Global NEST Journal, Vol 27, No X, pp 1-11

Weather forecasting model using attentive residual gated recurrent unit for urban flood prediction

Vivek Balaji K.[1], Sugumar R.[1], Mahendran R.[2] and Subramanian P.[3,*]

[1]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai 602105, Tamil Nadu, India.

[2]Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur 522502, Andhra Pradesh, India

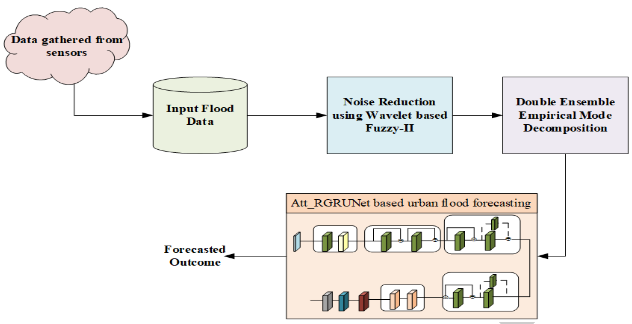
[3]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, 602105, India

Received: 10/11/2024, Accepted: 21/03/2025, Available online: 26/03/2025

*to whom all correspondence should be addressed: e-mail: subramanianp.sse@saveetha.com

https://doi.org/10.30955/gnj.06985

Graphical abstract



Abstract

Flood forecasting is significant for hydrology and disaster management due to the complex and nonlinear nature of flood-related data. There are several conventional forecasting methods often failed to effectively capture both spatial and temporal dependencies that leads to inaccurate predictions. In order to overcome the challenges, a novel method is introduced with data de-noising and deep learning models for flood prediction. The proposed approach comprises of three key steps, wherein de-noising of flood data is employed initially using wavelet transform with Fuzzy II threshold selection to eliminate noise in capturing significant features. Then, Intrinsic Mode Function (IMF) extraction is employed using Double Ensemble Empirical Mode Decomposition (DEEMD) to acquire appropriate flood patterns. Finally, flood forecasting is employed using an Attentive Residual Gated Recurrent Unit (Att_RGRUNet) model, wherein ResNet is utilized extracts spatial features, GRU model is utilized for temporal dependencies, and coordinate attention mechanism for enhancing the feature representation. The combined approach ensures high predictive accuracy and enhances early warning systems. The proposed model is evaluated based on RMSE, MAE and MAPE and acquired the values of 0.76984, 0.8 and 2.94 respectively.

Keywords: flood forecasting, attention mechanism, gated recurrent unit, double ensemble EMD, spatio-temporal feature extraction

1. Introduction

Weather forecasting is devised for predicting atmospheric conditions and mitigating risks related to extreme weather conditions (Anuradha *et al*. 2024). It is essential in safeguarding lives, infrastructure, and economic activities by providing timely warnings about storms, heavy rainfall, heat-waves, and other meteorological phenomena. Precise weather prediction is utilized by governments, businesses, and individuals to eliminate hazards caused due to extreme weather conditions by reducing the impact of natural disasters (Akinyoola *et al*. 2024; Qin *et al*. 2024). In addition, weather forecasting is utilized in several domains like agriculture, aviation, and transportation to detect the weather variations that significantly affect operations. Also, the development in meteorological science, remote sensing, and artificial intelligence have improved forecasting accuracy for enhancing disaster awareness and climate resilience. Urban flood forecasting is essential for weather forecasting due to the increasing frequency and intensity of extreme rainfall events in cities (Albahri *et al*. 2024). Rapid urbanization, poor drainage systems and climate change contribute to severe flood that causes disruptions in transportation, waterlogging, and damage to buildings. Real-time flood forecasting is efficient in implementing effective disaster response strategies for reducing casualties and economic losses (Yang *et al*. 2024; Zhou *et al*. 2024).

Traditional urban flood forecasting methods depends on hydrological and hydraulic models like Rational Method (Fernández-Nóvoa *et al*. 2024), the Soil Conservation Service (SCS) Curve Number Model (Aja *et al*. 2020), and the Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) (Verma *et al*. 2024). These models utilize historical rainfall data, topographical maps, and drainage system characteristics to predict flood levels and water flow patterns. Still, the traditional methods have been widely used, they face significant challenges in urban settings (Khan *et al*. 2024; Peker *et al*. 2024). The major limitation is the inability of the model to accurately model rapid urbanization and land-use changes, which significantly alter surface runoff patterns (El-Bagoury and Gad, 2024). Also, the traditional models struggle with real-time data processing that leads to delays in flood prediction. The requirement of extensive calibration and large datasets makes the traditional methods as a computationally expensive and less adaptable to dynamic urban environments (Zheng *et al*. 2024; Suwarno *et al*. 2021). Besides, traditional models are highly sensitive to parameter uncertainties that result in inaccurate flood forecasts during risky weather events.

Machine learning (ML) has been designed to overcome the limitations of traditional flood forecasting methods. ML models like Artificial Neural Networks (ANNs) (Doan and Le-Thi, 2025), Support Vector Machines (SVMs) (Sengupta, 2025), and Random Forests (Vamsi and Amudha, 2024) approaches are utilized to analyze large and complex datasets in real time by identifying nonlinear relationships between features. The ML models continuously learn and improve over time by making the model adaptive to changing environmental conditions. However, ML technique results with improved accuracy in forecasting; still, feature extraction and long-term dependencies in flood prediction was challenging task (Ren *et al*. 2024). To address this, hybrid deep learning models like Convolutional Neural Networks (CNNs) integrated with Long Short-Term Memory (LSTM) networks for providing an efficient solution for real-time urban flood forecasting. The hybrid approach enhances prediction accuracy by integrating spatial and temporal information for more precise and timely flood alerts (Wahba *et al*. 2024). Additionally, the deep learning models process the real-time data gathered from Internet of Things (IoT) sensors, weather stations, and remote sensing technologies for providing highly efficient in operational flood forecasting (Kumshe *et al*. 2024). The objectives of the flood forecasting model are:

To enhance the flood forecasting accuracy by utilizing wavelet transforms for signal decomposition and noise reduction.

To design a DEEMD model for extracting relevant IMFs from hydrological time-series data.

To design an Att_RGRUNet model, wherein ResNet is utilized extracts spatial features, GRU model is utilized for temporal dependencies, and coordinates attention mechanism for enhancing the feature representation.

To compare the proposed model with the existing urban flood forecasting model to demonstrate the superiority of the proposed model.

The research is organized as: Section 2 details the related works with the problem statement and the detailed proposed flood forecasting model in Section 3. The experimental results are presented in Section 4 and the conclusion in Section 5.

