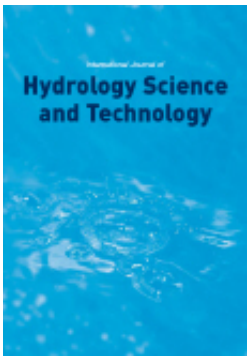
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Hydrology Science and Technology

ISSN online: 2042-7816 - ISSN print: 2042-7808
https://www.inderscience.com/ijhst

Analysis of flood severity using intelligent deep networks and sentinel image for the Kerala region

Supriya Kamoji, Mukesh Kalla

DOI: 10.1504/IJHST.2025.10067872

Article History:

Received:	27 March 2023
Last revised:	26 August 2023
Accepted:	30 August 2023
Published online:	03 December 2024

Analysis of flood severity using intelligent deep networks and sentinel image for the Kerala region

Supriya Kamoji* and Mukesh Kalla

Department of Computer Science and Engineering,
Sir Padampat Singhania University,
Bhatewar, Udaipur, India
Email: supriya.kamoji@spsu.ac.in
Email: mukesh.kalla@spsu.ac.in
*Corresponding author

Abstract: Accurate flood prediction and the classification of severity levels are vital for assessing the impact of floods and ensuring the safety of affected populations. Despite introducing numerous systems, many existing models require significant time to generate prediction results. Intelligent techniques like neural networks have been proposed to enhance prediction accuracy to address this. However, the complexity of data sources, such as satellite or sentinel data, poses challenges due to their unstructured nature and noisy attributes. This study introduces a pioneering approach known as the virtual bee-based recurrent model (VBRM) for flood prediction and severity classification. Specifically, the model categorises floods into low, medium, and high severity levels. The initial step involves gathering sentinel data related to the floods in Ernakulum, Kerala. This data is then used to train the VBRM, followed by a pre-processing stage that effectively filters out irrelevant features and noise. Subsequently, the refined data is fed into the classification layer, where the model extracts pertinent features and determines the severity level of floods. The ultimate objective is to achieve high prediction accuracy with minimal errors. Various performance metrics are employed and a comparative analysis is conducted against other existing models to evaluate the model's performance.

Keywords: deep networks; feature extraction; flood severity classification; optimisation; sentinel data; virtual bee-based recurrent model; VBRM.

Reference to this paper should be made as follows: Kamoji, S. and Kalla, M. (2025) 'Analysis of flood severity using intelligent deep networks and sentinel image for the Kerala region', *Int. J. Hydrology Science and Technology*, Vol. 19, No. 1, pp.1–23.

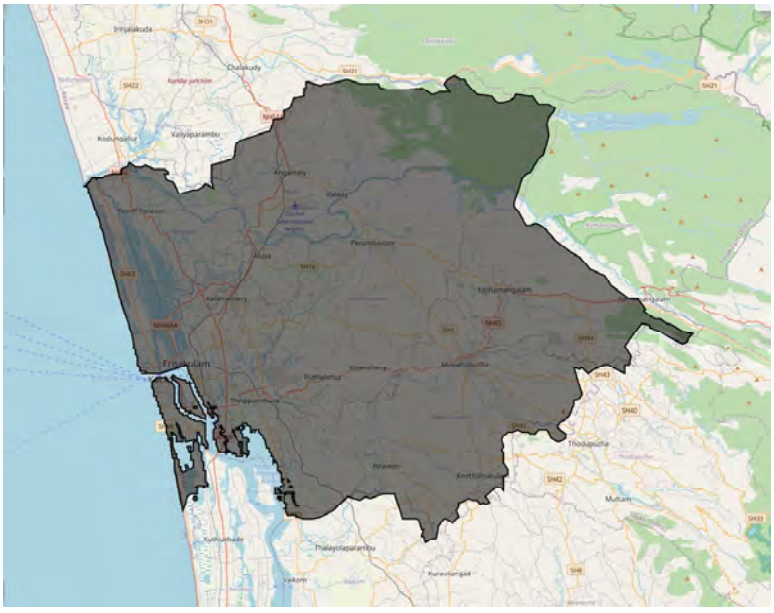
Biographical notes: Supriya Kamoji has 20 years of experience in teaching and research. She has completed her Master's in Computer Science from Thadomal Shahani college of Engineering, Mumbai University. She has published more than 17 papers in international Journals and conferences. She has worked on many research projects funded by university of Mumbai. Her areas of interest are artificial intelligence, machine learning, deep learning and cloud computing. She is a lifetime member of ISTE.

Mukesh Kalla has more than 15 years of experience in teaching and research. Currently, he is working as the Head of the Department of Sir Padampat Singhania University, Udaipur. He has completed his PhD from Rajasthan University in Computer Science Engineering. He has more than 20 international conference and Journal papers to his credit. His areas of interest are artificial intelligence, machine learning, deep learning and ad hoc networks.

1 Introduction

Floods are caused by prolonged periods of flooding caused by river overflow, excessive rainfall, and increasing sea levels (Jahangir et al., 2019). Hence, flooding is typically explained by natural elements, behaviours, and human behaviour (Fahy et al., 2019). Moreover, flooding is caused by various factors, including the size of the rain catchments, intensity and duration of the land use, rain, and topographic characteristics (Kankanamge et al., 2020). Artificial greenhouse gas emissions cause global climate change, brought on by physical reactions in Earth's ecosystems (Waghwala and Agnihotri, 2019). Global warming is a phenomenon that occurs when regular climate variables and patterns that can change on a local scale (Al-Juaidi et al., 2018).

Figure 1 Flood area mapping (see online version for colours)



Moreover, the variation in climate can result in weather systems that are not expected changes in average rainfall quantities that differ from what was considered 'normal' decades ago (Seejata et al., 2018). The frequency of heat waves, droughts, or cold spells is increased each year (Fassoni and Rodrigo, 2019). In addition, the increased incidence of rainstorms or blizzards has changed the temporal patterns of autumn and spring (Tang et al., 2020). Also, it has increased the frequency of occurrence of flood events (Shatnawi et al., 2020). Climate change events are defined as others that represent altered spatial patterns, frequencies, and normal conditions magnitudes (Ali et al., 2019). In addition, changes in flood frequency might result in climate variations that include temperature, rain, fall, and humidity (Luu et al., 2018). However, flooding is not solely determined by seasonal rainfall or the melting of mountainous glaciers or snowpack (Amitrano et al., 2018).

Furthermore, environmental Impacts may also contribute to flooding factors (Shafapour et al., 2019). Floods are natural physical occurrences throughout geological

history due to excessive runoff (Zhao et al., 2019). On the other hand, when rivers overflow their banks, surplus water is channelled into the floodplain (Wang et al., 2020).

Hence, the floodplains are ideal for human settlement because they are productive, close to water, and easy to excavate (Jha and Haripriya, 2019). These characteristics have aided in the growth and urbanisation of floodplains (Toosi et al., 2019), increasing the probability of disaster-causing floods. Despite the development of flood mitigation measures such as rivers, dams, ponds, and river channelling, flooding severity has increased (Singh et al., 2018). Several methods have been implemented to map flood severity, such as the neural model (Khoirunisa et al., 2021) and the ensemble model (Jhong et al., 2022). But, relevant results were not found, so the present work has developed based on deep parameters to predict the flood severity.

