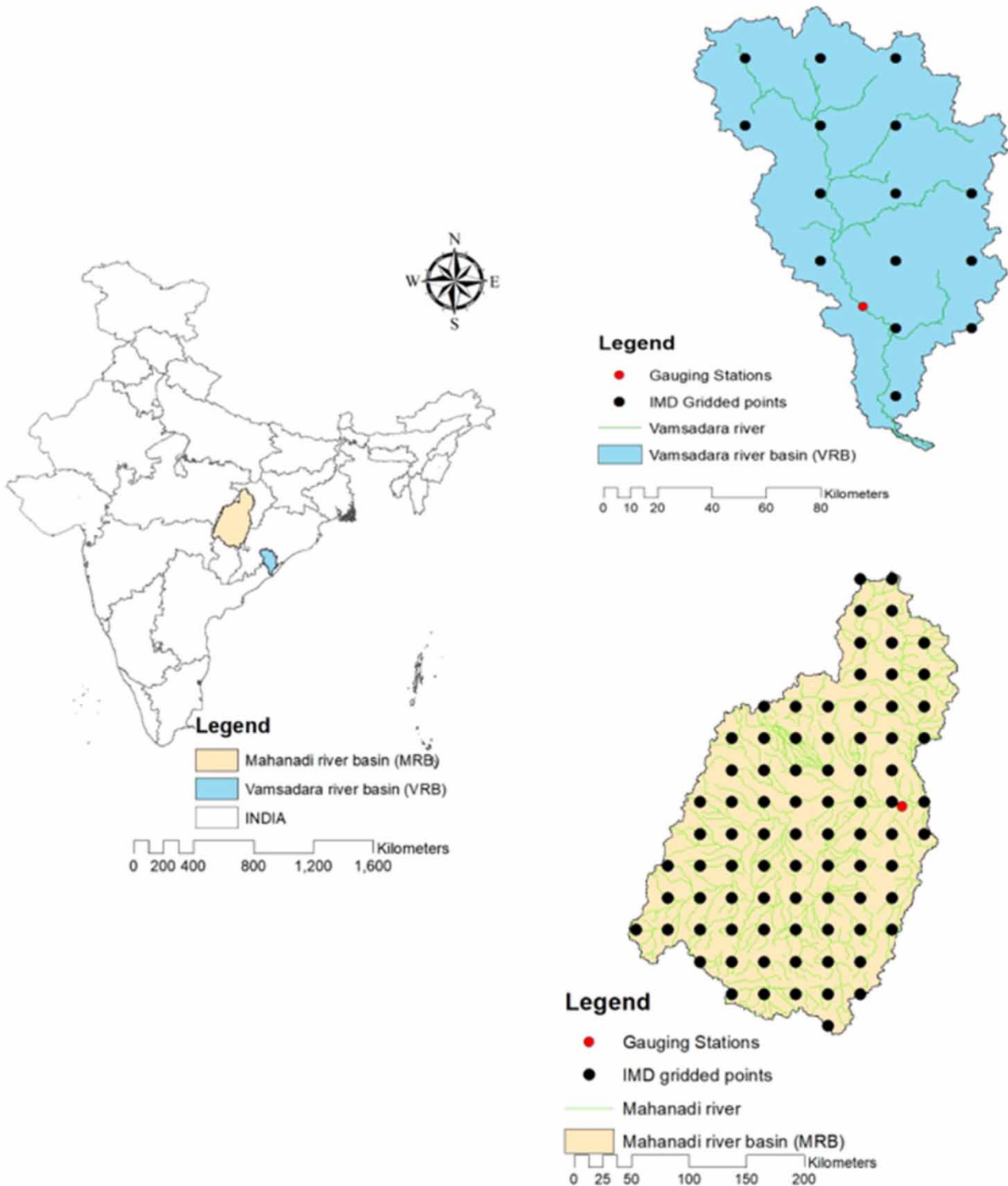
Figure 1 | Geographical location of the VRB and the MRB. The red and black colour dots indicate the gauging stations and IMG-gridded stations available at the basin.

TRMM, and the precipitation radar is more advanced (Tan & Santo 2018). More importantly, the spatial coverage (60° S–60°N) and the spatiotemporal resolution (30 min and (0.1°×0.1°)) have been improved compared to the previous SPPs.

The GPM team has developed GPM-IMERG precipitation products using the IMERG algorithm, which combines three SPP features, including PERSIANN-CDR, CMORPH, and TRMM (Prakash 2019). Three types of IMERG products are made available: early run, late run, and final run with a latency period of 4 h, 14 h, and 3.5 months. The first two products are useful in a real-time application, whereas the latter can be used for water balance studies. All these products are available at 0.1°×0.1° spatial resolution over the fully global domain.

Presently, IMERG is at its Ver 06 stage (https://gpm.nasa.gov/missions/two-decades-imerg-resources). The latest update is Ver 06 IMERG that allows for fusing the data collected during the TRMM era (2000–2014) with the precipitation estimates compiled during the GPM era. Owing to this, IMERG is now available from June 2000 (https://pmm.nasa.gov/data access/downloads/gpm). The 'Final run' of IMERG combines the GPCC monitoring product, the V8 Full Data Analysis, for most of the time (currently 1998–2019). Here, we have used the GPM-IMERG V6 final run precipitation product.

2.2.2. CHIRPS data

CHRPS is a satellite-based precipitation product that provides a quasi-global rainfall data set at about three decades. The data set is available at 0.05° with the integration of station data from 1981 to near the current date. For extracting CHIRPS data, IMD-gridded point coordinates over the catchment areas of Vamsadhara and Mahanadi were considered. They were downloaded from the website of the Climate hazards Center, UC Santa Barbara (https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_daily/netcdf/p25/).

The downloaded data sets are obtained as NetCDF files and are extracted using a script in MATLAB 2019a.

2.2.3. IMD data

We used IMD-gridded daily data (IMD 4), having a spatial resolution of 0.25°×0.25° and covering the entire Indian landmass, as ground validation. The data sets were developed based on the rainfall records from 6,955 rain gauges spread across the country (Pai et al. 2014). These rain gauges were partly maintained by the IMD, Agromet Department and mainly by the state government. Using the Shepard Interpolation method, the rain gauge collected rainfall has been converted into the gridded data set. For detailed information, see Pai et al. (2014). The data set is available from 1901 to date; however, the data set for 2000–2018 is considered for this study. The data sets can be obtained from the IMD, Pune (http://www.imdpune.gov.in/).

2.2.4. Streamflow data

The daily streamflow data sets used to develop forecasting models were obtained from the Central Water Commission and were downloaded from https://indiawris.gov.in/wris/. The gauging stations chosen for the VRB and the MRB are Kashinagar and Basantpur stations, respectively (Figure 1). This study used the data from 2000 to 2018 to calibrate and validate the models.

3. METHODOLOGY AND METHODS

We have shown a schematic of the proposed methodology in Figure 2. We used two different SPPs (GPM-IMERG and CHIRPS) for VRB and MRB catchments to understand the relationship between rainfall–runoff and predict 1 day ahead of streamflow. A sensitive analysis was carried out to understand the accuracy by comparing SPP with IMD data sets based on detection and error components in the SPP. Furthermore, correlation analysis was carried out for the three data sets (GPM-IMERG, CHIRPS, and IMD) with runoff to understand their relationship and determine the lagged variables used as input models. Linear regression (MLR) and machine learning models (FFNN, ELM, and LSTM) were developed for 1-day ahead runoff prediction using the selected variables. After successful calibration and validation of the models, the accuracy of the resulted runoff from the models was checked using evaluation measures like root-mean-squared error (RMSE), Nash–Sutcliffe Efficiency (NSE), and correlation. Based on the values of the evaluation measures, the best models were chosen, and their accuracy for the prediction of peak flow was tested. A few case studies of the peak flow analysis are carried out using the best models, and its accuracy is analyzed using evaluation measures. The best-performing model and the best-performing SPP are determined for each considered catchment.
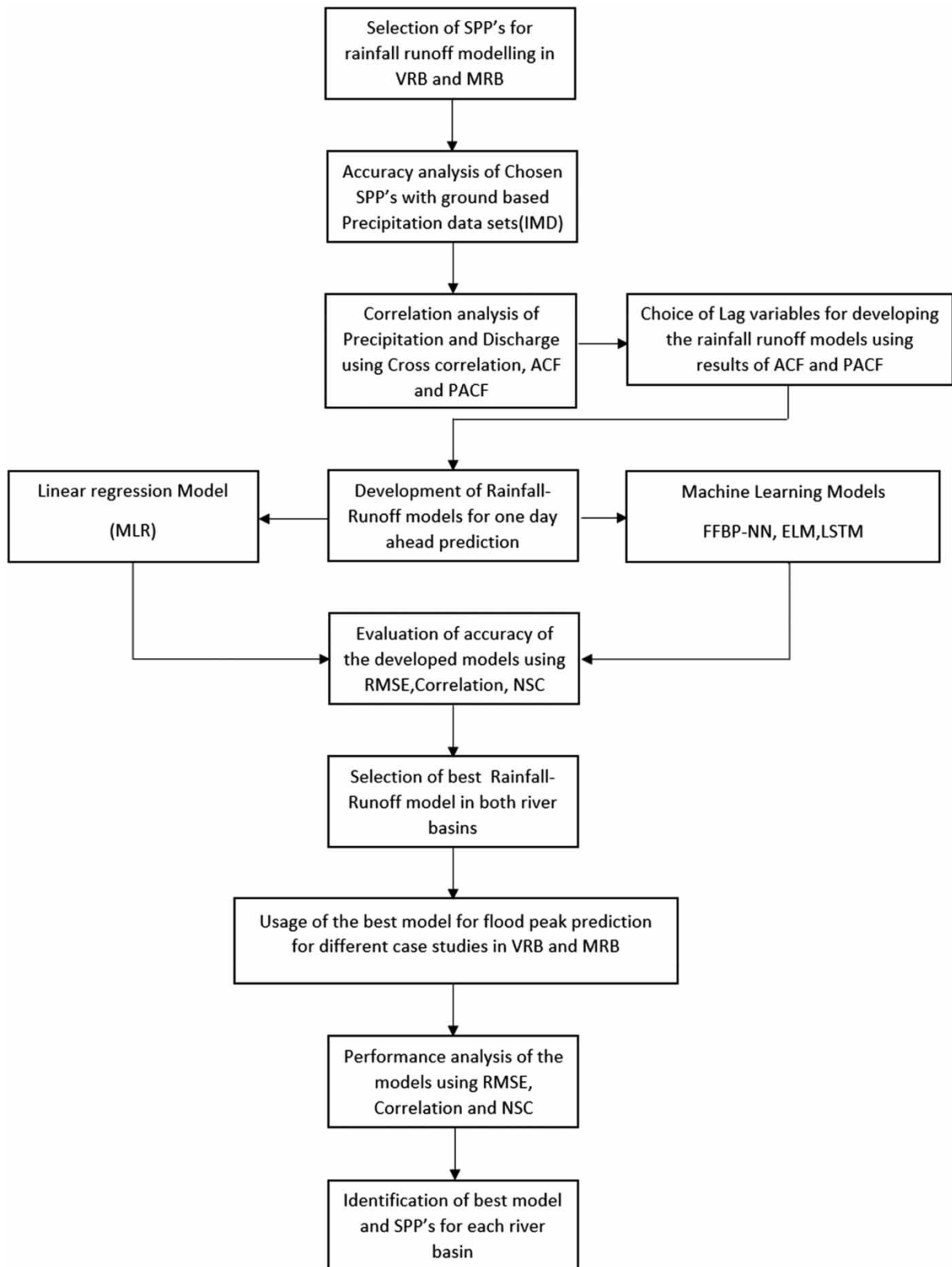
Figure 2 | Schematic representation of the methodology adopted for rainfall–runoff modelling.