2. Related works

The author (Pokharel and Roy, 2024) proposed a multimodule based approach for explainable DL-based monthly rainfall prediction. The study composed of four different modules such as attention mechanism, GRU-based decoder and encoder modules, and expected-gradient module. Initially, climate and weather features were given as input to the encoder to produce associated hidden states. Next, based on the decoder and encoder hidden states, a series of attention values were generated by the attention mechanism module. In the third step, each attention values were integrated with the rainfall predicted value in the decoder module at the final time stamp. Then, the obtained values were provided to the GRU cell to generate recent hidden state. This hidden state was passed through fully connected layer of the model to achieve associated rainfall prediction value. Finally, the expected gradient was used to quantify the significance of input feature to the output feature. The introduced model struggle in capturing long-term dependencies.

To accurately predict the ponding depth at specific urban flood points, (Pokharel and Roy, 2024) designed a Gradient Boosting Decision Tree (GBDT) model. In this, the rainfall data and ponding depth data were gathered from specific urban locations and Moran's I technique was employed for spatial analysis. Moran's I was employed to confirm the spatial independence of ponding points. The designed model portrayed the accuracy of the early warning system increases as the prediction period decreases. The model's accuracy relies heavily on the quality and quantity of rainfall and ponding depth data.

To create a model capable of predicting urban flood water levels in real-time was designed by (Zhou *et al*. 2024) using the Artificial Neural Network connected with Multilayer Architectures. In this, ANN was employed to capture spatial patterns in flood inundation and MLP was employed to process temporal relationships and integrate information. The designed model achieved promising prediction performance with low computational time that enables real-time applications. Balancing prediction accuracy with computational efficiency for real-time applications was challenging aspect.

Hybrid Deep Learning Model with Graph Convolutional Network (GCN) to capture spatial dependencies and Long Short-Term Memory (LSTM) to model temporal dependencies and patterns in human mobility flows was designed by (Berkhahn and Neuweiler, 2024). In this, the feature aggregation was employed to maintain the spatial heterogeneity for planning and managing emergency responses during urban floods. Enhanced prediction accuracy was employed for disaster management. Still, the designed model was heavily reliant on the quality and accuracy of the graph's adjacency matrix, which degrades the performance.

Machine learning model-based Flood Susceptibility Prediction (FSP) was designed by (Tang *et al*. 2024) to enhance the accuracy of FSP models by optimizing the representation and selection of flood conditioning factors (FCFs). In this, the feature selection was employed using collinearity and mean decrease impurity technique for reducing the computation burden. The designed model demonstrated that using extreme rainfall indices, enhanced accuracy was accomplished. Still, for processing large datasets, machine learning models was computationally intensive and makes over-fitting issues.

Unsupervised change detection (CD) Network was designed by (Asfaw *et al*. 2024) for flood extent detection using Spatiotemporal Variational Autoencoder (VAE). The designed model employed contrastive learning techniques to improve the discriminative power of the learned features. The designed model demonstrated better generalizability on unseen sites compared to supervised models. Still, the designed model was inefficient to handle uncertainties in reference flood maps.

Urban flooding is essential for rapid urbanization, climate change, and inadequate drainage systems (Yadav *et al*. 2024). The unpredictability of extreme weather conditions like heavy rainfall cause significant risks to lives, infrastructure, and the economy (Karthik *et al*. 2025). Traditional flood management approaches often fail to provide timely and accurate predictions that leads to severe consequences (Babu *et al*. 2024). Urban flood forecasting is designed to design appropriate models that integrate real-time hydrological, meteorological, and topographical data to predict flood occurrences and intensities (Sundarapandi *et al*. 2024). Using the conventional machine learning, remote sensing, and Geographic Information Systems, accurate flood predictions is employed to enhance early warning systems, aid emergency response, and minimize damage. Thus, effective urban flood forecasting is essential for sustainable urban planning and resilience against climate-induced disasters. Hence, a novel deep learning model with data pre-processing technique is introduced in this research.

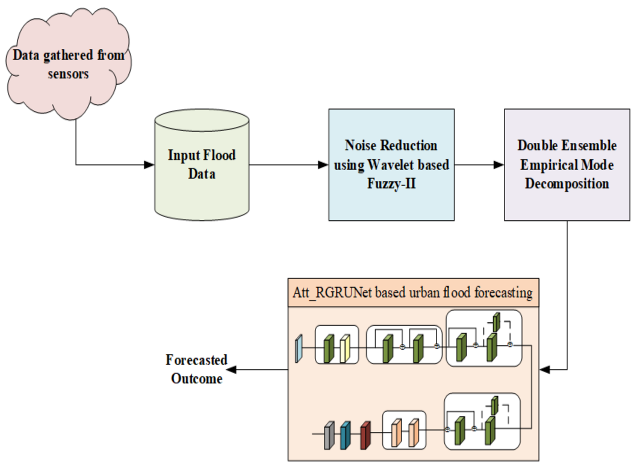


Figure 1. Proposed urban flood forecasting model

3. The proposed model

The proposed methodology of urban flood forecasting model is designed by integrating data processing and deep learning techniques. Initially, input flood data with various attributes like rainfall, temperature, and humidity is applied to the de-noising approach using wavelet transform with Fuzzy II threshold selection to remove noise and retain critical features. Then, Intrinsic Mode Functions (IMFs) are extracted using Double Ensemble Empirical Mode Decomposition (DEEMD), which enhances the signal decomposition process and extracts significant flood-related components. Finally, an Attentive Residual Gated Recurrent Unit (Att_RGRUNet) is employed for flood forecasting, wherein ResNet extracts the spatial features, GRU captures temporal dependencies, and a coordinate attention mechanism enhances feature representation. A fully connected layer is then used for final prediction, which is portrayed in **Figure 1**.

3.1. Data gathering

The input data for the urban flood forecasting is acquired from the publically available dataset (Flood Forecasting Dataset).

3.2. Noise reduction using wavelet based fuzzy-II

Wavelet-based Type II fuzzy de-noising offers a highly effective approach for processing noisy time-series data. The wavelet transform (WT) enables multi-scale analysis for precise localization of both high and low-frequency components, which is helpful in separating required and noise data form the input flood forecasting data. Meanwhile, Type II fuzzy sets introduce an additional layer of uncertainty modeling by defining upper and lower membership functions for obtaining robust noise filtering process. The first step in de-noising the flood prediction data is to transform the input data into the wavelet domain. It is done using the Wavelet Transform (WT), which decomposes the signal into wavelet coefficients at different scales and shifts. The decomposition of data is expressed as:

$$K = \frac{\sum_{x=1}^{n} (y_x \bar{y})(v_x - \bar{v})}{\sqrt{\sum_{x=1}^{n} (y_x - \bar{y})^2 (v_x - \bar{v})^2}} d_{m,n} = \sum_u A_r(u)\varphi_{m,n}(u) \tag{1}$$