Additionally, the critical curiosity of this current review is fusing the upgraded profound organisations highlighted in the Google Earth Engine (GEE) to foresee the flood highlights from the planned explicit locale. Moreover, before detecting the flood-affected area in the Ernakulum, the area's water features were traced and visualised before and after flood visualisation. Moreover, after-flood image is considered for the different processes, then features extraction and flood severity level prediction in each area of Ernakulum district. Here, the severity-affected region of the Ernakulum district 2018 has been specified as low flood harshness, high flood harshness, and medium flood harshness.

The critical process of this article is depicted as follows,

- At first, the Ernakulum, Kerala flood sentinel picture dataset was viewed in this exploration and prepared for the framework.
- Therefore, a new virtual bee-based recurrent model (VBRM) has been designed to predict flood severity.
- In the essential stage, the current noise in the dataset has been taken out utilising the pre-processing function then the noiseless information is given to the classification module of VBRM.
- Here, the presence of the virtual bee in the recurrent layer has inclined to get the optimum prediction outcome.
- Finally, the gained results are compared with recent other models in terms of accuracy, Kappa coefficient, producer accuracy, consumer accuracy, and error rate.

The current exploration concentrates on segments organised as follows; the new related works are portrayed in Section 2; the issue in flood planning and seriousness order is depicted in Section 3. Section 4 provides solutions to the problems raised, and the Section 5 represents the after-effect of the cure. Section 6 has finished up the exploration contentions.

2 Related work

A portion of the new related literature based on the present research work is elaborated below:

Rafiei et al. (2021) have designed a flood mapping system based on neural features to evaluate the risk rate of flood occurrence. Hence, the decision model has been designed to find the flood density by incorporating the required flood features. Then during the execution process, the thickness of the flood was determined from the test image using the saved flood features. Nonetheless, it has taken more time to process.

Nachappa et al. (2020) have designed the ensemble model with the optimisation function for the flood mapping system to gain better mapping results. Here, the optimisation process is executed during the classification phase of the ensemble model. Finally, the implementation outcome verified the planned method's robustness by earning the highest mapping result; nevertheless, the optimisation procedure took longer.

Darabi et al. (2019) have designed a relative study for the different ML schemes to identify the appropriate model for predicting flood severity. Moreover, evaluating the prediction parameters and the kappa statistics has specified the finest model. In that, the deep features have recorded outstanding results than other approaches. Moreover, all the ML models have gained better outcomes for particular objectives. But some models have needed additional time for the executive function.

The optimisation with an extreme ML model has been considered by Bui et al. (2019) for flood mapping. Here, the floods were mapped from the image data, so the error parameters have been analysed in different ways like root square, mean value, etc. Hence, the designed extreme model has gained wide frequency in predicting the flood severity from the mapped location. However, if the data is complex, the prediction value has decreased.

The flood-potential index (FPI) has been calculated using ten flood predictor variables, 158 non-flood, and 158 flood pixels by Romulus (Costache, 2019). Here, the Prahova Rivers are the study region considered in the present study. Usually, the Prahova River's middle and upper basins have a high or extremely high flood threat. Moreover, The ROC parameters were used to validate the results. Furthermore, the flood density has been mapped with the help of fuzzy models. But, designing the FPI is complex in design.

Kerala has a high annual rainfall rate that frequently causes flooding and other natural disasters. In this study, Saravanan et al. (2023) created maps of flood susceptibility in the Idukki district using data from remote sensing, machine learning (ML), and geographic information systems. In this work, the flood susceptibility of the Idukki district of Kerala is assessed by employing five different ML models: Adaboost, gradient boosting, extreme gradient boosting (XGB), CatBoost, and stochastic gradient boosting (SGB). A total of 16 hydrometeorological variables were considered. The accuracy of several strategies was assessed using the area under the curve (AUC) in terms of both achievement rates and forecasts. The effectiveness of each system remained demonstrated by the validation results.

In this work, it is recommended by Zhang et al. (2023) to use the high-resolution multi-source remote sensing dataset GF-FloodNet to pinpoint flood zones. GF-FloodNet contains 13,388 data from the images of the Gaofen-3 (GF-3) and Gaofen-2 (GF-2). Researchers created it using a multi-level sample selection and interactive annotation method based on active learning. GF-FloodNet differs from flood-related datasets because it offers pixel-level labelling, a spatial resolution of up to 1.5 m, and multi-source remote sensing data. Findings from experiments show that multi-source data are beneficial for GF-FloodNet. It is possible to support a variety of deep learning models when training for extracting flood regions. A potential optimal boundary region for

model training should be present in any deep-learning dataset. The barrier in GF-FloodNet appears to be close to 4824 samples.

This article (Abdollahi and Pradhan, 2023) demonstrates how the findings from a Shapley additive explanations (SHAP) model may be interpreted using a deep learning (DL) model created to predict wildfire vulnerability. Several SHAP graphs are used to comprehend which factors influence a prediction model, their order of significance, and the reasoning behind certain decisions. Numerous contributing factors are fed into the model, including topographical, land cover/vegetation, and climatic features. SHAP plot results show the significance of variables such as humidity, wind speed, rainfall, elevation, slope, and normalised difference moisture index (NDMI) in the outcome of the suggested approach for mapping wildfire risk. Specialists accept that making a reasonable model would make it more clear the model's decision to plan fierce blaze helplessness, recognise high-contributing components in the forecast model, and effectively oversee fire dangers.

RS and GIS advancements have become progressively critical lately in various precipitation-actuated landslip examination features. This article (Yang et al., 2023) broadly uses different perception logical advancements for inside and out investigation of 1,161 records procured by the WOS information stage throughout recent years by joining quantitative and subjective procedures to grasp their application situation methodically. The accompanying segment of this paper focuses on sub-space examination from four points: landslip discovery and checking, forecast models, responsiveness planning, and hazard evaluation. As per the overview, how much writing in this space has been slowly developing and will probably do as such from now on. This study can help navigation seeing avalanche anticipation and control as well as the accomplishment of the supportable advancement objectives set out by the unified countries by providing analysts in this field with the essential exploration questions, hot examination subjects, and future improvement patterns of future precipitation actuated avalanches.

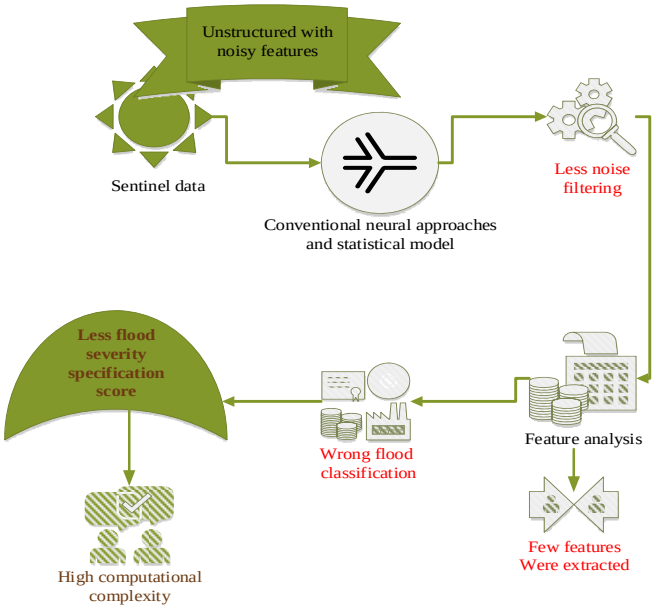
The ongoing review (Youssef et al., 2023) investigates the impacts of flood danger demonstrated nearby. Execution appraisals of three strategies – logistic regression (LR), extreme gradient boosting (EGB), and random forest (RF) – was completed to distinguish the best calculation for planning flood powerlessness. With field trips, remote detecting, geographical, geologic, and meteorological information were utilised to make the spatial data set and stock datasets the models required. As our exploration indicates, RF is the best calculation for planning flood helplessness, trailed by EGB and LR. Extraordinary upsides of ROC (93%), OAC (88%), and Kappa record (0.85), as well as the least upsides of RMSE (0.34) and MAE (0.12), show that RF offers incredible execution. Sentinel-1 photographs for genuine floods in 2016 and 2021 confirmed the RF model and the outcomes show great understanding.