3.1. Categorical analysis of satellite precipitation

To assess the SPP's ability to capture the occurrence of rainfall, we have used specific statistical measures known as detection metrics. In this study, GPM-IMERG and CHIRPS data sets are compared with the IMD data sets to evaluate the chosen satellite data sets for 16 years from 2000 to 2015. The following metrics were used in this study.

The probability of detection (POD) determines the number of events correctly detected by the satellite data sets concerning IMD data sets shown mathematically in the following equation:

$$POD = \frac{Hit}{Hit + Miss} \tag{1}$$

Hit represents the number of days when rainfall events are recorded at the rain gauge and the satellite. *Miss* represents the number of days where rainfall event is not recorded in the satellite. Still, its presence is shown in the IMD data set. The POD value ranges from 0 to 1. If the obtained value of POD tends towards 0, the data sets have less accuracy, whereas it represents high accuracy if it tends towards 1.

The false alarm ratio (FAR) determines the fraction of satellite values where the satellite data have reported precipitation, but no precipitation was indicated in IMD data sets. The FAR is mathematically represented in the following equation:

$$FAR = \frac{False\ Alarm}{Hits + False\ Alarm} \tag{2}$$

False Alarm shows that the rainfall event is recorded at the satellite when no event is represented in IMD data. If the obtained FAR value tends towards 0, it denotes that the data are accurate, whereas if the value tends towards 1, it represents that the value is inaccurate regarding IMD data.

The critical success index (CSI) shows the correct estimation of the rain occurrence while removing the negative occurrence (Ayoub *et al.* 2020). It provides a critical success rate that blends the POD and the FAR, where a perfect score of 1 implies zero *False Alarm* and misses and can be mathematically represented in the following equation:

$$CSI = \frac{Hits}{Hits + Miss + False\ Alarm} \tag{3}$$

3.2. Machine learning models

3.2.1. Long short-term memory

Recurrent neural networks are the models which use input data over long sequences to obtain outputs based on the previous computations. These models help deal with non-linear time series but have drawbacks when dealing with predetermining time lags to learn temporal sequence processing (Gers *et al.* 1999; Gers & Schmidhuber 2001). To overcome the disadvantage of long-distance dependence tasks and diminish gradient in RNNs, LSTMs were introduced by Hochreiter (1997). Typically, an LSTM is made up of memory cells, and memory is transferred based on cell state and hidden state. The cell state's primary role is to pass the unchanged data forward, and the data in this cell state can be changed using sigmoid gates. Gates are similar to layers that are assigned individual weights. To avoid dependency problems for long intervals, LSTMs use gates for control of the memory process.

1. *Input gate*: The input gate of the model generates input coefficients (i_t) generated based on the current input (x_t) and the previously hidden layer (d_{t-1}), and based on the input coefficient (i_t), the cell state (L_t) is calculated based on the amount of information used out of the current input (x_t).
2. *Forget gate*: The amount of information of the previous cell state (L_{t-1}) to be stored in the current cell state (L_t) is determined based on the coefficient of forget gate (f_t).
3. *Output gate*: The final output of the network (d_t) is governed by the output coefficient (O_t) based on the relation $d_t = O_t$ tan$h$ (L_t)

The structure and equations of each gate used for the development of LSTM are presented in Figure 3.

LSTMs capture short-term and long-term relationships in time series using these three gate mechanisms and avoid vanishing the gradient. The main feature of the LSTM is to store the components of input for each time step in input features for
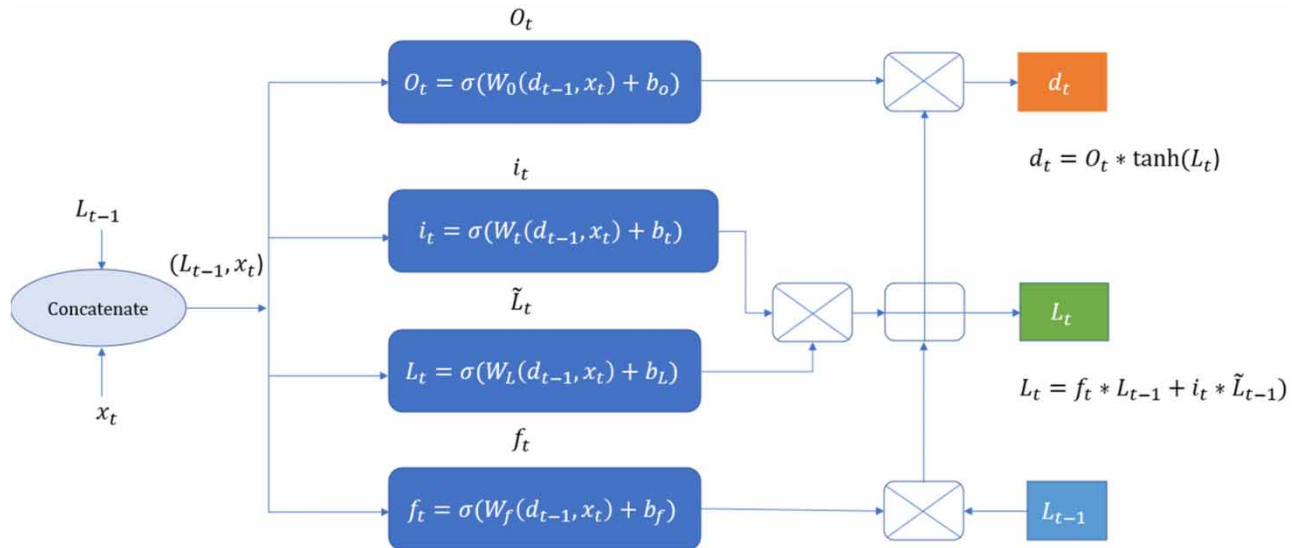
Figure 3 | A typical representation of the network of LSTM.

developing long-term memory. In addition, all the information features up to the current moment are stored in the hidden layers. Generally, these hidden layers are expressed using vectors of defined length and hence as time passes, the network compresses all the information from input components.

3.2.2. Extreme learning machines

Capturing the relationship between rainfall and runoff is quite tricky due to the high non-linear relationships between the variables. The interpretation of these non-linear relationships among these variables is out of the capacity of traditional models. A new approach using single-layer feedforward neural networks was proposed by Huang *et al.* (2004) to understand in-depth relationships in the data sets with fast learning capabilities and high generalization. In this approach, training of complex applications can be done in seconds which cannot be quickly done using orthodox neural network algorithms and produced the smallest training errors (Huang *et al.* 2006). ELM works on the same principles as that of backpropagation in prediction problems and data classification. Based on these features, analysis of components of climate like air temperature, rainfall, and solar energy with complex data sets can be easily carried out; hence, it can be considered as an ideal algorithm for forecasting atmospheric features (Wu & Chau 2010; Şahin 2013; Deo & Şahin 2015). The parameters that mainly govern the capacity of ELMs are the hidden nodes in the network, which can be mathematically expressed using the following equation as seen in the works of Huang *et al.* (2006), Şahin (2013), and Yaseen *et al.* (2016):

$$\sum_{i=1}^{s} M_i n_i \, (\alpha_i x_i + \beta_i) = P_i \tag{4}$$

where M represents the output vector weight between s number of nodes to the output nodes, $G(\alpha, \beta, x)$ represents the activation function of the hidden layer with α, β being the biases in the algorithm, and the ELM model's output is given by P. The hidden neuron index is represented by i.

The output from the network is governed by an activation function ($f(x)$). In this work, the sigmoidal function is chosen based on the previous works of Deo & Şahin (2015, 2016) and is represented in the following equation:

$$f(x) = \frac{1}{1 + \exp(-x)} \tag{5}$$

and the number of sample sets (N) can be given by the following equation:

$$\sum_{t=1}^{N} \| p_t - q_t \| \tag{6}$$

where the output of the ELM model at periods $(t) = 1,2,3…N$ is given by p_t, and the response variables used for developing the network are denoted by q_t.

In agreement with Equation (4), the output vector M can be estimated using the N number of data sets with a chain of linear equations

$$Y = GM \tag{7}$$

$$G(\alpha, \beta, x) = \begin{bmatrix} n(x_i) \\ . \\ . \\ n(x_N) \end{bmatrix} = \begin{bmatrix} n_i(\alpha_i x_i + \beta_i) \dots \dots n_L(\alpha_s x_i + \beta_s) \\ . \\ . \\ n_i(\alpha_N x_N + \beta_1) \dots \dots n_L(\alpha_s x_N + \beta_H) \end{bmatrix}_{N X s} \tag{8}$$

and

$$M = \begin{bmatrix} M_i^j \\ . \\ . \\ M_s^j \end{bmatrix}_{s X 1} \quad \text{and} \quad Y = \begin{bmatrix} y_i^j \\ . \\ . \\ y_N^j \end{bmatrix}_{N X 1} \tag{9}$$

where j represents the transpose of the vector or a matrix, and the hidden layer output matrix is represented as G.