Here, the scale and shift factor concerning the wavelet transform is notated as m, n respectively and the function that notates the wavelet is indicated as ϕ. The input data is symbolized as $A_r(u)$ and the wavelet coefficient is indicated as $d_{m,\,n}$. Then, the threshold selection is employed to identify the wavelet coefficients concerning the noise and required data. Normally, noise data is assigned with small wavelet coefficients and useful data is assigned with large wavelet coefficients. Here, the threshold value is represented as β, wherein the wavelet coefficients below β are considered noise and set to zero and others are retained. Once the threshold is determined, it is applied to the wavelet coefficients using either hard thresholding or soft thresholding for de-noising the data. Here, the consideration of fixed threshold introduce biased outcome. Thus, Fuzzy-II based threshold selection is introduced to overcome the challenge faced by the fixed threshold based hard or soft thresholding criteria. Here, Type II fuzzy sets introduce an additional layer of uncertainty by defining an upper and lower membership function, allowing us to better handle imprecise or ambiguous data. The Type II fuzzy index for an iterative threshold u^* is expressed as:

$$\beta\left(u^*\right)=\frac{1}{R}\sum_{p=0}^{E-1} I(p)\times\left[K(p)-G(p)\right] \tag{2}$$

where, R is the total number of elements in the set, E is the length of the data, $I(p)$ is the weighting function at index p, $K(p)$ is the upper membership function and $G(p)$ is the lower membership function. Here, upper and lower membership functions are defined as:

$$K(p)=\frac{1}{\propto_p}, \qquad G(p)=\propto_p \tag{3}$$

where $\propto$ is a positive parameter that determines how much uncertainty is present in the data For the flood forecasting model, the range of $\propto$ is generally set within [0, 2]. The Gaussian function is used to model uncertainty in the fuzzy set:

$$p=\exp\left(-\frac{\left|p-\max Q\right|^2}{s_l^2}\right) \tag{4}$$

where, max Q is the maximum value of the coefficients, N_i is the average level of the coefficients and s_l^2 is the bandwidth of the Gaussian membership function. The bandwidth of the Gaussian function is set to allow useful signal information while attenuating high-frequency noise:

$$s_l=\max\left(u^*-N_i,\max Q-u^*\right) \tag{5}$$

Using the index of vagueness, the optimal threshold is calculated as:

$$opt=\arg\max \beta\left(u^*\right) \tag{6}$$

where arg max finds the threshold value that maximizes the Type II fuzzy index, ensuring the best separation between noise and useful data. Thus, Type II fuzzy set-based de-noising offers a powerful method for processing noisy time-series data in flood detection. By using fuzzy membership functions, Gaussian thresholds, and wavelet transformation, this approach effectively separates signal from noise and improves the accuracy of flood prediction systems.

3.3. Double ensemble empirical mode decomposition

EMD is a data-driven method for analyzing nonlinear and non-stationary time series. It decomposes a complex signal into multiple Intrinsic Mode Functions (IMFs) that represent different frequency components of the original data. Unlike traditional decomposition methods, EMD does not use pre-defined basis functions but derives them adaptively from the data itself. It makes EMD particularly useful in hydrology and flood forecasting, where water levels and rainfall data exhibit complex, fluctuating behaviors. The IMF feature extraction process is portrayed in **Figure 2**.

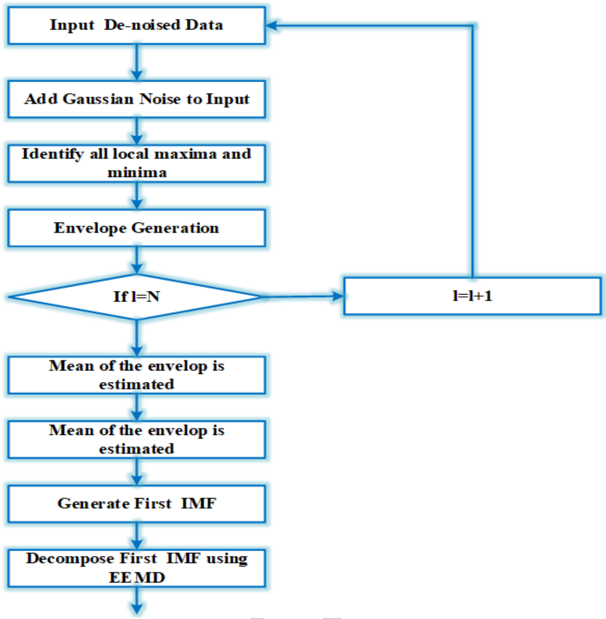


Figure 2. IMF extraction using DEEMD model

Consider a time series input $u(m)$, where m = 1, 2,…g. The EMD method follows these steps:

(i) Identify all local maxima and minima in $u(m)$.

(ii) Envelope Generation: The local maxima are connected using cubic spline interpolation to form the upper envelope $LMa(m)$. Also, the local minima are connected similarly to form the lower envelope $LMi(m)$.

(iii) Mean of the envelop is estimated using:

$$n(m)=\frac{LMa(m)+LMi(m)}{2} \tag{7}$$

(iv) The first IMF candidate is extracted using

$$p(m)=u(m)-n(m) \tag{8}$$

(v) The conditions for extracting the IMF are:

- The number of extrema both maxima and minima should be the same or differ by at most one.
- The mean envelope should be zero at all points.
- If $p(m)$ satisfies these conditions, it is accepted as the first IMF $k_1(m)$.
- Otherwise, replace $u(m)$ with $p(m)$ and repeat steps 1-4 until a valid IMF is extracted.

(vi) Once an IMF is extracted, compute the residual outcome as:

$$q(m)=u(m)-k_1(m) \tag{9}$$

(vii) Continue the IMF extraction from $q(m)$ until the residual becomes a monotonic function.

(viii) The original data is decomposed as:

$$u(m)=\sum_{l=1}^{b} k_l(m)+q_k(m) \tag{10}$$

where $k_l(m)$ are the IMFs and $q_k(m)$ is the final residual representation of the features. Here, the challenging aspect is the IMF with different scales or data with same scale for representing multiple IMFs. Thus, a small noise component is added with the data prior to performing the EMD process, which is termed as Ensemble Empirical Mode Decomposition (EEMD). Here, the addition of Gaussian white noise to the data reduces mode mixing and improves the separation of frequency components. Initially, the ensemble size represented as N that symbolizes the number of noise realizations and the noise amplitude are assigned. Then, generate a white noise sequence $k_n(m)$ and add it to the original input data. It is expressed as:

$$u_n(m) = u(m) + k_n(m) \tag{11}$$

Decompose $u_n(m)$ using EMD and extract IMFs and perform the decomposition for multiple iterations, each time using a different realization of white noise. The final IMFs are obtained by averaging over all decompositions:

$$d_k = \frac{1}{N}\sum_{n=1}^{N} IMF_{k,n} \tag{12}$$

where $IMF_{k,n}$ represents the k^{th} IMF in the n^{th} realization. Here, addition of noise helps improve IMF separation and the extracted IMFs are less likely to contain overlapping frequency components. EEMD provides more accurate hydrological data analysis for enhancing the reliability of flood prediction model. Double Ensemble Empirical Mode Decomposition (DEEMD) is a refinement of EEMD that specifically targets the first IMF, which contains the highest frequency components. It is crucial in flood forecasting, where high-frequency variations correspond to rapid changes in water levels. The steps considered in the DEEMD are:

- The first IMF is extracted using the EEMD process.
- Decompose the first IMF again using EEMD to further separate high-frequency time series data.
- Extract the refined IMFs to acquire more accurate representation of short-term variations.