3 System model and problem description

Predicting the flood rate from trained parameters is difficult because of complex, unique data for different regions. Several ML models were implemented, but those have achieved the best results for the specific area and conditions because of the data complexity (Brunner et al., 2021). Also, in some cases, the prediction score became very

low and required more duration for the prediction process. If the technique has taken more resources to predict the flood severity, it has resulted in high computational costs.

Figure 2 System model with problem (see online version for colours)



So, efficient deep features models are required to classify the flood region in a short duration with a very high exactness score. The conventional flood prediction framework with issues is discussed in Figure 2. But in many cases designing the efficient deep network-based classification model has increased the design and implementation complexity because of the several additional layers and number of neurons. These problems have inspired this investigation to design an intelligent structure for predicting flood severity in numerical data.

4 Proposed VBRM for flood prediction

A novel VBRM has been implemented for flood density prediction applications. Initially, the system is trained with flood mapping datasets that include rainfall parameters, in that the Ernakulum, Kerala region has been mapped, and the flood prediction process has been continued. Consequently, the flood range has been calculated based on rainfall features. The area that has been considered in this present work is Kerala state, India. Initially, the Kerala flood region was mapped then, based on the trained parameters, the flood range was predicted.

Figure 3 describes the suggested architecture. At last, the essential parameters were assessed, and a comparison analysis was done with other existing approaches. Subsequently, the percentage of improvement has been noted based on the comparison results.

Figure 3 Proposed methodology (see online version for colours)

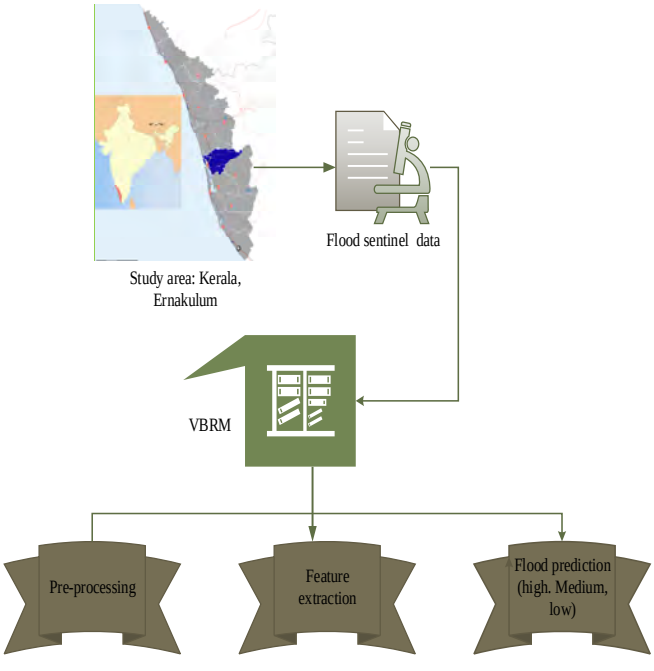
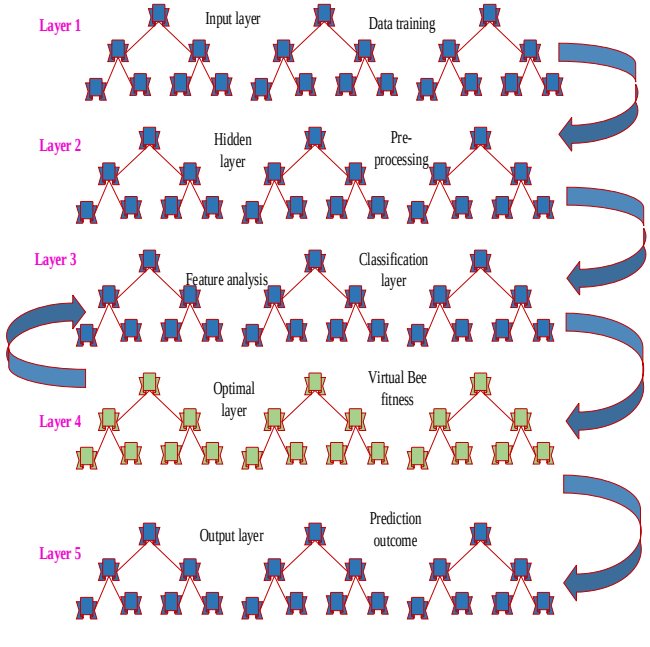


Figure 4 Layers of VBRM (see online version for colours)



4.1 Proposed design of VBRM

The virtual bee algorithm was the basis for creating the novel VBRM (Zhu et al., 2021) and recurrent neural system (Ackerson et al., 2021). The present work has five phases processed with the principle of a novel VBRM. Moreover, the five layers are the input phase, hidden layer, classification module, optimisation module, and output layer.

Here, the Kerala region flood sentinel data is taken as the database, which is trained to the input layer. Then the noisy contents are removed in the hidden phase with the help of error-eliminating formulations. Moreover, flood features were extracted, and the classification module specified the affection range. The optimal layer tunes the classification parameter; virtual bee fitness is added to the optimal layer. Finally, the detection efficiency score has been gained in the output phase. Hence, the discussed layers are represented in Figure 4.

4.1.1 Pre-processing

In order to make classification and feature analysis less complex, the pre-processing function was performed in the designed novel VBRM's hidden phase. Moreover, the dataset training process has been equated in the equation (1)

$$F\left(s_d\right) = s_d\{1, 2, 3, ..., n\} \tag{1}$$

Now, the Kerala flood sentinel information has been signified s_d, and the training function of the sentinel data is determined as $F(s_d)$ and 1, 2, 3, …, n has determined the total number of information. The error elimination process has been defined in equation (2) following the training phase.

$$P^* = \sum_{n=0}^{n-1} B\left(s_d\left(mi - ai\right)\right) \tag{2}$$

Here, the pre-processing variable is denoted as P^*, the tracking variable is determined as B, standard features are measured as mi, and the noisy features are described as ai. Here, the loud parts are removed from the sentinel dataset defined as $s_d(mi - ai)$ in equation (2). Data with fewer errors have been obtained in the end. In the code of the Kerala dataset, the intermediate outputs of pre-processing include transforming the data into long-format, mapping months to numerical values, creating a 'date' column, and scaling rainfall values.

4.1.2 Feature analysis

The present meaningful flood features in the trained refined sentinel data have been analysed and extracted to find the flood-affected region. The LSTM model is trained on the pre-processed data for feature analysis, and predictions are made on training and testing sets for rainfall forecasting. However, the code does not explicitly demonstrate additional feature engineering or in-depth feature analysis beyond the data preparation and model training steps. Hence, the feature extraction module has been executed in the designed model.