Referring to the works of Yaseen *et al.* (2016), the output weights of the hidden layer network are show using the Moon-Penrose generalized inversion function (+) by inverting the hidden layer matrix

$$\hat{M} = G^+ Y \tag{10}$$

The output weights from the N data sets used in the modelling procedure are represented as $\hat{M}$.

Using Equation (11) with testing data sets (x_t), the predicted values of y can be obtained as

$$y = \sum_{i=1}^{s} \hat{M}_i n_i \ (\alpha_i x_t + \beta_i) \tag{11}$$

Readers are encouraged to refer to the works of Huang *et al.* (2006, 2015a, 2015b) for a detailed explanation of ELM. Figure 4 represents a distinct network of ELM. The descriptions of FFNN and MLR are presented in the Appendix.
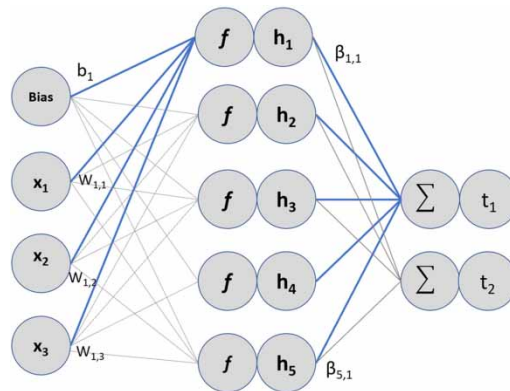
Figure 4 | A typical representation of a network of ELMs. Here, W represents the input layer weights, b represents randomly generated biases that are not involved in further training, f is the activation function, h_i represent the hidden neurons, X_i are the input data sets, β are the optimal values, and t_i are the target/output data sets.

3.3. Experimental setup

3.3.1. Selection of input variables

Choosing the potential input variables based on their relationship with output variables is more important than the algorithm's choice. This work's primary purpose is to understand the capability of SPPs in capturing the rainfall–runoff relationship, which can be an alternative ground-based rainfall data sets in real-time applications.

In this study, a rainfall–runoff modelling framework is considered where the dependent variable is runoff (output variable) and precipitation and historical runoff are the independent variables (input variables).

The framework of the adopted methodology can be mathematically represented as

$$S_t = f(R_t R_{t-1} \ldots R_{t-m} \ldots \ldots \ldots S_{t-1} \ldots S_{t-n}) \tag{12}$$

where $f(.)$ denotes the machine learning model, S_t represents the runoff at a time 't', and rainfall at the time 't' is represented by R_t. m and n denote the lag memory in the rainfall and runoff time series, respectively, and the number of inputs for the model is a function of m and n. Cross-correlation between rainfall and runoff and autocorrelation of rainfall and runoff has been used to get the m and n. For each precipitation data set and catchment, the importance of m and n was estimated, and accordingly, input variables were selected for the model training.

3.3.2. Model training

The total data length is divided into 70:30 for calibration and validation of the developed models. All the models were developed in MATLAB 2019a.

For MLR models, the regression model was fitted using the least-squares approach, and the error statistics were estimated for both calibration and validation periods. In FFNN, the number of hidden neurons used varied between 10 and 15. The transfer function used is TANSIG, and the Levenberg Marquardt (LM) algorithm is used as a training function. Several studies (Adamowski & Sun 2010; Maheswaran & Khosa 2013) have shown that the LM algorithm is the most robust training algorithm for feedforward networks.

For the ELM model, the number of neurons is determined using a trial and error approach to obtain the best calibration result. Validation data sets are then used as inputs at the same neurons for obtaining the forecasted results. For LSTM models, in order to determine the hyperparameters such as gradient threshold, initial learn rate, learn rate drop period, and learn rate factor, an optimization algorithm, namely Adam optimization (Kingma & Ba 2015), is used as it has been proven to be fruitful in determining the hyperparameters in LSTM models. Following the approach of Le *et al.* (2019), the maximum number of epochs was taken as 100,000 and early stopping techniques were used for the prevention of overfitting of the model.

3.4. Performance measures

Performance measures have been used as an index to determine the performance and understand the confidence limits of the model. The performance measures used in this study are RMSE, correlation (R^2), and NSE.

If the value of R^2 and NSE is inclined towards 0, it represents that the model's output is not a good representation of the original values. Whereas, if the value is inclined towards 1, the model is a good representation of the original system and its values. RMSE represents the amount of error in the output value of the model when compared to the original value. The lower the value of RMSE, the better the fitness of the model to the original system.

The performance measures are mathematically represented as

1. RMSE

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_i^{obs} - X_i^{sim})^2}{n}} \tag{13}$$

2. Correlation (R^2)

$$R^2 = \left[\frac{\sum_{i=1}^{n} (x_i^{obs} - \bar{x}^{obs})(x_i^{sim} - \bar{x}^{sim})}{\sqrt{\sum_{i=1}^{n} (x_i^{obs} - \bar{x}^{obs})^2 (x_i^{sim} - \bar{x}^{sim})^2}} \right]^2 \tag{14}$$

3. NSE

$$NSE = 1 - \left[\frac{\sum\limits_{i=1}^{n} (X_i^{obs} - X_i^{sim})^2}{\sum\limits_{i=1}^{n} (X_i^{obs} - \bar{x}^{obs})^2} \right] \qquad (15)$$

where X_i^{obs} and X_i^{sim} denote the observed and model-simulated streamflow values at any time instant 'i'; $\bar{x}^{obs}$ and $\bar{x}^{sim}$ represent the mean values of 'n' observed and simulated values, respectively.

4. RESULTS

We presented results in three sections. The first section shows the statistical analysis of SPP compared with the IMD data using metrics such as POD, FAR, and CSI. The selection of dominant timestep for input variables, namely CHIRPS, GPM-IMERG, and IMD, using Auto Correlation Function (ACF), Partial Auto Correlation Function (PACF), and cross-correlation to develop rainfall–runoff models is shown in section 4.2. Furthermore, the performance analysis of all the developed rainfall–runoff models is discussed. Finally, the accuracy of the best-performing models is shown to predict flood peaks in section 4.3.

4.1. Evaluation of precipitation products

This section presents the results from comparing precipitation products with ground-based (IMD) data sets using the metrics mentioned in section 3.1. In this study, the satellite-derived products are GPM-IMERG and CHIRPS, which are compared with IMD data sets (actual values) for accessing the capabilities of SPPs for the prediction of runoff. The cumulative distribution function (CDF) plots of IMD, GPM-IMERG, and CHIRPS data during the entire period for the VRB and the MRB are shown in Figure 5(a) and (b) with x-axis representing rainfall (x) and y-axis representing the probability. Based on the CDF plots, both the SPPs have similar CDF patterns compared to IMD data sets for the VRB and the MRB (Figure 5(a) and (b)). However, for Figure 5(a) in the rainfall range of 10–30 mm, underestimation is evident from the inset figure (zoomed plot) for both the data sets. In Figure 5(b), an underestimation of the rainfall event is observed in GPM-IMERG data for the range of 10–30 mm, whereas an underestimation in the 10–40 mm range is observed in the CHIRPS data.

In Table 1, POD values are in the range of 0.67–0.75, highlighting that both SPPs perform satisfactorily in capturing light rainfall (1–10 mm). However, for moderate (10–50 mm) and heavy (>50 mm) rainfalls, the POD values range between 0.29 and 0.45 for both the SPPs in both catchments. The same is in congruence with the FAR, highlighting false alarm for moderate and heavy rainfall in both the catchments for both SPPs. The CSI values show that SPPs have more than a 50% success rate for capturing light rainfall. However, the success rate ranges from 27 to 38% in capturing medium and heavy rainfall. All the observations from statistical parameters were further supported by RMSE, both the SPPs exhibited low RMSE values (<1.5) for the rainfall range of 1–10 mm, confirming that the accuracy of SPPs in capturing lower rainfall is robust. Whereas for medium and high rainfall (10–50 and >50 mm), higher RMSE values are observed, indicating a significant deviation of the rainfall values from the original values (IMD).

Based on the above analysis, both the SPPs are close to IMD data sets, but CHIRPS data are better than GPM-IMERG for medium and high rainfall. Based on this understanding, models for 1-day ahead discharge prediction and peak flow prediction are developed. Also, this analysis helps understand the probable reason for abnormalities in the predictive capacity of the model inflows, along with the proposal of the best replacement of IMD data.