The refined IMFs are further utilized for performing the urban flood forecasting model using the proposed Att_RGRUNet model.

3.4. Proposed att_rgrunet urban flood forecasting model

The urban flood forecasting is devised using the proposed Attentive Residual Gated Recurrent Unit Network (Att_RGRUNet) for extracting the spatial and temporal attributes. The designed Att_RGRUNet model is efficient in learning the spatial features from IMF data through the residual connections. Then, the temporal dependencies among the features are acquired through the gated recurrent unit (GRU). In addition, the consideration of co-ordinate attention module is employed to assign higher weights to critical flood factors that lead to precise forecasting. The structure of Att_RGRUNet is presented in **Figure 3**.

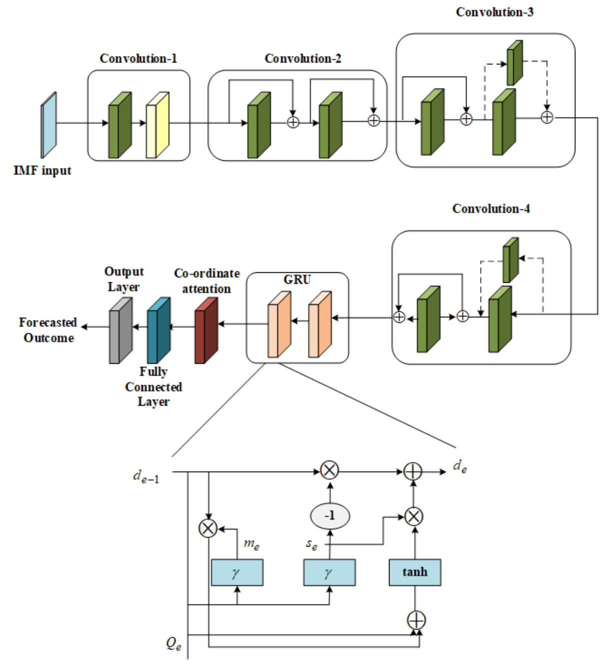


Figure 3. Structure of Att_RGRUNet Model

The input acquired by the proposed Att_RGRUNet model is the IMF features that are shared through various convolution layers with skip connections for capturing the spatial features. Conventional neural networks have the problem of degradation issue with the increase in depth size. When gradients shrink too much during backpropagation for making the earlier layers learn very slowly that leads to vanishing gradient issue. When gradient become excessively large that leads to instability in training termed exploding gradients. These two issue are effectively solved through the network with residual (skip) connections. The residual block is the core component of ResNet. It allows a neural network to learn a residual function instead of trying to learn the entire transformation directly. Let, D be the input to the Att_RGRUNet model and $N(p)$ be the actual mapping the network needs to learn. $K(p)$ is the residual function that the network will learn. Instead of learning $N(p)$ directly, the residual block learns:

$$N(p) = K(p) + D \tag{13}$$

where, $K(p)$is the transformation applied to D and D is directly added to the output through a skip connection. Since the identity mapping is passed forward, the network only needs to learn the difference $K(p) = N(p) - D$. In this, gradients flow directly through the shortcut connections that prevent the model from shrinking too much. Also, when the additional layers learn nothing $K(p) = 0$, the identity mapping ensures that performance does not degrade. The skip connection helps to maintain stable gradient values for avoiding exploding gradients. Batch Normalization (BN) is applied to stabilize the training by normalizing the activations in a neural network. It is used in ResNet to speed up convergence and maintain a stable distribution of activations. The features extracted through the various convolutional layers with skip connections are fed into the GRU module for capturing the temporal features. In the GRU module, the consideration of dropout layer assists in minimizing the over-fitting issues of Att_RGRUNet model. At each time step e, GRU updates its hidden state using the following formulas:

The update gate m_e determines how much of the previous hidden state d_{e-1} should be carried forward.

$$m_e = \gamma\left(G_m Q_e + N_m d_{e-1} + l_m\right) \tag{14}$$

where, m_e is the update gate at e, Q_e is the input data at e, d_{e-1} is the previous hidden state, G_m, N_m is the weight matrices, l_m is the bias vector, and γ is the sigmoid activation function. The reset gate s_e decides how much of the previous hidden state should be forgotten.

$$s_e = \gamma\left(G_s Q_e + N_s d_{e-1} + l_s\right) \tag{15}$$

A candidate hidden state $\hat{d}_e$ is computed using the reset gate for selective memory retention.

$$\hat{d}_e = \tanh\left(G_d Q_e + s_e \Theta N_d d_{e-1} + l_d\right) \tag{16}$$

The hidden state is helpful to generate an intermediate flood prediction feature based on both past and current information. The final hidden state d_e is computed as a combination of the previous hidden state and the candidate state.

$$d_e = (1-m_e)\Theta d_{e-1} + m_e \Theta \hat{d}_e \tag{17}$$

Here, $(1-m_e)$ controls the influence of past states, and $m_e\Theta\hat{d}_e$ incorporates new flood-related information. Then, the co-ordinate attention mechanism is employed for capturing the significant features that enhances the forecasting accuracy. The co-ordinate attention mechanism is employed for capturing both spatial and temporal dependencies among the features by using both horizontal and vertical global pooling operations. The structure of co-ordinate attention mechanism is presented in **Figure 4**.

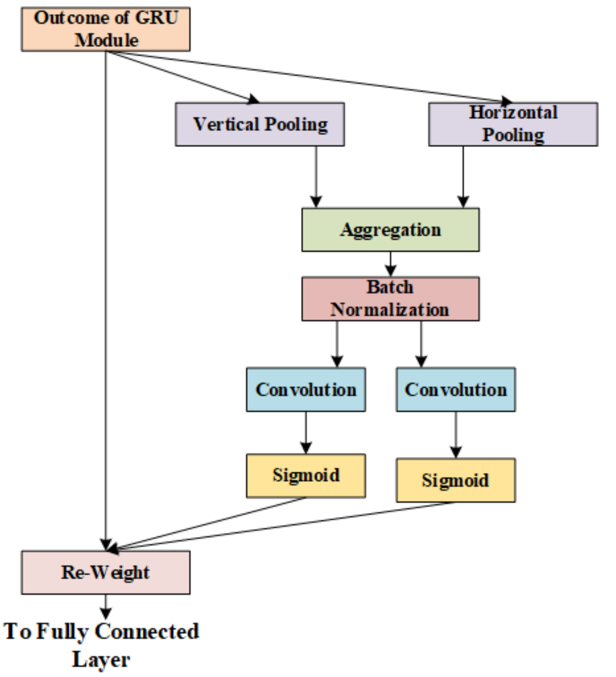


Figure 4. Structure of co-ordinate attention mechanism

Using the outcome of the GRU module, the vertical global pooling operation is performed by the co-ordinate attention mechanism and is formulated as:

$$b_f^m(m) = \frac{1}{O}\sum_{u=1}^{O} q_f(m,u) \tag{18}$$

where, $b_f^m(\mathrm{m})$ is the height-wise feature representation, $q_f(m, u)$ is the input feature at channel f, height m, and width $u$and O is the width of the input feature map.