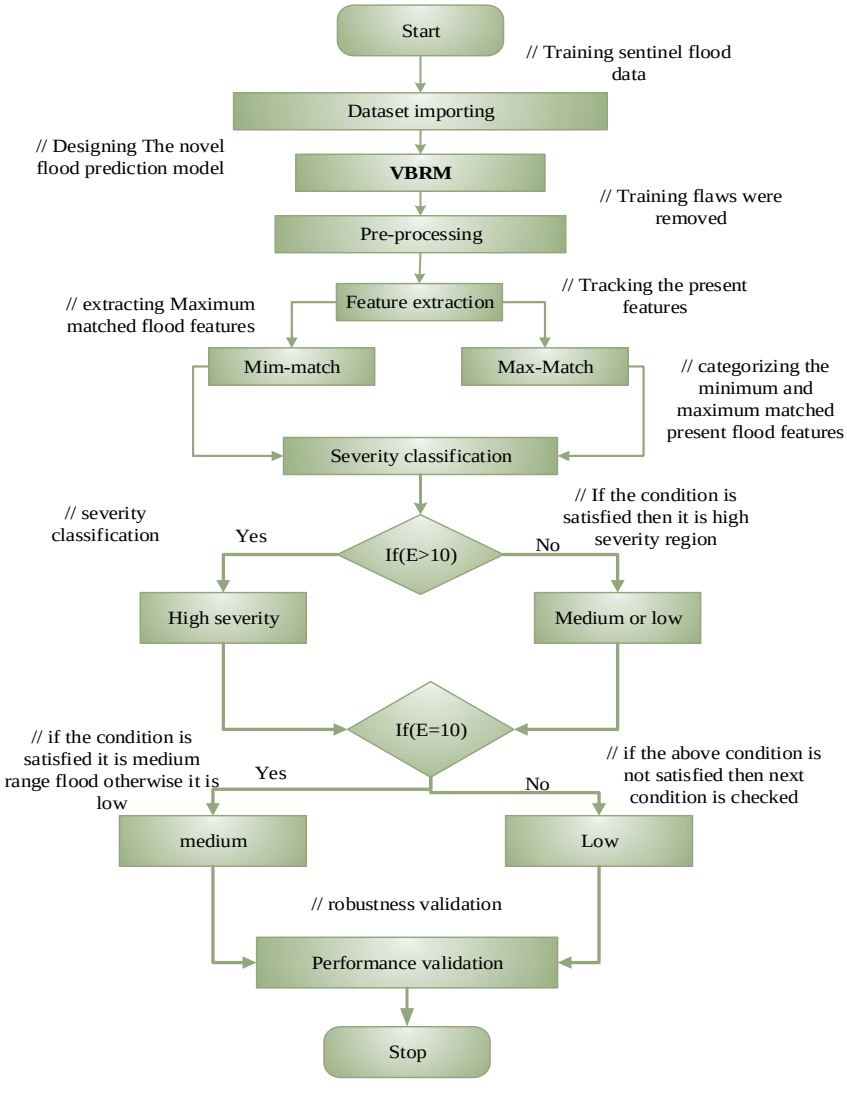
$$D'F\left(s_d\left(\mathit{features}\right)\right) = \sum_{n} s_d\left(\min, \max\right) \tag{3}$$

The feature analysis process has been functioned by equation (3). Here, the term min represents minimum matched and max maximum matched flood features. Moreover, $F(s_d(features))$ it denotes the flood features set. Hereafter, the maximum matched flood features were tracked and extracted for further classification functions.

$$E\left(s_d(\max\_\ features) = s_d\left(\frac{\min + \max}{\max}\right)\right. \tag{4}$$

Thus, equation (4) was used to extract the most closely matching flood features. The current flood features were extracted as quickly as possible using equation (4) above as the extraction variable.

Figure 5 Flow of VBRM (see online version for colours)



4.1.3 Classification

Following the extraction and tracking of the flood features, the flood range has been categorised as high, medium, and low. Hence, to classify the flood severity range classes, the fitness value of the virtual bee has been considered.

$$classification = \begin{cases} if\,(E(\max\_\ feature) > 10) & High \\ f\,(E(\max\_\ feature) = 10) & Medium \\ f\,(E(\max\_\ feature) > 10) & Low \end{cases} \tag{5}$$

Moreover, the flood severity classification has been executed by equation (5). Here, the value 10 is the fitness solution of the virtual bee. Moreover, if the extracted features are more significant than the value of 10, then it is classified as high flood severity; if the extracted flood feature is similar to 10, then it is a medium-range flood, and if the extracted flood feature is less than ten then is categorised as low severity flood.

Algorithm 1 VBRM

Start
{
Data initialisation ()
{
int s_d = 1, 2, 3, …, n;
// Kerala, flood sentinel data, has been initialised
}
Pre-processing ()
{
int P^* = mi, ai, B;
//pre-processing variable initialisation
$P^*(s_d) \rightarrow remove\_ai(s_d)$
// removing noise features from the trained set
}
Feature extraction
{
int min, max, E
// feature extraction parameter initialisation (minimum and maximum variable)
extract $\rightarrow s_d(max\_features)$
// extracting the available maximum disease features
}
Classification ()
{
$f(E(max\_feature) > 10)$
{
High
} **else** $f(E(max\_feature) = 10)$

```
        {
        Medium flood region
        }else (low)
    }
}
Stop
```

The pseudo-code structure used in the discussed mathematical formulations is described in Algorithm 1. Effectiveness of this approach has been evaluated, and the code was created in the Python environment. In addition, the GEE framework has been utilised to map the study location and flood-affected region. Moreover, the designed novel VBRM workflow has been diagrammatically stepwise, as is determined in Figure 5.

5 Results and discussion

The planned model is tried in Python and running on the Windows 10 stage. Thus, to esteem the functioning presentation of the proposed model, the sentinel flood information has been taken and prepared in the framework. The execution boundaries are arranged in Table 1.

Table 1 Parameterisation of the execution

Parameter description	
Programming platform	Python
Operating system	Windows 10
Version	3.8
Database type	Sentinel data
Study area	Ernakulum (Kerala)
Deep learning	Recurrent neural model
Optimisation	Virtual bee

5.1 Dataset description

The dataset at hand encompasses a comprehensive record of monthly rainfall patterns across 36 distinct meteorological sub-divisions within India, spanning an extensive temporal scope from 1901 to 2015. This dataset holds a monthly granularity, presenting a detailed breakdown of rainfall measurements over this substantial time frame. The geographical coverage extends across India's meteorological sub-divisions, encompassing well-known and more localised areas. Rainfall measurements are quantified in millimetres (mm), providing a standardised unit for assessing and comparing precipitation levels across the sub-divisions. This dataset is valuable for analysing historical rainfall trends and patterns, enabling researchers and meteorologists to gain insights into the variations and changes in India's rainfall over more than a century.

5.2 Case study

The designed model is applied to the Kerala flood sentinel database, and the outcome of this testing function is described as the case study, which is detailed as follows. Here, the training and testing process has functioned in the ratio of 80:20, which is training 80% and testing 20%.