4.2. Model parameters and performance analysis

The cross-correlation function between the input variables such as CHIRPS, GPM-IMERG, and IMD data and runoff is plotted in Figure 6(a)–(c), (e) and (f) for the VRB and the MRB, respectively. To select the number of lags in rainfall time series having significant correlation with runoff, a threshold value of 0.25 is considered. The cross-correlation (Figure 6) reveals that rainfall–runoff dependency is up to 4 days, 1 day, and 7 days for the VRB and 3 days, 3 days, and 4 days for the MRB, respectively. The ACF and PACF corresponding to the river basins' runoff variable are plotted in Figure 7(a)–(d), respectively.
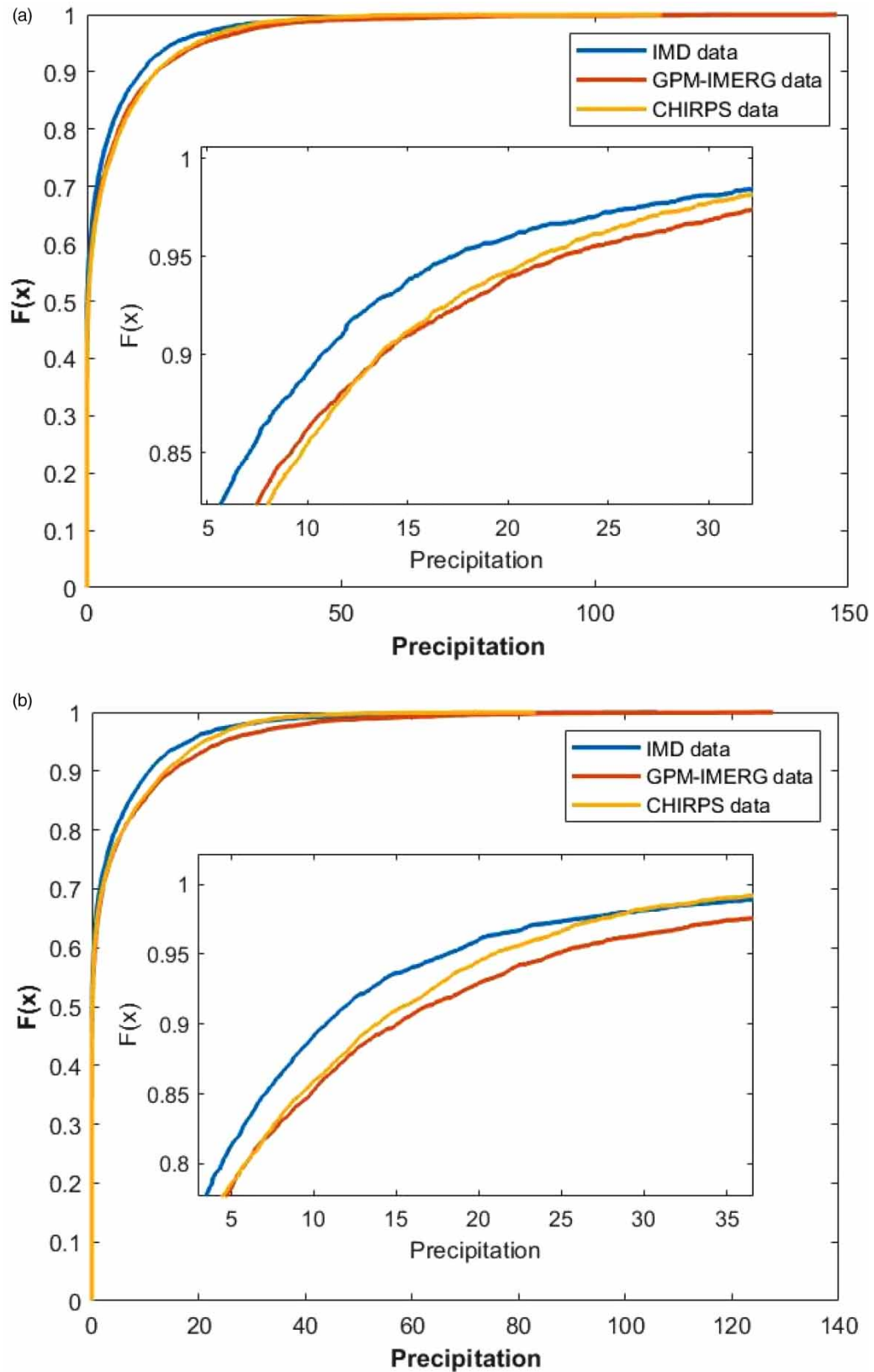
Figure 5 | Cumulative distributive function for satellite data sets and IMD-based precipitation for the VRB (a) and MRB (b) The inner plot shows the zoomed plot.

Table 1 | Estimating the detection metrics for the SPPs with reference to the gauge-based IMD data in Vamsadhara and Mahanadi river basins for different rainfall thresholds

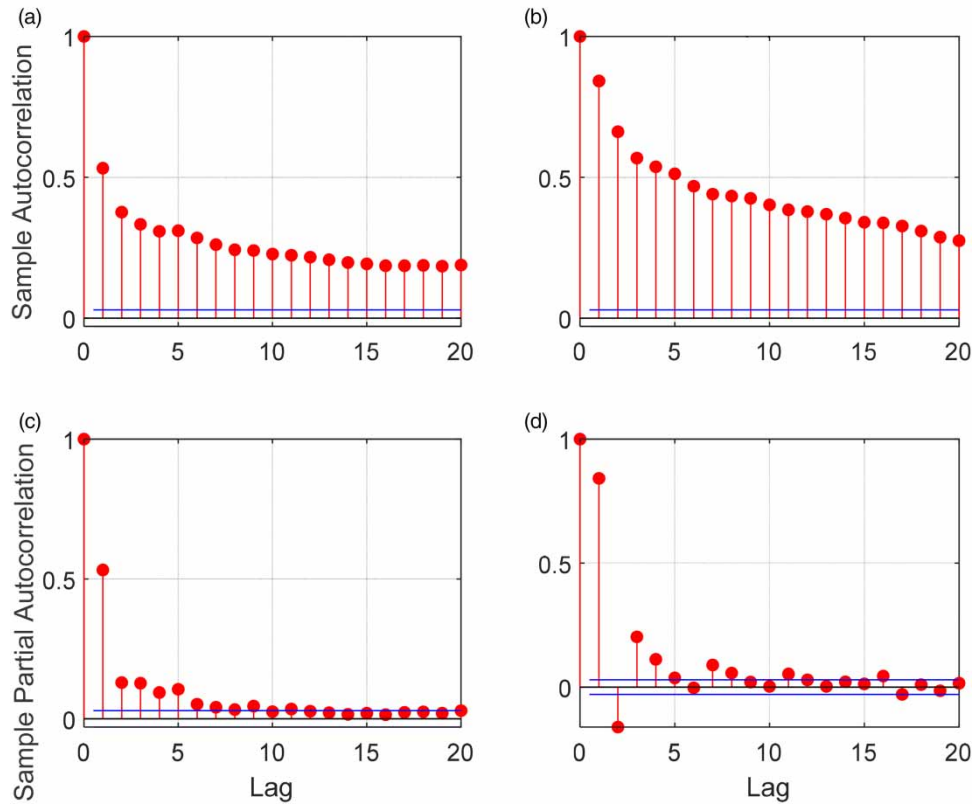| Basin	Vamsadhara						Mahanadi					
	GPM-IMERG			CHIRPS			GPM-IMERG			CHIRPS		
Metrics	1–10 mm	10–50 mm	> 50 mm	1–10 mm	10–50 mm	> 50 mm	1–10 mm	10–50 mm	> 50 mm	1–10 mm	10–50 mm	> 50 mm
POD	0.71	0.38	0.29	0.73	0.41	0.34	0.67	0.37	0.30	0.75	0.45	0.35
FAR	0.32	0.62	0.72	0.30	0.60	0.68	0.31	0.57	0.67	0.25	0.56	0.65
CSI	0.54	0.35	0.30	0.56	0.36	0.32	0.56	0.34	0.27	0.65	0.38	0.30
RMSE (mm)	1.24	7.34	10.14	1.10	6.5	9.14	1.31	8.23	9.61	1.45	8.76	7.12

The ACF and PACF show that the runoff has significant lag memory up to 7 days and 8 days. Selected lagged variables are used to develop the rainfall–runoff models, including linear regression (MLR) and machine learning models like FFNN, ELM, and LSTM. All the models are calibrated and validated for rainfall–runoff using the approaches mentioned in section 3.4. Each model is trained until the calibration error decreases and reaches a minimum. For the least error model, validation data set is used for obtaining the predicted runoff. In MLR, a simple linear relation is used for training, whereas in FFNN, ELM, and LSTM models, the least error is obtained by trial and error by changing the number of neurons and learning rates. The obtained predicted runoff data sets are compared to the observed data sets and are evaluated using the performance measures in section 3.5.

Table 2 represents the results of all the rainfall–runoff models generated with different rainfall products resulting from the different models considered in the study. The values of the performance measures for accessing the quality of the predicted
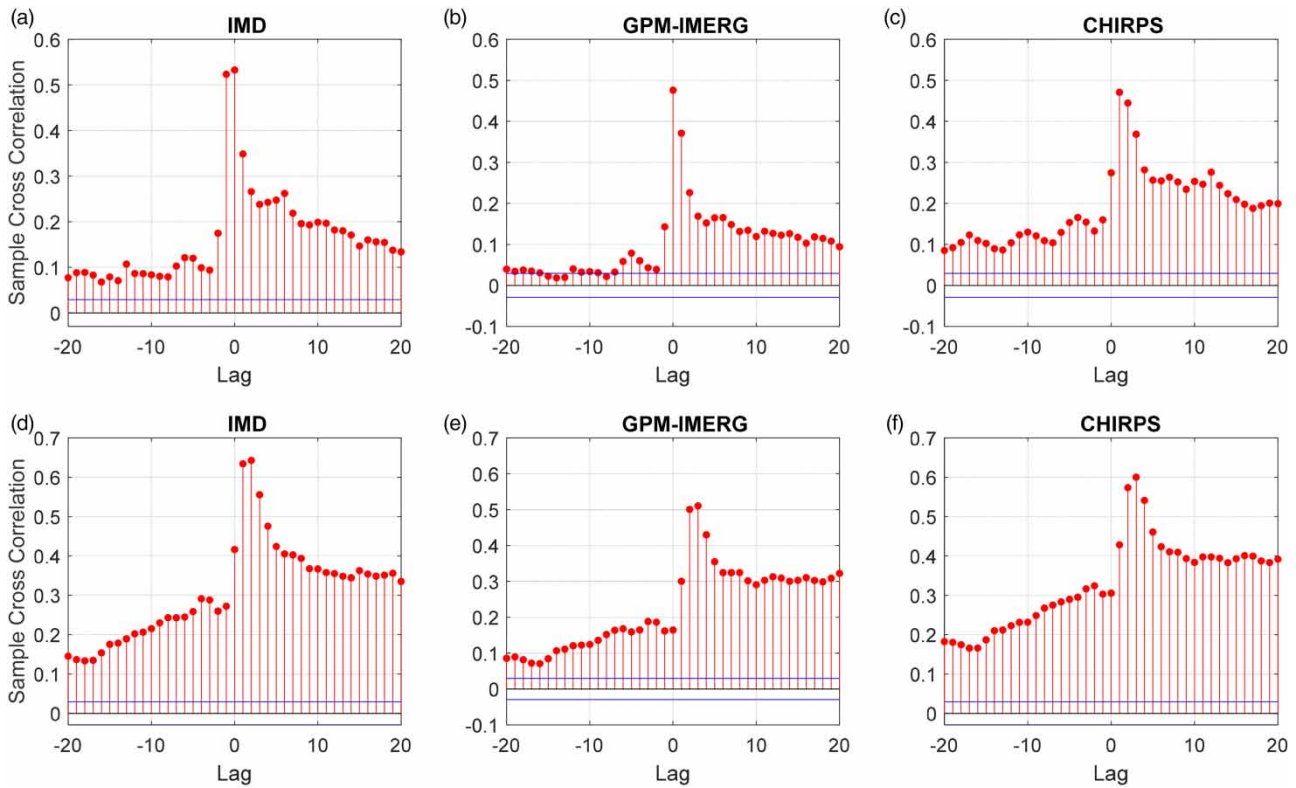
Figure 6 | The cross-correlation function between the input variables (IMD, GPM-IMERG, and CHIRPS data) and runoff (a–c) and (d–f) for VRB and MRB, respectively.