The horizontal global pooling operation is formulated as:

$$b_f^o(o) = \frac{1}{M}\sum_{v=1}^{M} q_f(o,u) \tag{19}$$

where, $b_f^o(\mathrm{o})$ is the width-wise feature representation, $q_f(o, u)$ is the input feature at channel f, height v, and width o and M is the height of the input feature map. Here, the consideration of both transformations allows the model to capture global dependencies along one direction while retaining precise spatial information in the other direction. The aggregation of features extracted through both the horizontal and vertical pooling operations is defined as:

$$w = \mathrm{Re}LU\left(W_1\left(\left[b_m, b_o\right]\right)\right) \tag{20}$$

where, W_1 is the shared 1×1 convolution, $[b_m, b_o]$ is the concatenated height and width feature maps and ReLU is the activation function. After aggregating the features, the attention weights are estimated for both horizontal and vertical features. It is formulated as:

$$x_f^m = \lambda\left(W_m\left(w_m\right)\right) \tag{21}$$

$$x_f^o = \lambda\left(W_o\left(w_o\right)\right) \tag{22}$$

Here, x_f^m, x_f^o is the attention weights for height and width, λ is the sigmoid activation function, and W_m, W_o is the convolution operations for height and width. The features acquired at the output of the co-ordinate attention mechanism $z_f(u, v)$ are expressed as:

$$z_f(u,v) = q_f(u,v)\times x_f^m(u)\times x_f^o(v) \tag{23}$$

The extracted GRU features are passed through fully connected termed dense layers to produce the final flood prediction:

$$UF = G_o z_f + l_o \tag{25}$$

where, UF is the predicted flood value, G_o is the weight matrix for output layer, l_o is the bias and z_f is the final outcome of the co-ordinate attention mechanism at the last time step. The overall step by step algorithm of the proposed urban flood forecasting model is presented in Algorithm 1.

4. Results and discussion

The proposed urban flood forecasting method is implemented in PYTHON programming tool and is assessed based on various measures. The proposed Att_RGRUNet model is compared with existing methods like GCN-LSTM [23], CD network [25], ANN_MLP [22], and GBDT [21] to demonstrate the superiority. For the evaluation of proposed urban flood forecasting model, the publically available [26] dataset is utilized.

Algorithm 1: Pseudo-code for proposed flood forecasting model

```
Pseudo-code for proposed flood forecasting model
Start
Step 1: Data Acquisition
Input flood data from publically available dataset
Step 2: Pre-processing (De-Noising)
Apply discrete WT on input data to decompose into sub-bands
For each sub-band
{
  Compute Fuzzy II threshold using entropy-based selection
  Apply soft thresholding on coefficients to remove noise
}
 End For
Apply Inverse Wavelet Transform (IWT) to reconstruct de-noised
Return de-noised data
Step 3: IMF extraction using DEEMD
Add Gaussian Noise for Ensemble Approach
Generate Envelopes
Iterative Decomposition Process
Apply DEEMD
Output Intrinsic Mode Functions (IMFs) and Residues
Return IMF_features
Step 4: Att_RGRUNet based urban flood forecasting model
# Spatial Feature Extraction using ResNet
Initialize ResNet layers for spatial feature extraction
Extract spatial features
# Temporal Feature Extraction using GRU
Initialize GRU layers for sequential learning
Extract temporal features
# Coordinate Attention Mechanism
Compute horizontal and vertical attention weights
Aggregate Features
# Fully Connected Layer for Final Prediction
 Apply Fully Connected Layer on enhanced features z_f acquired by co-ordinate attention mechanism
Output Forecasted Flood Levels UF
Stop
```

Dataset Description: Dataset from https://www.kaggle.com/datasets/s3programmer/flood-risk-in-india?select=flood_risk_dataset_india.csv is utilized for evaluating the proposed urban flood detection model. Here, 80% of the data from the dataset is employed for training the flood detection model and remaining 20% of data is employed for forecasting the urban flood. The description of the dataset is presented in **Table 1**.

Table 1. Dataset Description

Parameters	Values
Total Records	1000
Attributes	14
Location	India
Time Period	2014 to 2020
File Format	Csv
Data Size	1.8MB

The dataset distribution by considering three various factors like temperature, humidity and rainfall are presented in **Figure 5**.

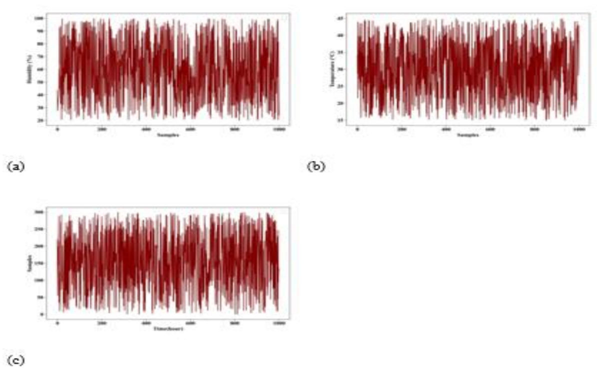


Figure 5. Dataset Distribution: (a) humidity, (b) temperature and (c) rainfall

Assessment Measures: The proposed urban flood forecasting model is assessed using three various measures like Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). The formulation for the assessment measures are interpreted as:

$$MAPE = \frac{1}{D}\sum_{l}\left|\frac{A_l - \hat{A}_l}{A_l}\right| \tag{26}$$

$$RMSE = \sqrt{\frac{1}{D}\sum_{l}\left(A_l - \hat{A}_l\right)^2} \tag{27}$$

$$MAE = \frac{1}{D}\sum_{l}\left|A_l - \hat{A}_l\right| \tag{28}$$

where, the total number of samples is symbolized as D, the actual outcome is interpreted as A_l and the predicted outcome is notated as $\hat{A}_l$.

Implementation Details: The proposed model has been simulated using the Python programming language. The simulation uses the L1 loss as the information loss term and the L2 norm as the punishment term. To balance the information loss and punishment terms, the hyperparameter should be taken into account before the penalty term. The Att_RGRUNet model is optimized using L1 and L2 losses. An ADAM optimizer is used for training, with an initial learning rate of 0.001 and a weight decay of 0.025. The training process is completed after 100 iterations with batch size for each iterations is eight. All experiments are constructed with PyTorch 1.2.0 and Python 3.6 and run on the Nvidia Titan RTX GPUs.