Figure 6 Mapped the Kerala region (see online version for colours)

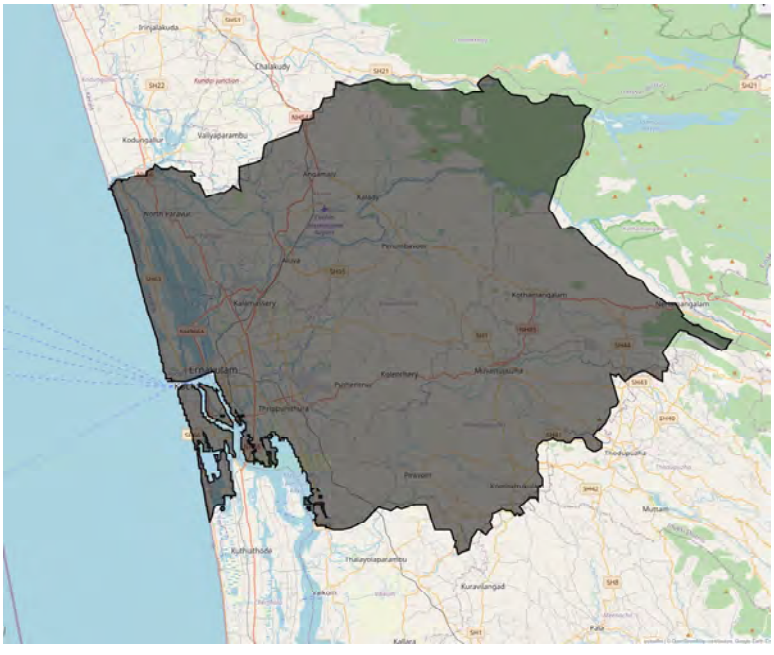
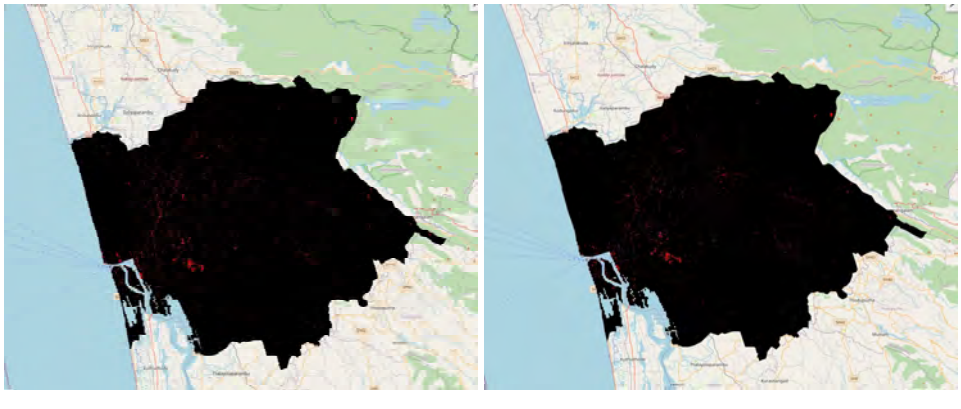


Figure 7 Feature extraction, (a) before the flood, (b) after the flood (see online version for colours)



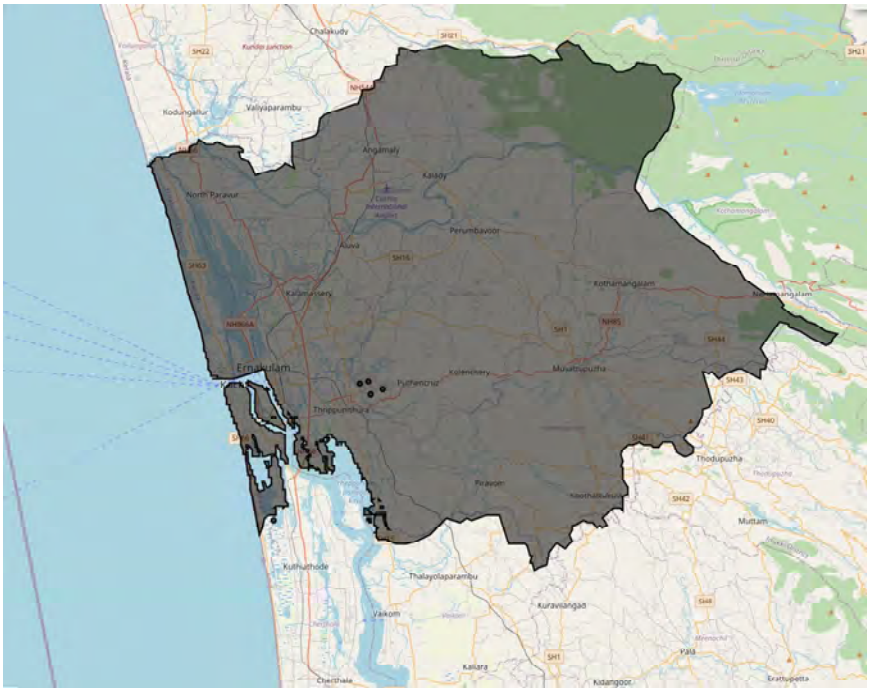
(a) (b)

The mapped Kerala region using the GEE is described in Figure 6. The present research work has adopted the Kerala region as the study area for detecting the flood affected area and severity range.

The Kerala region contains several water-covered areas, so the water-covered region has been analysed before and after the flood occurrences described in Figure 7. Moreover, the water features in the sentinel data are extracted in Figure 7(b). Consequently, the severity range prediction has been performed.

The high-severity region mapped location is described in Figure 8. The black dots represent the area affected by high flood severity. Hence, those regions are Irumpanam, Varikkoli, and Thripunithura, mapped in Figure 8.

Figure 8 High severity region (see online version for colours)



Moreover, the present model has mapped ten regions as low flood-affected areas, as determined in Figure 9. Hence, the region affected by a low range of flood is North Paravur, Piravom, Muvattupuzha, Thodupuzha, Valliya Parambu and Neriamangalam.

The regions like Alangad, Kodikuthmala, Angamaly, and Kanakkankadavu have been affected by floods in a medium range. Hence, these mapped areas are described in Figure 10. We have employed a ML model, specifically long short-term memory (LSTM), to forecast and categorise the severity of different regions into three distinct flood categories: high, medium, and low. This categorisation is likely based on pertinent information such as rainfall data. However, the code does not overtly indicate that we are engaged in a multi-class classification task. Primarily, the code centres around pre-processing the data, training the LSTM model, assessing the model's performance using the test dataset, and computing diverse evaluation metrics such as accuracy, Kappa coefficient, producer accuracy, and consumer accuracy.

Figure 9 Low flood severity region (see online version for colours)