Figure 7 | Correlation plots for a runoff with (a) and (c) are representing ACF and PACF for the VRB, and (b) and (d) for the MRB, respectively.

Table 2 | Performance measures of various models in calibration and validation for the Vamsadhara and Mahanadi river basins

Data	IMD											
	Vamsadhara river basin						**Mahanadi river basin**					
	Calibration			**Validation**			**Calibration**			**Validation**		
Models	RMSE (m³/s)	R^2	NSE	RMSE (m³/s)	R^2	NSE	RMSE (m³/s)	R^2	NSE	RMSE (m³/s)	R^2	NSE
MLR	230.84	0.50	0.23	163.47	0.625	0.360	325.30	0.65	0.39	373.70	0.69	0.46
FFNN	202.45	0.64	0.41	145.67	0.702	0.492	228.02	0.83	0.74	243.12	0.77	0.72
ELM	120.42	0.82	0.67	110.38	0.83	0.69	190.25	**0.92**	**0.89**	**198.34**	**0.92**	**0.87**
LSTM	**105.42**	**0.86**	**0.69**	**93.60**	**0.85**	**0.71**	191.52	0.92	0.87	201.30	0.92	0.84
	GPM-IMERG											
MLR	227.95	0.52	0.25	162.20	0.62	0.37	356.30	0.73	0.70	394.34	0.78	0.71
FFNN	184.70	0.71	0.51	142.12	0.72	0.51	238.60	0.83	0.79	293.42	0.82	0.76
ELM	128.24	0.81	0.63	129.45	0.81	0.63	221.45	**0.89**	**0.86**	**215.45**	**0.89**	**0.87**
LSTM	**105.72**	**0.82**	**0.68**	**94.10**	**0.81**	**0.67**	223.42	0.86	0.82	220.56	0.88	0.85
	CHIRPS											
MLR	206.52	0.62	0.39	157.18	0.64	0.40	357.07	0.72	0.69	391.80	0.79	0.70
FFNN	159.68	0.76	0.57	145.31	0.79	0.59	243.81	0.82	0.78	254.92	0.81	0.75
ELM	138.45	0.83	0.60	135.52	0.83	0.61	200.47	**0.91**	**0.89**	**208.27**	**0.92**	**0.88**
LSTM	**107.54**	**0.84**	**0.68**	**98.32**	**0.84**	**0.65**	220.87	0.89	0.83	217.50	0.89	0.83

Bold values represent the models with best performance.

discharge are mentioned in section 3.5, i.e., RMSE, R^2, and NSE for both the VRB and the MRB in calibration and validation stages. Values of these measures are used to understand the best-performing model and the best SPP as a replacement of IMD data sets. Results are discussed for each considered basin separately.

4.2.1. Vamsadhara river basin

4.2.1.1. Models based on the IMD data.   Comparing the models developed using IMD data sets (Table 2) shows that the LSTM model performs better than other MLR, FFNN, and ELM models. For example, during the calibration and validation stages, the LSTM model has an NSC value of 0.69 and 0.71, whereas MLR, FFNN, and ELM models have values of 0.23, 0.41, 0.67 and 0.36, 0.49, 0.69, respectively. Similarly, lower values of RMSE and higher correlation were observed for LSTM in validation with values equal to 93.6 and 0.85, respectively.

4.2.1.2. Models based on the GPM-IMERG data.   Investigating the performance of the models with the GPM-IMERG data, models such as MLR, FFNN, and ELM were unable to produce good modelling results during both calibration and validation stages. In this case, LSTM models produced comparatively better results than the rest of the models with NSC values equal to 0.68 and 0.67 in calibration and validation stages. In contrast to this, the values of 0.25, 0.51, 0.3 and 0.37, 0.51, 0.63 were observed for MLR, FFNN, and ELM models, respectively. Similarly, better values are observed for RMSE and correlation in LSTM models.

4.2.1.3. Models based on the CHIRPS data.   Using CHIRPS as input data, Table 2 shows that the LSTM model produces better results than the other developed models. Higher values of RMSE and correlation were observed for LSTM models along with high NSC values. For example, MLR, FFNN, and ELM models produced results with NSC values of 0.39, 0.57, 0.60 and 0.40, 0.59, 0.61 in calibration and validation stages. Whereas, LSTM models produced values of 0.68 and 0.65, respectively, in calibration and validation stages.

4.2.2. Mahanadi river basin

4.2.2.1. Models based on the IMD data.   Investigating the performance of several linear regression (MLR) as well as machine learning models (FFNN, ELM, and LSTM) with IMD data sets, results (Table 2) show that MLR, FFNN, and LSTM have obtained NSC values of 0.39, 0.74, 0.87 and 0.69, 0.77, 0.92 in calibration and validation stages. The best model output was observed in ELM models with NSC values of 0.89, 0.92 in calibration and validation stages. Similarly, the least RMSE and high correlation values are observed in ELM models.

4.2.2.2. Models based on the GPM-IMERG data.   Several models were developed using GPM-IMERG data as input for the rainfall–runoff modelling. ELM models produced the best results with high correlation, NSC with least RMSE in both calibration and validation stages. For example, ELM obtained NSC values of 0.86 and 0.87 in calibration and validation stages, respectively. Whereas the MLR, FFNN, and LSTM produced values of 0.70, 0.79, 0.82 and 0.71, 0.76, 0.85 in calibration and validation stages, respectively.

4.2.2.3. Models based on the CHIRPS data.   Comparing the results obtained using CHIRPS data sets as input, ELM models showed better modelling results with the least RMSE and high NSE and correlation. For example, MLR, FFNN, and LSTM models have NSE values of 0.69, 0.78, 0.82 and 0.70, 0.75, 0.88 in calibration and validation stages, respectively, when compared to the ELM models having NSE values of 0.89 and 0.88.

Overall, LSTM models in the VRB and the MRB outperform other models with all precipitation sets. Comparing LSTM models with GPM-IMERG and CHIRPS data sets for the VRB, both the models were found to produce similar results, but comparatively, CHIRPS model results were closer to the results of models with IMD data. For example, the correlation of IMD-based models is 0.85, whereas GPM-IMERG- and CHIRPS-based models have a significance of 0.81 and 0.84, respectively, in the validation stage. Similarly, comparing the results of ELM models with GPM-IMERG and CHIRPS data sets for the MRB, both the models were found to produce similar results. Still, the CHIRPS model results were closer to the results of models with the IMD data. For example, the correlation of IMD-based models is 0.92, whereas GPM-IMERG- and CHIRPS-based models correlate at 0.89 and 0.92, respectively.

## 4.3. Application of the best model for forecasting flood peaks

Although promising results were observed in models for 1-day ahead runoff prediction, there is yet a concern regarding the capabilities of the models for peak flow prediction during extreme events. Models have been applied to estimate 1-day ahead flood prediction during the monsoon season to assess their ability to capture the peak events. For this purpose, we analyzed four flood events, each in the VRB and the MRB, and the corresponding results obtained from the models. In this section, only the best models for each river basin are reported, as their superiority over other models has been proven in the earlier section. Accordingly, for Vamsadhara, the LSTM model is considered, and the ELM model is considered for the Mahanadi basin. Figure 8(a) and (b) represents the spatial variation of 1-day GPM-IMERG precipitation in the VRB and the MRB for one of the extreme events that occurred on 8 August 2009. Figure 9(a) and (b) shows the results of LSTM and ELM models with different precipitation products along with the observed runoff for the VRB, and Figure 9(c) and (d) shows the runoff forecasts from LSTM and ELM models for the MRB. The result statistics only for the best models are presented in Table 3.

In the VRB, for the year 2009, based on the value of peak error percentage, the IMD-LSTM model produced a reasonable forecast of the flood peak with 17.65%, followed by CHIRPS-LSTM with 27.23%. Similarly, in 2010, the percentage error in estimating the flood peak was least for IMD-LSTM with 9.60%, followed by CHIRPS-LSTM with 16.20%. In the flood event in 2011, the least peak error was observed in IMD-LSTM, followed by the model with CHIRPS as input. In 2012, CHIRPS-LSTM was the best model with a peak error percentage of 25.61%, which outperformed the model with IMD data, an error percentage of 35.85%. In all the years, the performance of GPM-IMERG-based models in terms of flood peak is not satisfactory.