4.1. Analysis of proposed urban flood forecasting model

The urban flood forecasting by considering the rainfall factor is presented in **Figure 6**, wherein the outcome arrived for RSRUNet, Att_RGRUNet, Att_RGRUNet with DEEMD, and proposed method is illustrated. The proposed method illustrated the close relationship between the actual (green line) and predicted (red line) outcome. Thus, the RMSE (brown line) estimated by the proposed method is minimal compared to other methods.

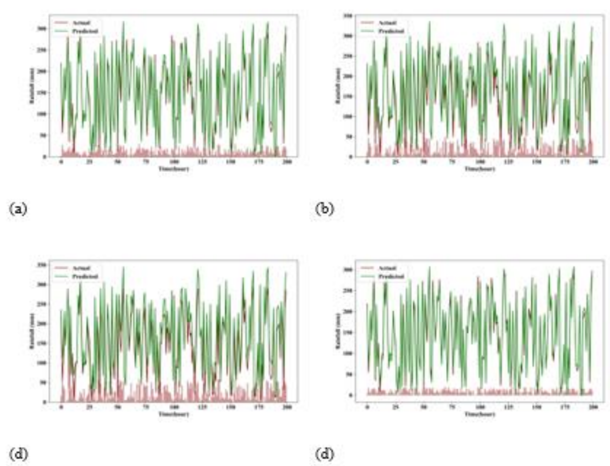


Figure 6. Forecasting outcome for Rainfall: (a) RSRUNet, (b) Att_RGRUNet with DEEMD, (c) Att_RGRUNet and (d) Proposed (Denoising+DEEMD+ Att_RGRUNet)

The urban flood forecasting by considering the humidity factor is presented in **Figure 7**, wherein the outcome arrived for RSRUNet, Att_RGRUNet, Att_RGRUNet with DEEMD, and proposed method is illustrated.

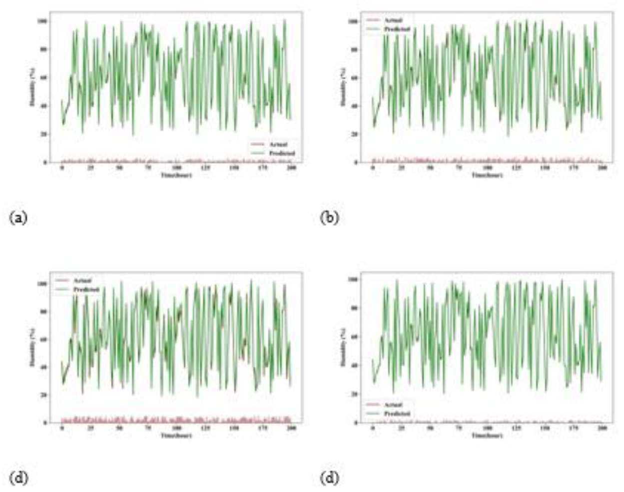


Figure 7. Forecasting outcome for Humidity: (a) RSRUNet, (b) Att_RGRUNet with DEEMD, (c) Att_RGRUNet and (d) Proposed (Denoising+DEEMD+ Att_RGRUNet)

The urban flood forecasting by considering the temperature factor is presented in **Figure 8**, wherein the outcome arrived for RSRUNet, Att_RGRUNet, Att_RGRUNet with DEEMD, and proposed method is illustrated.

The ablation study of the proposed model in terms of RMSE, MAPE and MAE for the parameters temperature, humidity and rainfall is presented in **Table 2**. The proposed method employed Wavelet transform with Fuzzy II based threshold selection removes the noise effectively. Then, DEEMD model assist to separate the critical flood-related IMFs for enhancing forecasting accuracy. Finally, Att_RGRUNet captures both spatial and temporal dependencies efficiently. Thus, the enhanced outcome is derived by the proposed model. The proposed model without de-noising capability (Att_RGRUNet with DEEMD) comprises of noisy component that leads to biased outcome in forecasting. Thus, poor performance compared to proposed method is acquired by Att_RGRUNet with DEEMD. Then, the proposed model without de-noising and DEEMD (Att_RGRUNet) provides the degraded outcome due to the failure in removing the artifacts and feature extraction. Similarly, the model designed with only the spatio-temporal feature extraction model RSRUNet accomplished degraded outcome compared to all the other methods. Thus, the combined effect of all the techniques assists the proposed model to acquire enhanced outcome with minimal error.

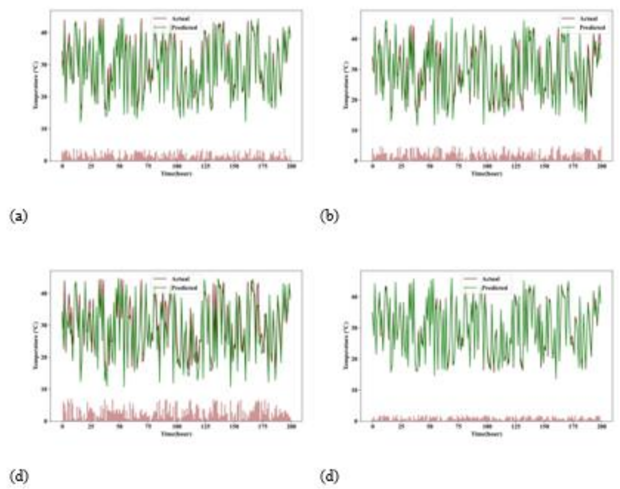


Figure 8. Forecasting outcome for Temperature: (a) RSRUNet, (b) Att_RGRUNet with DEEMD, (c) Att_RGRUNet and (d) Proposed (Denoising+DEEMD+ Att_RGRUNet)

Table 2. Ablation Study

Parameters	Metrics	RSRUNet	Att_RGRUNet	Att_RGRUNet with DEEMD	Proposed (DEEMD+ Att_RGRUNet)
Temperature	MAPE (%)	4.510	3.920	3.280	2.940
	RMSE	2.490	1.860	1.350	0.836
	MAE	2.170	1.580	1.030	0.800
Humidity	MAPE (%)	10.430	9.520	8.450	7.830
	RMSE	8.360	7.850	7.160	6.463
	MAE	4.080	3.570	2.860	2.350
Rainfall	MAPE (%)	16.810	16.460	15.950	15.310
	RMSE	1.927	1.527	0.911	0.770
	MAE	2.410	1.870	1.570	1.200
Average	MAPE (%)	10.583	9.967	9.227	8.693
	RMSE	4.259	3.746	3.140	2.690
	MAE	2.887	2.340	1.820	1.450

The urban flood forecasting for various time steps is presented in **Figure 9**. Here, for the analysis, the proposed urban flood forecasting model is compared with existing approaches like GCN-LSTM, CD network, ANN_MLP, and GBDT. The conventional GCN-LSTM struggles with high spatial-temporal dependency complexity that leads to larger absolute errors when spatial relationships change dynamically. While considering CD network failed to capture nonlinear dependencies in flood forecasting effectively that leads to increased errors. In addition, ANN-MLP model lacks sequential memory capabilities and hence it less efficient for time-series flood prediction. The GBDT approach is effective for structured data but does not perform well with complex sequential data. The proposed model with both spatial and temporal features and attention mechanism model selectively focus on the most influential flood-related features. Thus, minimal MAE is estimated by the proposed model for all the three parameters.