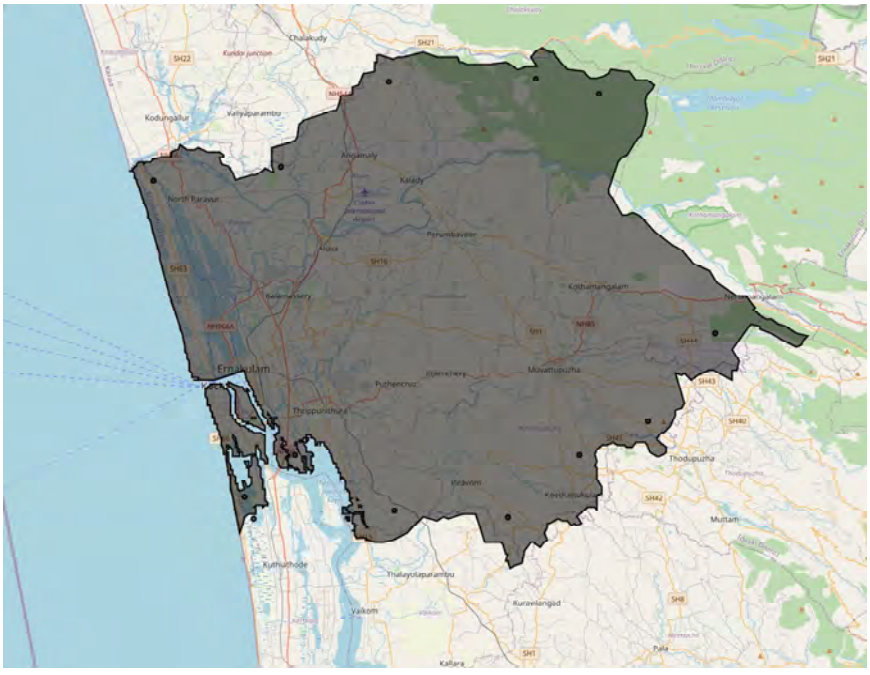
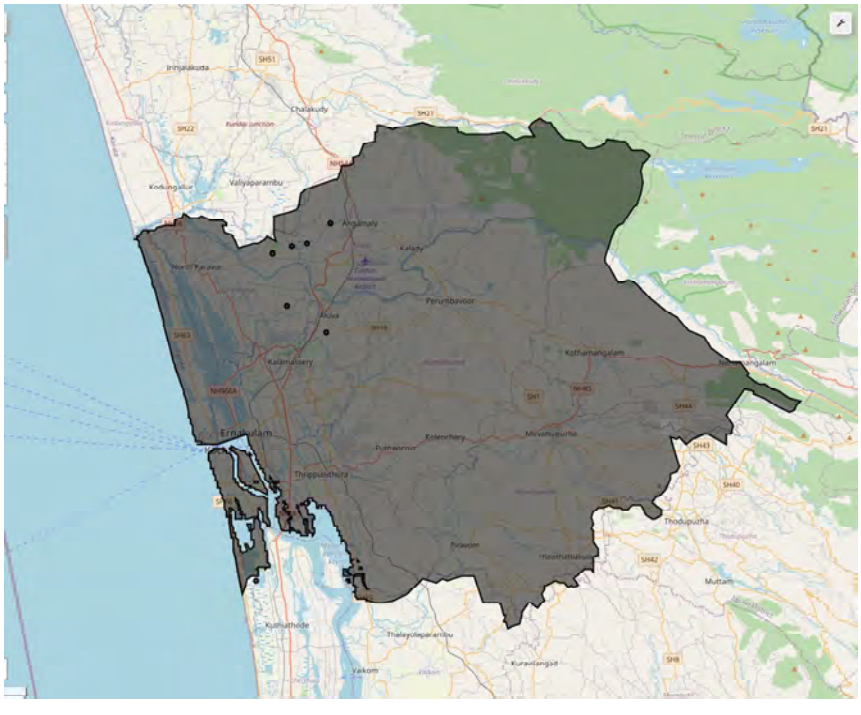


Figure 10 Medium severity region (see online version for colours)



5.3 Performance metrics

This section covers the experimental setting and the suggested way's effectiveness. Several metrics, including accuracy, producer accuracy, consumer accuracy, and kappa coefficient and error rate, are used to assess the framework's efficiency for flood severity analysis using the VBRM method and sentinel image for the Kerala region.

5.3.1 Accuracy

The accuracy in this context refers to measuring how accurately the intelligent deep networks and Sentinel imagery can predict and classify the severity of floods within the Kerala region. It is derived by dividing the entire amount of predictions made using the computational models by the number of flood severity levels that were accurately anticipated. This metric assesses the models' capability to effectively and precisely categorise the extent of flooding in the specified area. The prediction exactness score has been measured as classification accuracy, formulated in equation (6)

$$Accuracy = \frac{P_B + N_B}{P_B + N_B + P_C + N_C} \tag{6}$$

Here, true-positive is represented as P_B true-negative is determined, false positive is indicated, and false negative is exposed as N_B.

5.3.2 Error rate

The error rate refers to the proportion or percentage of erroneous predictions or classifications a predictive or analytical model makes. It is frequently used to evaluate the accuracy and dependability of a model's performance and is derived by dividing the number of wrong predictions by the total number of predictions. A lower error rate denotes greater accuracy and a better match between the model and the data. In contrast, a higher error rate suggests that the model's predictions are less reliable and may require refinement. The error metrics validation has been formulated in equation (7).

$$Error\ Rate = \frac{predicted\ outcome}{extracted\ \ prediction} \tag{7}$$

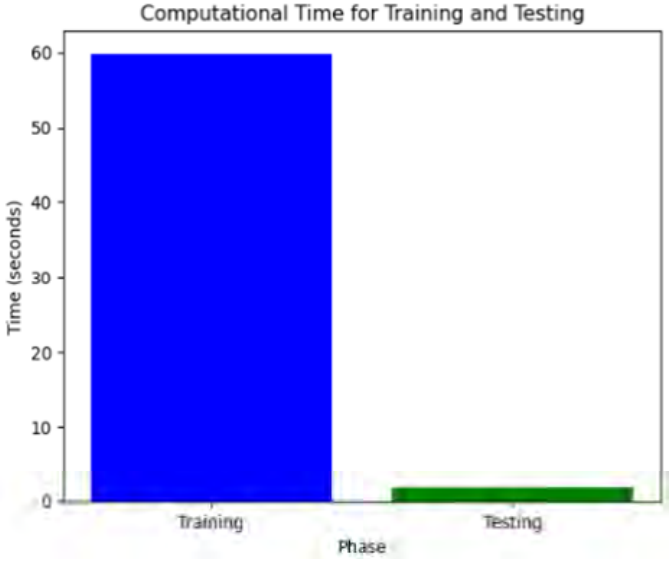
5.3.3 Kappa coefficient

When working with categorical data, the kappa coefficient, commonly called Cohen's kappa, is a measure of statistics that assesses the degree of consensus between two rates or classifiers. It accounts for the potential for coincidental concurrence and offers a more complex evaluation of agreement than simple correctness. The kappa coefficient has a range of values from –1 to 1, with one denoting perfect arrangement, 0 denoting understanding by chance, and negative values denoting agreement worse than chance. It is commonly used in inter-rate reliability studies, where multiple observers or classifiers assess the same set of items, and the degree of agreement needs to be quantified while accounting for random arrangement.

5.3.4 Computation time of VBRM

Computational time is required for a computer or computational system to perform specific tasks, calculations, or operations. It measures how long it takes for a computer to complete a set of instructions or algorithms, which could involve functions like data processing, calculations, simulations, or any other form of computation. Computational time is typically measured in units of seconds, milliseconds, or even nanoseconds, depending on the task's complexity and the hardware's processing speed. The computational time for the proposed virtual bee recurrent model (VBRM) is mentioned in Figure 11.

Figure 11 The computation time of VBRM (see online version for colours)



The computation time associated with the VBRM, also known as the VBRM, signifies the duration required to perform various computations involved in utilising the model. For this specific model, the training process consumes 59.44 seconds, encompassing tasks such as optimising model parameters, weight adjustments, and fine-tuning to ensure optimal predictive capabilities. Subsequently, during the testing phase, the model demonstrates efficiency by completing its computations within a mere 1.78 seconds. These time measurements shed light on the speed at which the VBRM can process and analyse data, enabling timely predictions and classifications. Such insights into computation time are crucial in assessing the model's practicality and suitability for real-time applications where swift and accurate predictions are essential.