In the MRB, for the year 2009, the CHIRPS-ELM model produced good flood peak forecasts based on the peak error percentage with a value of 7.75%, followed by IMD-ELM models with −7.32%. Similarly, in 2010, the percentage error in the estimation of flood peak was least for IMD-ELM with −8.30%, followed by CHIRPS-ELM with 19.23%. In the flood event in 2011, the least error was found for the ELM model with CHIRPS input. For the flood event in 2012, the CHIRPS-ELM model
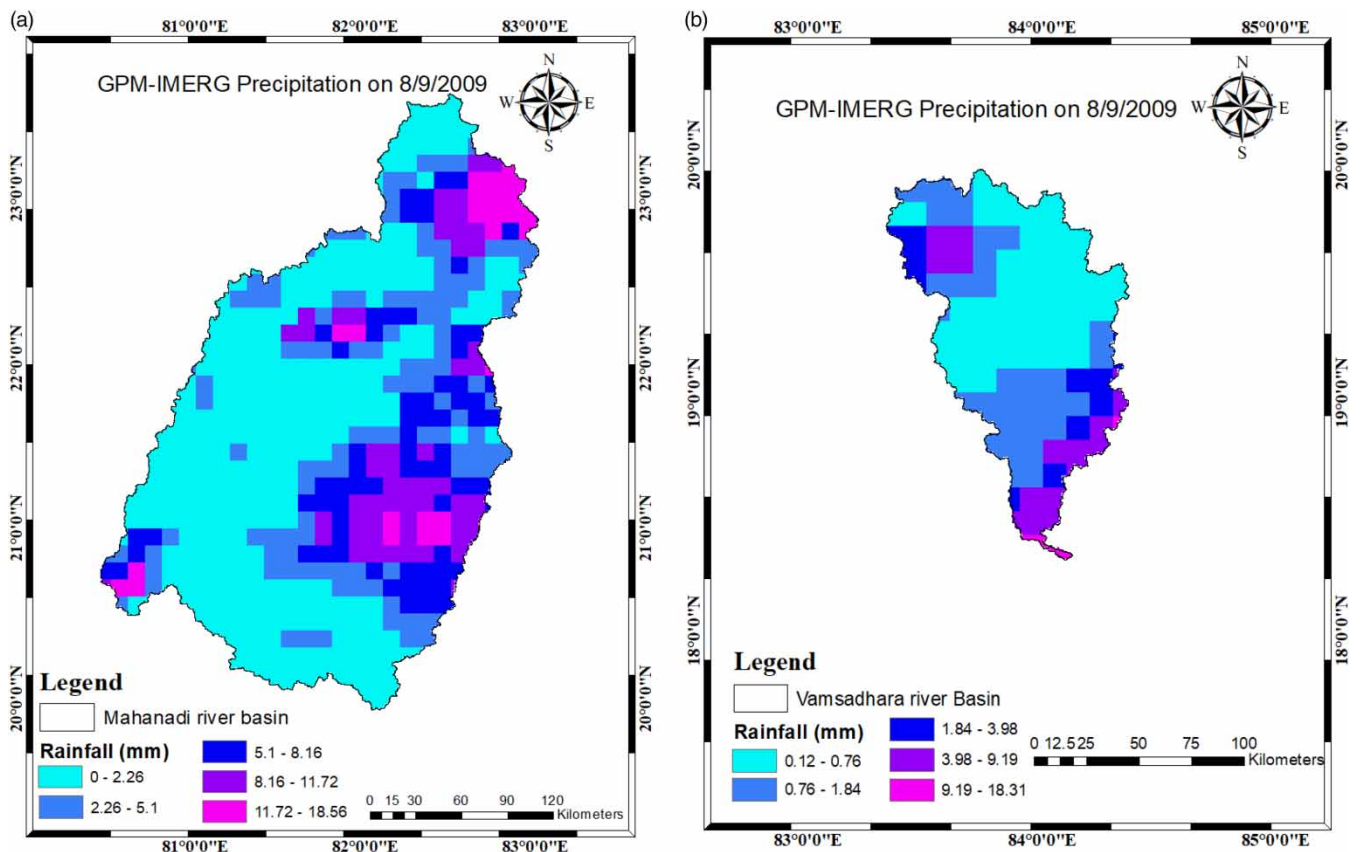
Figure 8 | (a) The spatial distribution of the rainfall on 8 August 2009 as obtained using the GPM-IMERG data set in (a) the MRB and (b) the VRB.
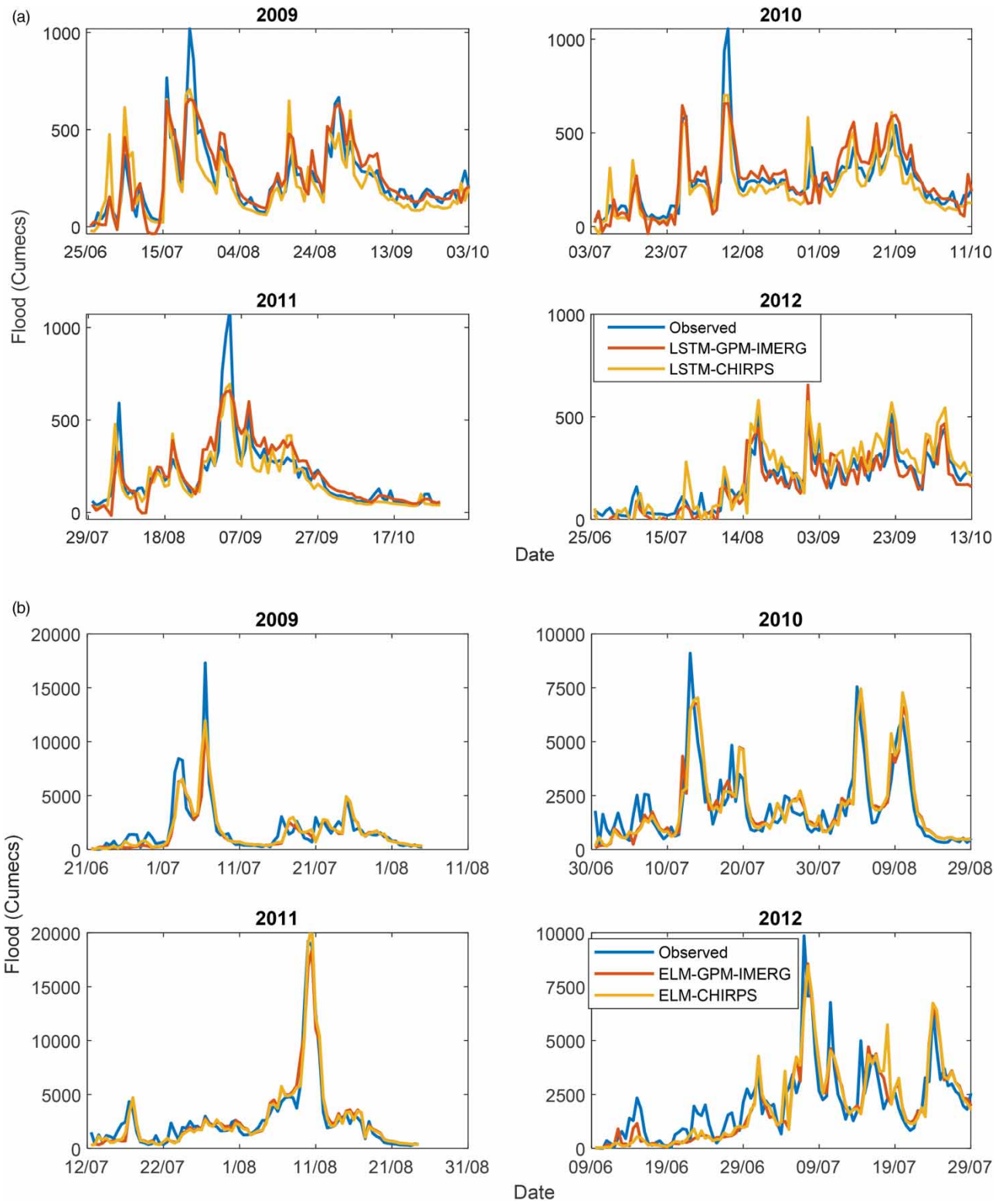
Figure 9 | (a) Hydrographs of flood peak in the VRB with LSTM models. (b) Hydrographs of flood peak in the MRB with ELM models.

obtained an error percentage of 3.05%, which outperformed IMD-ELM. The results show that CHIRPS models have pro-
duced good results in predicting flood peaks with the least errors in most flood events and were closely followed by
models with IMD data sets, which over-predicted the flood events.

Table 3 | Accuracy of the best models in each basin to predict the flood peak timing and the amount of error in the predicted peak compared to the observed values

Year	Model	Observed flood value (m³/s)	Predicted flood value (m³/s)	Error (%)	Model	Observed flood value (m³/s)	Predicted flood value (m³/s)	Error (%)
	Vamsadhara river basin				Mahanadi river basin			
2009	IMD_LSTM	1,553	1,320	17.65	IMD_ELM	16,120	17,310	−7.32
	GPM-IMERG_LSTM		1,020	34.32	GPM-IMERG_ELM		13,220	17.99
	CHIRPS_LSTM		1,130	**27.23**	CHIRPS_ELM		14,960	**7.75**
2010	IMD_LSTM	1,563	1,426	9.60	IMD_ELM	8,406	9,104	−8.30
	GPM-IMERG_LSTM		1,056	48.01	GPM-IMERG_ELM		6,765	24.25
	CHIRPS_LSTM		1,345	**16.20**	CHIRPS_ELM		7,050	**19.23**
2011	IMD_LSTM	1,948	1,654	17.77	IMD_ELM	23,370	19,230	21.52
	GPM-IMERG_LSTM		1,452	34.15	GPM-IMERG_ELM		18,370	27.21
	CHIRPS_LSTM		1,545	**26.08**	CHIRPS_ELM		20,580	**13.55**
2012	IMD_LSTM	729.3	532.9	36.85	IMD_ELM	8,825	9,857	−10.40
	GPM-IMERG_LSTM		500.1	45.80	GPM-IMERG_ELM		8,459	4.32
	CHIRPS_LSTM		580.6	**25.61**	CHIRPS_ELM		8,563	**3.05**

Bold values represent the models with least error in prediction of flood peak.

Overall, for both river basins, the models were able to capture the peak floods and the time to peak to greater accuracy. Comparing the performance of the models between two rivers basins, the models perform less accurately in the VRB than the MRB. This difference in the model performance across the two river basins could be attributed to higher variations in the hydrograph and more complexity in rainfall–runoff relationships in Vamsadhara. Owing to the large catchment size, in the MRB, the storage-induced inertia dampens the non-linear features and makes the catchment more amenable for modelling (Maheswaran & Khosa 2012). Whereas, the non-linear features are more prominent in the VRB, which is 10 times smaller.

5. DISCUSSION

This study aims to develop a machine learning-based rainfall–runoff model using SPPs. The statistical analysis of the precipitation products showed that the performance of the SPPs was dependent on the rainfall ranges. GPM-IMERG is sensitive to light and medium rainfall; however, it underestimates heavy rainfall. On the contrary, the performance of the CHIRPS data set can capture the heavy rainfall more accurately than GPM-IMERG.

Further comparison of the model performance observed that the LSTM and ELM models performed better than the FFNN and linear regression models. Of the two best models, LSTM was more accurate for the VRB and ELM to be the best model for the MRB. Comparing the results for 1-day ahead forecasts in the MRB obtained by Nanda et al. (2016) from TRMM-driven wavelet-based non-linear autoregressive with exogenous inputs (WNARX) models found that that the results from the present are superior. In their study, Nanda et al. (2016) reported NSE values of 0.85 during the testing period using the TRMM-RT rainfall dataset-driven WNARX model. However, the present study showed that the GPM-IMERG-ELM models and CHIRPS-ELM model yielded NSE values of 0.87 and 0.88, respectively.

Furthermore, models using GPM-IMERG and CHIRPS data in combination with ELM and LSTM have been shown to outperform the TRMM-based WNARX models as reported by Nanda et al. (2016). Similarly, in another study, Yeditha et al. (2020) developed WNARX models using TRMM for the VRB and obtained NSE results of 0.629 during the validation period. Comparing that with the present study using GPM-IMERG-LSTM (NSE = 0.67) and CHIRPS-LSTM (NSE = 0.65) shows that the present study results are superior.