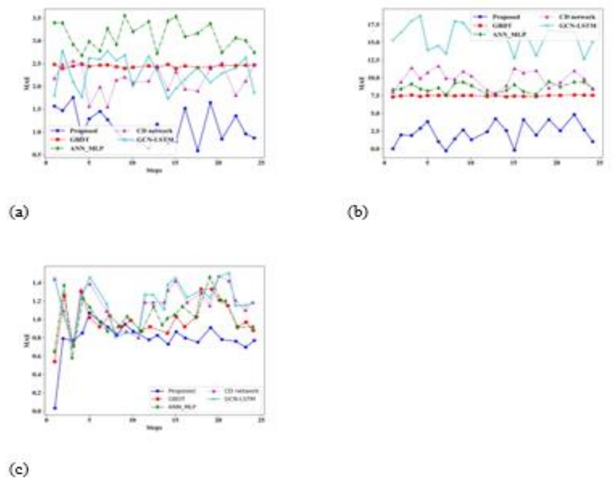


Figure 9. MAE Assessment: (a) rainfall, (b) humidity and (c) temperature

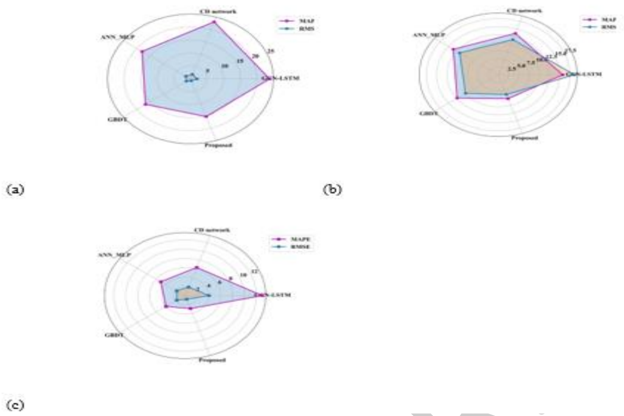


Figure 10. RMSE and MAPE Assessment: (a) rainfall, (b) humidity and (c) temperature

The assessment of MAPE along with the RMSE is presented in **Figure 10** for the parameters like temperature, humidity and rainfall. Here, the incapability of the GCN-LSTM to extract the local flood changes and degraded outcome in handling sequential data by the GBDT leads to poor performance. The inefficiency in sequential learning capabilities of ANN-MLP and failure in gathering nonlinear dependencies among the features by the CD network makes the models to enhance the error in forecasting. The proposed model with efficient de-noising, feature extraction and spatio-temporal feature extraction with attention mechanism assist the model to minimize the error.

The comparative discussion of the urban flood forecasting methods is presented in **Table 3**. The existing urban flood forecasting methods like GCN-LSTM, CD network, ANN_MLP, and GBDT are compared with proposed method. The existing GBDT approach Moran's I feature extraction and the forecasting using GBDT model, which was incapable in handling the large dataset that makes the over-fitting issues. The ANN_MLP model utilized ANN based feature extraction and MLP based forecasting, which was failed to consider the temporal feature extraction. The CD network utilized the unsupervised learning approach that was inefficient to handle uncertainties in reference flood maps. The GCN-LSTM model employed both spatial and temporal feature extraction; still, the representation of features through the graph matrix degrades the performance. The proposed method employed noise removal technique due to environmental and instrumental factors, which impact on the forecasting accuracy. To remove the noise, wavelet transform is used for signal decomposition and Fuzzy II threshold selection method is applied to distinguish between noise and required data. The noise removal process assists to improve the quality of input data for subsequent processing. IMF extraction using DEEMD is employed to capture the frequency components relevant to flood prediction. By extracting refined IMFs, DEEMD ensures the significant flood patterns that are helpful in enhancing the forecasting accuracy. Also, Att_RGRUNet assist in extracting the spatial and temporal features to generate precise flood predictions. Thus, enhanced outcome is derived by the proposed urban flood forecasting model.

Table 3. Comparative Analysis

Parameters	Metrics	GCN-LSTM	CD network	ANN_MLP	GBDT	Proposed
Temperature	MAPE (%)	12.98	6.369	5.07	3.98	2.94
	RMSE	4.0347	1.929	1.7461	1.7287	0.836
	MAE	2	1.3	1.25	1.2	0.8
Humidity	MAPE (%)	15.76	13.587	13.5337	12.3566	7.83
	RMSE	18.6453	11.5424	11.6242	9.82943	6.46314
	MAE	15	10	7.7	7.4	2.35
Rainfall	MAPE (%)	23.85	23.13	18	16.73	15.31
	RMSE	1.97756	1.84185	1.74185	1.53038	0.76984
	MAE	2.7	2.4	3.3	2.4	1.2

5. Conclusion

The proposed methodology of flood forecasting based on de-noising, decomposition and urban flood forecasting is introduced in this research. The de-noising process is devised using wavelet transform with Fuzzy II threshold selection for acquiring high-quality input data. Then, DEEMD-based IMF extraction effectively isolates significant flood-related attributes. Finally, the Attentive Residual Gated Recurrent Unit Network (Att_RGRUNet model) is devised for forecasting the urban flood by considering spatial and temporal feature extraction techniques along with the attention mechanism for enhancing the forecasting accuracy. Experimental results indicate that the proposed model outperforms conventional flood forecasting methods by demonstrating improved reliability and early warning capabilities. Still, the designed model has certain limitations like increased computational complexity due to multiple processing stages and the requirement for large-scale datasets for model training. Thus, in the future, an optimization based model will be designed for enhancing the computational efficiency by incorporating real-time data streams and extending the model to other hydrological applications.

Funding statement

No funding was received for this study.

Data availability statement

The datasets generated and or analyzed during the current study is publicly available of the submitted research work.

https://www.data.gov.in/catalog/departure-rainfall-data

https://www.kaggle.com/datasets/harinkl/rainfall-of-india

Conflict of interest

The authors declare they have no conflicts of interest to report regarding the present study.

References

Aja D., Elias E. and Obiahu O.H. (2020). Flood risk zone mapping using rational model in a highly weathered Nitisols of Abakaliki Local Government Area, South-eastern Nigeria. *Geology, Ecology, and Landscapes*, **4**(2), 131–139.