5.4 Performance estimation

A Python platform is used to execute the created model, and the efficiency of the classical was confirmed by relating its metrics to those of other models in standings of accuracy, producer accuracy, consumer accuracy, and kappa coefficient and error rate. A few of the current approaches have been taken into consideration in addition to checking

the model's successive score, which are high-resolution-net (HRNet) (Dong et al., 2021), bimodal-threshold scheme (BTS) (Dong et al., 2021), class variance (CV) (Dong et al., 2021), DenseNet (Dong et al., 2021), Segnet (Dong et al., 2021), and DeepLab (Dong et al., 2021). Quantum-particle swarm credal decision model (QPSCDM) (Ngo et al., 2021), support vector model (SVM) (Ngo et al., 2021), classification based on regression tree (CbRT) (Ngo et al., 2021), logistic based regression (LR) (Ngo et al., 2021), best-first tree (BFT) (Ngo et al., 2021), credal-decision-tree (CDT) (Ngo et al., 2021). The accuracy parameter was calculated to determine the correctness of the flood prediction range.

5.4.1 Comparison of the suggested method's accuracy with that of current approaches

It is necessary to assess how well the offered strategy predicts outcomes or categorises objects to compare the given method's accuracy. This evaluation often compares the percentage of accurate predictions produced using the suggested approach to the entire amount of forecasts made. By contrasting the recommended method's accuracy with other available strategies or benchmarks, one can determine its effectiveness in providing accurate results. This comparison aids in gauging the reliability and precision of the proposed method's outcomes and making informed decisions about its suitability for specific applications or tasks.

Figure 11 displays the examination of the proposed VBRM accuracy in terms of other existing techniques. The suggested VBRM accuracy rating is 99%. The accuracy level of the proposed VBRM is higher when compared to other current approaches. The accuracy rates for the HRNeT, DenseNet, SegNet, ResNet, DeepLab, CV and BTS currently in use are 97%, 96%, 96%, 96%, 95%, 69% and 76%, respectively. While comparing the values of existing techniques with the VBRM, it attains higher accuracy. Table 2 displays the accuracy comparison values between the suggested and currently used strategies.

Table 2 Comparison of the recommended method's accuracy values with those of current approaches

Methods	*Accuracy (%)*
HRNeT (Dong et al., 2021)	97
DenseNeT (Dong et al., 2021)	96
SegNet (Dong et al., 2021)	96
ResNet (Dong et al., 2021)	96
DeepLab (Dong et al., 2021)	95
CV (Dong et al., 2021)	69
BTS (Dong et al., 2021)	76
Proposed VBRM	99

5.4.2 Comparison of the suggested approach with the current approach in terms of error rate

The comparison of the suggested method with current methods concerning error rate involves analysing the performance of the newly introduced approach in making

predictions or classifications. This assessment primarily focuses on quantifying the proportion of incorrect predictions or types generated by the proposed method concerning the number of estimations made overall. By contrasting the error rates of the proposed method with those of established techniques or benchmarks, one can ascertain how well the new approach minimises or avoids prediction errors. This comparison aids in understanding the relative reliability and precision of the proposed method's outcomes and provides insights into its potential advantages or limitations compared to existing methods.

Figure 12 Comparison of the suggested model in standings of accuracy (see online version for colours)

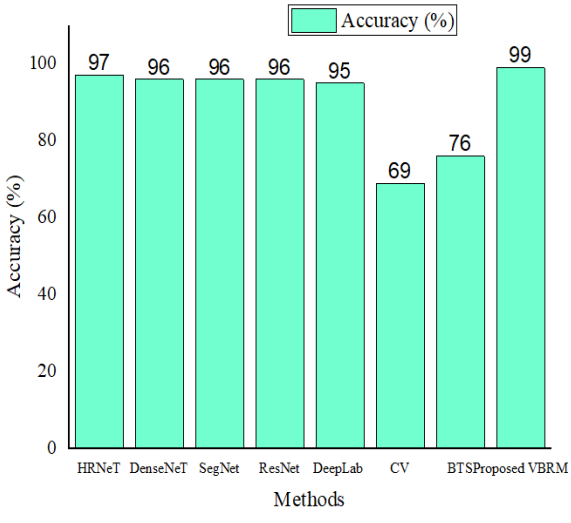


Figure 13 Error rate comparison between the suggested model and the other approach (see online version for colours)

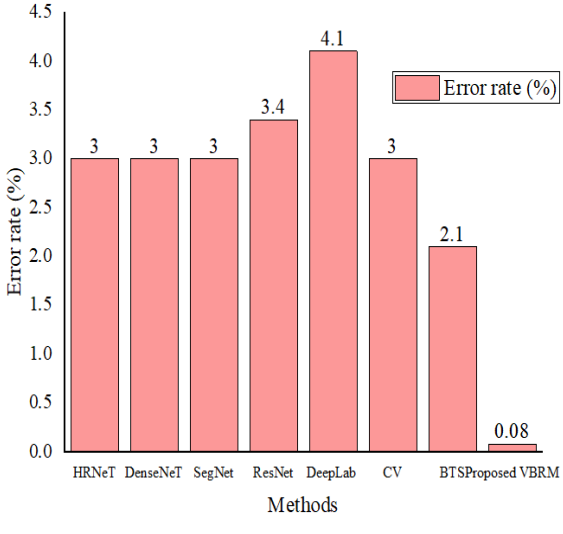


Figure 13 displays the examination of the proposed VBRM error rate in terms of other existing techniques. The suggested VBRM error rate is 0.08%. The error rate level is relatively low when comparing the suggested VBRM to other current methods. The accuracy rates for the HRNeT, DenseNet, SegNet, ResNet, DeepLab, CV and BTS currently in use are 3%, 3%, 3%, 3.4%, 4.1%, 3% and 2.1%, respectively. While comparing the values of existing techniques with the VBRM, it attains a meagre error rate. Table 3 displays the error rate comparison numbers between the suggested and other currently used strategies.

Table 3 Comparison of the suggested method's Error Rate values with those of current approaches

Methods	*Error rate (%)*
HRNeT (Dong et al., 2021)	3
DenseNeT (Dong et al., 2021)	3
SegNet (Dong et al., 2021)	3
ResNet (Dong et al., 2021)	3.4
DeepLab (Dong et al., 2021)	4.1
CV (Dong et al., 2021)	3
BTS (Dong et al., 2021)	2.1
Proposed VBRM	0.08

5.4.3 Kappa coefficient comparison of the suggested approach with other current techniques

Analysis of the kappa coefficient to compare the proposed method to other current methods involves evaluating the performance of the new approach regarding the level of agreement achieved between predictions or classifications. This assessment revolves around calculating the kappa coefficient, which considers both the observed agreement and the agreement expected by chance. By comparing the kappa coefficient of the suggested technique with that of established methods or benchmarks, one can gain insights into how well the new approach aligns with or improves upon existing methods in terms of agreement beyond chance. This comparison aids in understanding the comparative reliability and robustness of the proposed method's outcomes and determines its potential strengths or advancements in achieving higher levels of agreement in predictions or classifications. The prediction rate exactness has been valued by measuring the statistical test. Moreover, the kappa metrics have provided the sensitivity rate of the prediction. Hence, the kappa score has ranged from 0 to 1.

Figure 14 displays the examination of the proposed VBRM kappa coefficient in terms of other existing techniques. The suggested VBRM kappa coefficient is 0.99%. The kappa coefficient level is high when comparing the proposed VBRM to other current methods. The kappa coefficient rates for the QPSCDM, SVM, CbRT, LR, CDT and BFT currently in use are 0.8%, 0.71%, 0.76%, 0.71%, 0.71% and 0.71%, respectively. While comparing the values of existing techniques with the VBRM, it attains a more significant kappa coefficient. The comparison values for the proposed with other current methods concerning error rate are mentioned in Table 4.