From the above comparison of the results, the GPM-IMERG and CHIRPS data sets are performing better in capturing the rainfall than the TRMM data set for the two river basins. Similar observations were obtained by Liu & Zipser (2015) and Prakash et al. (2018). The authors show that the GPM-IMERG data set shows significant improvement in reduced false alarms and improved extreme rainfall detection over the TRMM data. Furthermore, the deep learning models, ELM and LSTM, showed a better rainfall–runoff relationship than the other contemporary models. Overall, these SPPs can be used as an alternative for ground-based data for rainfall–runoff modelling with prediction capability. Since GPM-IMERG data

are available in real time, they can be explored for real-time forecasting purposes. Additionally, these models can be applied reliably for multistep ahead forecasting when precipitation forecasts from weather models such as European Center for Medium Weather Forecast are used.

One possible alternate method to improve the model performance is developing a framework for correcting the SPPs and the same can be considered as a future direction. Similarly, considering the ensemble model approach and wavelet-based hybrid approach would reduce the uncertainty in the model results.

6. CONCLUSIONS

A novel data-driven rainfall–runoff model is proposed using two different satellite-based precipitation products. Satellite-based precipitation products GPM-IMERG and CHIRPS were evaluated with ground-based IMD data sets. Results showed that the CHIRPS data sets are a better proxy of the IMD data sets compared to the GPM-IMERG data. Following this, the model is tested on two flood-prone river basins. These SPPs were used to drive ELM, LSTM, FFNN, and MLR rainfall–runoff models. SPP-driven LSTM and ELM models produced reliable forecasts in both the river basins and are comparable to the IMD data-driven models. The CHIRPS-LSTM and CHIRPS-ELM models produced robust VRB and MRB results, respectively. For simulating the peak floods, the models driven by the CHIRPS data set were accurate than the GPM-IMERG data. Nevertheless, the results based on the latter were promising.

Nonetheless, the present outcomes regarding the applicability of the SPPs, their accuracy, and the capabilities of the proposed models need to be confirmed and supported further through their application to various river basins with different climate and catchment characteristics.

COMPETING INTERESTS

The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTION

P.K.Y., S.S.N., and M.R. conceptualized the work, performed primary data collection and analysis, and wrote the initial version of the paper. P.K.Y., M.R., and A.A. had an extensive discussion and structured the initial draft of the manuscript. B.B. participated in revising the work, providing important intellectual content as well as interpretations and conclusions. All the authors contributed to the revision.

DATA AVAILABILITY STATEMENT

All relevant data are available from an online repository or repositories. https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_daily/netcdf/p25/ http://www.imdpune.gov.in/. https://pmm.nasa.gov/data access/downloads/gpm.

REFERENCES

Adamowski, J. & Sun, K. 2010 Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. *Journal of Hydrology* **390** (1–2), 85–91.

Adamowski, J., Fung Chan, H., Prasher, S. O., Ozga-Zielinski, B. & Sliusarieva, A. 2012 Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resources Research* **48** (1), 1–14.

Ahani, A., Shourian, M. & Rahimi Rad, P. 2018 Performance assessment of the linear, nonlinear and nonparametric data driven models in river flow forecasting. *Water Resources Management* **32** (2), 383–399.

Ali, S. & Shahbaz, M. 2020 Streamflow forecasting by modeling the rainfall–streamflow relationship using artificial neural networks. *Modeling Earth Systems and Environment* **6** (3), 1645–1656.

Alizamir, M., Kisi, O. & Zounemat-Kermani, M. 2018 Modelling long-term groundwater fluctuations by extreme learning machine using hydro-climatic data. *Hydrological Sciences Journal* **63** (1), 63–73.

Atiquzzaman, M. & Kandasamy, J. 2016 Prediction of hydrological time-series using extreme learning machine. *Journal of Hydroinformatics* **18** (2), 345–353.

Ayoub, A. B., Tangang, F., Juneng, L., Tan, M. L. & Chung, J. X. 2020 Evaluation of gridded precipitation datasets in Malaysia. *Remote Sensing* **12** (4), 613.

Bitew, M. M. & Gebremichael, M. 2011 Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands. *Hydrology and Earth System Sciences* **15** (4), 1147–1155.

Brath, A. & Rosso, R. 1993 Adaptive calibration of a conceptual model for flash flood forecasting. *Water Resources Research* **29** (8), 2561–2572.

Chen, Y., Li, J. & Xu, H. 2016 Improving flood forecasting capability of physically based distributed hydrological models by parameter optimization. *Hydrology and Earth System Sciences* **20** (1), 375–392.

Couta, D., Zhang, Y. & Li, Y. 2019 River flow forecasting using long short-term memory. In: *DEStech Transactions on Computer Science and Engineering (ICAIC)*, pp. 142–147.

Deo, R. C. & Şahin, M. 2015 Application of the extreme learning machine algorithm for the prediction of monthly effective drought index in eastern Australia. *Atmospheric Research* **153**, 512–525.

Deo, R. C. & Şahin, M. 2016 An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland. *Environmental Monitoring and Assessment* **188** (2), 1–24.

Ding, Y., Zhu, Y., Wu, Y., Jun, F. & Cheng, Z. 2019 Spatio-temporal attention lst model for flood forecasting. In: *Proceedings – 2019 IEEE International Congress on Cybermatics: 12th IEEE International Conference on Internet of Things, 15th IEEE International Conference on Green Computing and Communications, 12th IEEE International Conference on Cyber, Physical and Social Computing*, pp. 458–465.

Ding, Y., Zhu, Y., Feng, J., Zhang, P. & Cheng, Z. 2020 Interpretable spatio-temporal attention LSTM model for flood forecasting. *Neurocomputing* **403**, 348–359.

Elsafi, S. H. 2014 Artificial neural networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Engineering Journal* **53** (3), 655–662.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. & Michaelsen, J. 2015 The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes. *Scientific Data* **2**, 1–21.

Gadhawe, M. A., Guntu, R. K. & Agarwal, A. 2021 Network-based exploration of basin precipitation based on satellite and observed data. *The European Physical Journal Special Topics* **123**, 3343–3357.

Gers, F. A. & Schmidhuber, J. 2001 LSTM recurrent networks learn simple context-free and context-sensitive languages. *IEEE Transactions on Neural Networks* **12** (6), 1333–1340.

Gers, F. A., Schmidhuber, J. & Cummins, F. 1999 Learning to forget: continual prediction with LSTM. *IEE Conference Publication* **2** (470), 850–855.

Guntu, R. K., Rathinasamy, M., Agarwal, A. & Sivakumar, B. 2020 Spatiotemporal variability of Indian rainfall using multiscale entropy. *Journal of Hydrology* **587** (December 2019), 124916.

Hadid, B., Duviella, E. & Lecoeuche, S. 2020 Data-driven modeling for river flood forecasting based on a piecewise linear ARX system identification. *Journal of Process Control* **86**, 44–56.

Hiep, N. H., Luong, N. D., Viet Nga, T. T., Hieu, B. T., Thuy Ha, U. T., Du Duong, B., Long, V. D., Hossain, F. & Lee, H. 2018 Hydrological model using ground- and satellite-based data for river flow simulation towards supporting water resource management in the Red River Basin, Vietnam. *Journal of Environmental Management* **217**, 346–355.

Hochreiter, S. 1997 Long short-term memory. *Neural Computation* **1780**, 1735–1780.

Hu, Y., Yan, L., Hang, T. & Feng, J. 2020 Stream-Flow Forecasting of Small Rivers Based on LSTM. arXiv preprint arXiv:2001.05681.

Huang, G.-B., Zhu, Q. Y. & Siew, C. K. 2004 Extreme learning machine: a new learning scheme of feedforward neural networks. *IEEE International Conference on Neural Networks – Conference Proceedings* **2** (August), 985–990.

Huang, G.-B., Zhu, Q. Y. & Siew, C. K. 2006 Extreme learning machine: theory and applications. *Neurocomputing* **70** (1–3), 489–501.

Huang, G., Huang, G.-B., Song, S. & You, K. 2015a Trends in extreme learning machines: a review. *Neural Networks* **61**, 32–48.

Huang, Y., Zhang, Z. & Jin, Q. 2015b Evaluation of TRMM Satellite Rainfall in Driving Variable Infiltration Capacity (VIC) Model in Ganjiang Basin (ICCET), pp. 1048–1051.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K. L., Joyce, R. J., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., Wolff, D. B. & Xie, P. 2020 Integrated multi-satellite retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). *Advances in Global Change Research* **67**, 343–353.

Islam, A. S. 2010 Improving flood forecasting in Bangladesh using an artificial neural network. *Journal of Hydroinformatics* **12** (3), 351–364.

Janowiak, J. E., Xie, P., Joyce, R. J., Chen, M. & Yarosh, Y. 2004 Validation of satellite-derived rainfall estimates and numerical model forecasts of precipitation over the United States. In: *Proceedings of the 2nd International Precipitation Working Group Workshop*, January, p. 11.

Kan, G., He, X., Ding, L., Li, J., Liang, K. & Hong, Y. 2017 Study on applicability of conceptual hydrological models for flood forecasting in humid, semi-humid semi-arid and arid basins in China. *Water (Switzerland)* **9** (10), 1–25.

Kasi, V., Yeditha, P. K., Rathinasamy, M., Pinninti, R., Landa, S. R., Sangamreddi, C., Agarwal, A. & Dandu Radha, P. R. 2020 A novel method to improve vertical accuracy of CARTOSAT DEM using machine learning models. *Earth Science Informatics* **13** (4), 1139–1150.

Khac-Tien Nguyen, P. & Hock-Chye Chua, L. 2012 The data-driven approach as an operational real-time flood forecasting model. *Hydrological Processes* **26** (19), 2878–2893.

Kim, K., Park, J., Baik, J. & Choi, M. 2017 Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over far-east Asia. *Atmospheric Research* **187**, 95–105.

Kingma, D. P. & Ba, J. L. 2015 *Adam: A Method for Stochastic Optimization*, pp. 1–15.

Kişi, Ö. 2008 River flow forecasting and estimation using different artificial neural network techniques. *Hydrology Research* **39** (1), 27–40.

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.

Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S. & Nearing, G. S. 2019 Toward improved predictions in ungauged basins: exploiting the power of machine learning. *Water Resources Research* **55** (12), 11344–11354.

Lai, C., Zhong, R., Wang, Z., Wu, X., Chen, X., Wang, P. & Lian, Y. 2019 Monitoring hydrological drought using long-term satellite-based precipitation data. *Science of the Total Environment* **649** (August), 1198–1208.