Akinyoola J.A., Oluleye A. and Gbode I.E. (2024). A review of atmospheric aerosol impacts on regional extreme weather and climate events. *Aerosol Science and Engineering*, **8**(3), 249–274.

Albahri A.S., Khaleel Y.L., Habeeb M.A., Ismael R.D., Hameed Q.A., Deveci M. and Alzubaidi L. (2024). A systematic review of trustworthy artificial intelligence applications in natural disasters. *Computers and Electrical Engineering*, **118**, 109409.

Anuradha T., Aruna Sri P.S.G. and RamaDevi J. (2024). Hybrid model for rainfall prediction with statistical and technical indicator feature set. *Expert Systems with Applications,* **249,** 123260.

Asfaw W., Rientjes T., Bekele T.W. and Haile A.T. (2025). Estimating elements susceptible to urban flooding using multisource data and machine learning. *International Journal of Disaster Risk Reduction*, **116**, 105169.

Babu T., Selvanarayanan R., Thanarajan T. and Rajendran S. (2024), Integrated Early Flood Prediction using Sentinel-2 Imagery with VANET-MARL-based Deep Neural RNN, *Global NEST Journal*, **26**(10). https://doi.org/10.30955/gnj.06554

Berkhahn S. and Neuweiler I. (2024). Data driven real-time prediction of urban floods with spatial and temporal distribution. *Journal of Hydrology X*, **22**, 100167.

Doan T.N. and Le-Thi D.N. (2025). A Novel Deep Learning Model for Flood Detection from Synthetic Aperture Radar Images. *Journal of Advances in Information Technology*, **16**(1).

El-Bagoury H. and Gad A. (2024). Integrated hydrological modeling for watershed analysis, flood prediction, and mitigation using meteorological and morphometric data, scs-cn, hec-hms/ras, and qgis. *Water*, **16**(2), 356.

Fernández-Nóvoa D., González-Cao J. and García-Feal O. (2024). Enhancing Flood Risk Management: A Comprehensive Review on Flood Early Warning Systems with Emphasis on Numerical Modeling. *Water*, **16**(10), 1408.

Flood Forecasting Dataset from https://www.kaggle.com/datasets/s3programmer/flood-risk-in-india?select=flood_risk_dataset_india.csv

Karthik S., Surendran R, Sam Kumar G.V. and Srinivasulu S. (2025), Flood Prediction in Chennai based on Extended Elman Spiking Neural Network using a Robust Chaotic Artificial Hummingbird optimizer, *Global NEST Journal*. https://doi.org/10.30955/gnj.07113.

Khan I.R., Elmahdy S.I., Rustum R., Khan Q. and Mohamed M.M. (2024). Floods modeling and analysis for Dubai using HEC-HMS model and remote sensing using GIS. *Scientific reports*, **14**(1), 25039.

Kumshe U.M.M., Abdulhamid Z.M., Mala B.A., Muazu T., Muhammad A.U., Sangary O. and Bala M.M. (2024). Improving short-term daily streamflow forecasting using an Autoencoder based CNN-LSTM Model. *Water Resources Management*, 1–17.

Peker İ.B., Gülbaz S., Demir V., Orhan O. and Beden N. (2024). Integration of HEC-RAS and HEC-HMS with GIS in flood modeling and flood hazard mapping. *Sustainability*, **16**(3), 1226.

Pokharel S. and Roy T. (2024). A parsimonious setup for streamflow forecasting using CNN-LSTM. Journal of Hydroinformatics, **26**(11), 2751–2761.

Qin F., Huang C. and Lin Z. (2024). Big data and artificial intelligence-driven natural disaster prediction and prevention: Technological advances and prospects. *Geographical Research Bulletin*, **3,** 381–398.

Ren H., Pang B., Bai P., Zhao G., Liu S., Liu Y. and Li M. (2024). Flood susceptibility assessment with random sampling strategy in ensemble learning (RF and XGBoost). *Remote Sensing*, **16**(2), 320.

Sengupta S. (2025). IoT-Based Flood Detection and Management Systems in Urban Areas. *Risk Assessment and Management Decisions*, **2**(2), 104–116.

Sundarapandi A.M.S., Navaneethakrishnan S.R., Hemlathadhevi A. and Rajendran S. (2024), A Light weighted Dense and Tree structured simple recurrent unit (LDTSRU) for flood prediction using meteorological variables, *Global NEST Journal*, **26**(8). https://doi.org/10.30955/gnj.06242.

Suwarno I., Ma'arif A., Raharja N.M., Nurjanah A., Ikhsan J. and Mutiarin D. (2021). IoT-based lava flood early warning system with rainfall intensity monitoring and disaster communication technology. *Emerging Science Journal*, **4**(0), 154–166.

Tang J., Wang J., Li J., Zhao P., Lyu W., Zhai W. and Yang C. (2024). Predicting human mobility flows in response to extreme urban floods: A hybrid deep learning model considering spatial heterogeneity. *Computers, Environment and Urban Systems*, **113**, 102160.

Vamsi C. and Amudha V. (2024, September). Accuracy improvement of flooded area detection from satellite images using support vector machine in comparison with novel logistic regression. *In AIP Conference Proceedings*, **2871**, 1. AIP Publishing.

Verma S., Verma R.K., Mishra S.K., Agarwal A. and Sharma N.K. (2024). Improved SCS-CN Methodology Incorporating Storm Duration and Temporally Decaying Retention for Enhanced Runoff Prediction. *Journal of Hydrologic Engineering*, **29**(6), 04024045.

Wahba M., Essam R., El-Rawy M., Al-Arifi N., Abdalla F. and Elsadek W.M. (2024). Forecasting of flash flood susceptibility mapping using random forest regression model and geographic information systems. *Heliyon*, **10**(13).

Yadav R., Nascetti A., Azizpour H. and Ban Y. (2024). Unsupervised flood detection on SAR time series using variational autoencoder. *International Journal of Applied Earth Observation and Geoinformation*, **126**, 103635.

Yang L., Yang Y., Shen Y., Yang J., Zheng G., Smith J. and Niyogi D. (2024). Urban development pattern's influence on extreme rainfall occurrences. *Nature communications*, **15**(1), 3997.

Zheng C., Yang W., Jiang X., Lian J., Hu D., Yan X. and Yan L. (2024). A novel integrated Urban flood risk assessment approach coupling GeoDetector-Dematel and clustering method. *Journal of Environmental Management*, **354**, 120308.

Zhou Y., Wu Z., Jiang M., Xu H., Yan D., Wang H. and Zhang X. (2024). Real-time prediction and ponding process early warning method at urban flood points based on different deep learning methods. *Journal of Flood Risk Management*, **17**(1), e12964.

Zhou Y., Wu Z., Liang Q., Xu H., Wang H. and Xue W. (2024). Threshold and real-time initiation mechanism of urban flood emergency response under combined disaster scenarios. *Sustainable Cities and Society*, **108**, 105512.