Le, X. H., Ho, H. V., Lee, G. & Jung, S. 2019 Application of Long Short-Term Memory (LSTM) neural network for flood forecasting. *Water (Switzerland)* **11** (7), 1387.

Li, D., Christakos, G., Ding, X. & Wu, J. 2018 Adequacy of TRMM satellite rainfall data in driving the SWAT modeling of Tiaoxi catchment (Taihu Lake Basin, China). *Journal of Hydrology* **556** (December), 1139–1152.

Li, W., Kiaghadi, A. & Dawson, C. N. 2020 *High Temporal Resolution Rainfall Runoff Modelling Using Long-Short-Term-Memory (LSTM) Networks*, pp. 1–21.

Liu, C. & Zipser, E. J. 2015 The global distribution of largest, deepest, and most intense precipitation systems. *Geophysical Research Letters* **42** (9), 3591–3595.

Maheswaran, R. & Khosa, R. 2012 Comparative study of different wavelets for hydrologic forecasting. *Computers and Geosciences* **46**, 284–295.

Maheswaran, R. & Khosa, R. 2013 Wavelets-based non-linear model for real-time daily flow forecasting in Krishna River. *Journal of Hydroinformatics* **15** (3), 1022–1041.

Mazrooei, A. & Sankarasubramanian, A. 2019 Improving monthly streamflow forecasts through assimilation of observed streamflow for rainfall-dominated basins across the CONUS. *Journal of Hydrology* **575** (May), 704–715.

Meng, J., Li, L., Hao, Z., Wang, J. & Shao, Q. 2014 Suitability of TRMM satellite rainfall in driving a distributed hydrological model in the source region of Yellow River. *Journal of Hydrology* **509** (November 1997), 320–332.

Munzimi, Y. A., Hansen, M. C. & Asante, K. O. 2019 Estimating daily streamflow in the Congo Basin using satellite-derived data and a semi-distributed hydrological model. *Hydrological Sciences Journal* **64** (12), 1472–1487.

Nair, S., Srinivasan, G. & Nemani, R. 2009 Evaluation of multi-satellite TRMM derived rainfall estimates over a western state of India. *Journal of the Meteorological Society of Japan* **87** (6), 927–939.

Nanda, T., Sahoo, B., Beria, H. & Chatterjee, C. 2016 A wavelet-based non-linear autoregressive with exogenous inputs (WNARX) dynamic neural network model for real-time flood forecasting using satellite-based rainfall products. *Journal of Hydrology* **539**, 57–73.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10** (3), 282–290.

Nguyen, P. K. T., Chua, L. H. C. & Son, L. H. 2014 Flood forecasting in large rivers with data-driven models. *Natural Hazards* **71** (1), 767–784.

Ni, L., Wang, D., Singh, V. P., Wu, J., Wang, Y., Tao, Y. & Zhang, J. 2020 Streamflow and rainfall forecasting by two long short-term memory-based models. *Journal of Hydrology* **583**, 124296.

Orouji, H., Bozorg Haddad, O., Fallah-Mehdipour, E. & Mariño, M. A. 2013 Modeling of water quality parameters using data-driven models. *Journal of Environmental Engineering* **139** (7), 947–957.

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S. & Mukhopadhyay, B. 2014 Development of a new high spatial resolution (0.25° × 0.25°) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam* **65** (1), 1–18.

Prakash, S. 2019 Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. *Journal of Hydrology* **571** (January), 50–59.

Prakash, S. & Gairola, R. M. 2014 Validation of TRMM-3B42 precipitation product over the tropical Indian Ocean using rain gauge data from the RAMA buoy array. *Theoretical and Applied Climatology* **115** (3–4), 451–460.

Prakash, S., Mitra, A. K., Momin, I. M., Pai, D. S., Rajagopal, E. N. & Basu, S. 2015 Comparison of TMPA-3B42 versions 6 and 7 precipitation products with gauge-based data over India for the southwest monsoon period. *Journal of Hydrometeorology* **16** (1), 346–362.

Prakash, S., Mitra, A. K., AghaKouchak, A., Liu, Z., Norouzi, H. & Pai, D. S. 2018 A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *Journal of Hydrology* **556** (February 2014), 865–876.

Rahman, S. H., Sengupta, D. & Ravichandran, M. 2009 Variability of Indian summer monsoon rainfall in daily data from gauge and satellite. *Journal of Geophysical Research Atmospheres* **114** (17), D17113.

Şahin, M. 2013 Comparison of modelling ANN and ELM to estimate solar radiation over Turkey using NOAA satellite data. *International Journal of Remote Sensing* **34** (21), 7508–7533.

Sahoo, B. B., Jha, R., Singh, A. & Kumar, D. 2019 Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. *Acta Geophysica* **67** (5), 1471–1481.

Setti, S., Maheswaran, R., Sridhar, V., Barik, K. K., Merz, B. & Agarwal, A. 2020 Inter-comparison of gauge-based gridded data, reanalysis and satellite precipitation product with an emphasis on hydrological modeling. *Atmosphere* **11** (11), 1–26.

Sit, M. A. & Demir, I. 2019 *Decentralized Flood Forecasting Using Deep Neural Networks*. arXiv.

Steyn, M., Wilms, J., Brink, W. & Smit, F. 2017 Short-term stream flow forecasting at Australian river sites using data-driven regression techniques. *International Conference on Time Series* 2017, 1–12.

Su, F., Hong, Y. & Lettenmaier, D. P. 2008 Evaluation of TRMM multisatellite precipitation analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *Journal of Hydrometeorology* **9** (4), 622–640.

Sulugodu, B. & Deka, P. C. 2019 Evaluating the performance of CHIRPS satellite rainfall data for streamflow forecasting. *Water Resources Management* **33** (11), 3913–3927.

Tan, M. L. & Santo, H. 2018 Comparison of GPM IMERG, TMPA 3B42 and PERSIANN-CDR satellite precipitation products over Malaysia. *Atmospheric Research* **202** (November), 63–76.

Tang, G., Ma, Y., Long, D., Zhong, L. & Hong, Y. 2016 Evaluation of GPM Day-1 IMERG and TMPA version-7 legacy products over Mainland China at multiple spatiotemporal scales. *Journal of Hydrology* **533**, 152–167.

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z. & Hong, Y. 2020 Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote Sensing of Environment* **240** (January), 111697.

Teegavarapu, R. S. V. & Chandramouli, V. 2005 Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. *Journal of Hydrology* **312** (1–4), 191–206.

Tiwari, A. & Vaibhaw, V. 2020 *Comparative Study on Time Series Forecasting using Deep learning Models*, March, 0–13.

Unduche, F., Tolossa, H., Senbeta, D. & Zhu, E. 2018 Evaluation of four hydrological models for operational flood forecasting in a Canadian Prairie watershed. *Hydrological Sciences Journal* **63** (8), 1133–1149.

Valipour, M. & Bateni, S. M. 2021 *Global Surface Temperature: A New Insight*, pp. 1–4.

Vieux, B. E., Vieux, J. E., Chen, C. & Howard, K. W. 2003 Operational deployment of a physics-based distributed rainfall-runoff model for flood forecasting in Taiwan. *IAHS-AISH Publication* (282), 251–257.

Vu, T. T., Li, L. & Jun, K. S. 2018 Evaluation of multi-satellite precipitation products for streamflow simulations: a case study for the Han River Basin in the Korean Peninsula, East Asia. *Water (Switzerland)* **10** (5), 642.

Wagena, M. B., Goering, D., Collick, A. S., Bock, E., Fuka, D. R., Buda, A. & Easton, Z. M. 2020 Comparison of short-term streamflow forecasting using stochastic time series, neural networks, process-based, and Bayesian models. *Environmental Modelling and Software* **126**, 104669.

Wei, G., Lü, H., Crow, W. T., Zhu, Y., Wang, J. & Su, J. 2018 Evaluation of satellite-based precipitation products from IMERG V04A and V03D, CMORPH and TMPA with gauged rainfall in three climatologic zones in China. *Remote Sensing* **10** (1), 30.

Wu, C. L. & Chau, K. W. 2010 Data-driven models for monthly streamflow time series prediction. *Engineering Applications of Artificial Intelligence* **23** (8), 1350–1367.

Wu, C. L., Chau, K. W. & Li, Y. S. 2009 Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. *Water Resources Research* **45** (8), 1–23.

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

Xu, R., Tian, F., Yang, L., Hu, H., Lu, H. & Hou, A. 2017 Ground validation of GPM IMERG and trmm 3B42V7 rainfall products over Southern Tibetan plateau based on a high-density rain gauge network. *Journal of Geophysical Research* **122** (2), 910–924.

Yadav, A. K., Chandola, V. K., Singh, A. & Singh, B. P. 2020 Rainfall-runoff modelling using artificial neural networks (ANNs) model. *International Journal of Current Microbiology and Applied Sciences* **9** (3), 127–135.

Yaghoubi, B., Hosseini, S. A. & Nazif, S. 2019 Monthly prediction of streamflow using data-driven models. *Journal of Earth System Science* **128** (6), 1–15.

Yaseen, Z. M., El-shafie, A., Jaafar, O., Afan, H. A. & Sayl, K. N. 2015 Artificial intelligence based models for stream-flow forecasting: 2000–2015. *Journal of Hydrology* **530**, 829–844.

Yaseen, Z. M., Jaafar, O., Deo, R. C., Kisi, O., Adamowski, J., Quilty, J. & El-Shafie, A. 2016 Stream-flow forecasting using extreme learning machines: a case study in a semi-arid region in Iraq. *Journal of Hydrology* **542** (February 2018), 603–614.

Yaseen, Z. M., Allawi, M. F., Yousif, A. A., Jaafar, O., Hamzah, F. M. & El-Shafie, A. 2018 Non-tuned machine learning approach for hydrological time series forecasting. *Neural Computing and Applications* **30** (5), 1479–1491.

Yeditha, P. K., Venkatesh, K., Rathinasamy, M. & Agarwal, A. 2020 Forecasting of extreme flood events using different satellite precipitation products and wavelet-based machine learning methods. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30** (6), 063115.

Zhang, X. & Tang, Q. 2015 Combining satellite precipitation and long-term ground observations for hydrological monitoring in China. *Journal of Geophysical Research: Atmospheres* **120** (13), 6426–6443.

Zhou, Y. & Ushiama, T. 2019 *LSTM-Based Recommendation Approach for Interaction Records*, pp. 950–962.