Figure 14 Comparison of the suggested method's kappa coefficient with that of current techniques (see online version for colours)

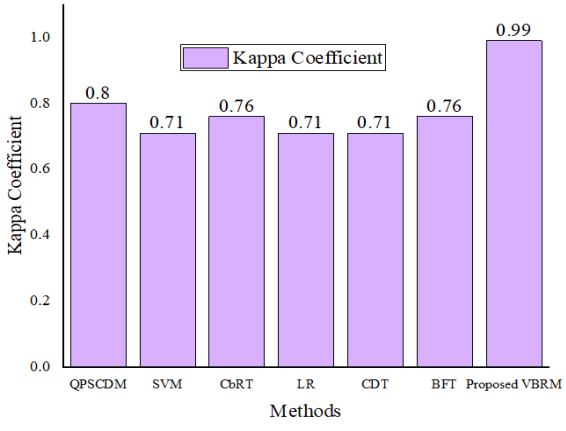


Table 4 Kappa coefficient values of the suggested method compared to those of current approaches

Methods	*Kappa coefficient*
QPSCDM (Ngo et al., 2021)	0.8
SVM (Ngo et al., 2021)	0.71
CbRT (Ngo et al., 2021)	0.76
LR (Ngo et al., 2021)	0.71
CDT (Ngo et al., 2021)	0.71
BFT (Ngo et al., 2021)	0.76
Proposed VBRM	0.99

5.5 Discussion

The presented VBRM has recorded the finest performance score in detecting and classifying the flood region and severity range. It has demonstrated the provided model's robustness score. The overall performance values are displayed in Table 5, along with the Kappa coefficient, accuracy, producer accuracy, consumer accuracy, and error rate.

Table 5 Overall performance statistics

	Performance statistics
Accuracy	99.10%
producer accuracy	99.10%
Consumer accuracy	99.10%
Kappa score	0.99%
Error rate	0.08%

The comprehensive performance evaluation is presented in Table 5. In this table, three distinct accuracy metrics have been computed to gauge the developed model's operational efficacy effectively. Specifically, the producer accuracy delineates the

training accuracy rate, while the consumer accuracy elucidates the accuracy rate of feature predictions. Remarkably, the showcased model has consistently demonstrated an exceptional performance of 99.1% across all assessments. This uniformity underscores the model's robustness and adaptability, signifying the reliability of the novel VBRM that has been crafted.

Furthermore, the kappa coefficient records an impressive 99.00%, indicating a high level of agreement, and the error rate is a mere 0.08%. Compared to existing methods, the proposed approach outperforms the rest, securing the highest level of performance. The novel amalgamation of recurrent neural networks and Virtual Bee-based computation within the VBRM effectively tackles the intricacies arising from intricate and noisy flood data. This unique blend results in heightened precision, faster processing, noise reduction proficiency, and enhanced severity classification. Consequently, the VBRM stands as a superior performer compared to conventional flood prediction models. Its superiority translates into more potent flood management and mitigation strategies, significantly enhancing their effectiveness.

6 Conclusions

The current research has introduced the novel VBRM for flood prediction application in the Ernakulum, Kerala district. Moreover, the presence of the virtual bee function in the recurrent model has helped to gain the maximum prediction outcome. The Ernakulum, Kerala region was initially mapped, and the features were extracted. Consequently, the fold area has been mapped in three cases: high flood region, average flood region, and low flood region. Also, the planned framework is executed Python GEE platform. Hence, the presented novel VBRM has gained 99.1% of flood prediction rate and severity classification score respectively. It has maximised the detection score by 2% compared to other models. Therefore, the designed framework is suitable for flood prediction applications in any region. However, this model requires more parameter features for predicting the flood affected that has maximised the resource usage. In the future, designing a hybrid multi-modal feature analysis system will afford the most exemplary predicting results within minimum resource cost.

References

Abdollahi, A. and Pradhan, B. (2023) 'Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model', *Science of the Total Environment*, Vol. 879, No. 1, p.163004.

Ackerson, J.M., Rushit, D. and Naeem, S. (2021) 'Applications of recurrent neural network for biometric authentication & anomaly detection', *Information*, Vol. 12, No. 7, p.272.

Ali, S.A. et al. (2019) 'Application of GIS-based analytic hierarchy process and frequency ratio model to flood vulnerable mapping and risk area estimation at Sundarban region, India', *Modeling Earth Systems and Environment*, Vol. 5, No. 3, pp.1083–1102.

Al-Juaidi, A.E.M., Ayman, M.N. and Omar, E.M.A. (2018) 'Evaluation of flood susceptibility mapping using logistic regression and GIS conditioning factors', *Arabian Journal of Geosciences*, Vol. 11, No. 24, pp.1–10.

Amitrano, D. et al. (2018) 'Unsupervised rapid flood mapping using Sentinel-1 GRD SAR images', *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 56, No. 6, pp.3290–3299.

Brunner, M.I. et al. (2021) 'Challenges in modeling and predicting floods and droughts: a review', *Water*, Vol. 8, No. 3, p.e1520, Wiley Interdisciplinary Reviews.

Bui, D.T. et al. (2019) 'A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping', *Catena*, Vol. 179, No. 1, pp.184–196.

Costache, R. (2019) 'Flood susceptibility assessment by using bivariate statistics and machine learning models-a useful tool for flood risk management', *Water Resources Management*, Vol. 33, No. 9, pp.3239–3256.

Darabi, H. et al. (2019) 'Urban flood risk mapping using the GARP and QUEST models: a comparative study of machine learning techniques', *Journal of Hydrology*, Vol. 569, No. 1, pp.142–154.

Dong, Z. et al. (2021) 'Monitoring the summer flooding in the Poyang Lake area of China in 2020 based on Sentinel-1 data and multiple convolutional neural networks', *International Journal of Applied Earth Observation and Geoinformation*, Vol. 102, No. 1, pp.102–400.

Fahy, B. et al. (2019) 'Spatial analysis of urban flooding and extreme heat hazard potential in Portland, OR', *International Journal of Disaster Risk Reduction*, Vol. 39, No. 1, pp.101–117.

Fassoni, A.AC. and Rodrigo, C.D.P. (2019) 'Mapping spatial-temporal sediment dynamics of river-floodplains in the Amazon', *Remote Sensing of Environment*, Vol. 221, No. 1, pp.94–107.

Jahangir, M.H., Seyedeh, M., Mousavi, R. and Mahnaz, A. (2019) 'Spatial predication of flood zonation mapping in Kan River Basin, Iran, using artificial neural network algorithm', *Weather and Climate Extremes*, Vol. 25, No. 1, pp.100–215.

Jha, R.K. and Haripriya, G. (2019) 'An integrated assessment of vulnerability to floods using composite index – a district level analysis for Bihar, India', *International Journal of Disaster Risk Reduction*, Vol. 35, No. 1, pp.101–174.

Jhong, Y.D. et al. (2022) 'Real-time neural-network-based ensemble typhoon flood forecasting model with self-organizing map cluster analysis: a case study on the Wu River Basin in Taiwan', *Water Resources Management*, Vol. 38, No. 9, pp.1–25.

Kankanamge, N. et al. (2020) 'Determining disaster severity through social media analysis: testing the methodology with South East Queensland Flood tweets', *International Journal of Disaster Risk Reduction*, Vol. 42, No. 1, pp.101–360.

Khoirunisa, N., Cheng, Y.K. and Chih, Y.L. (2021) 'A GIS-based artificial neural network model for flood susceptibility assessment', *International Journal of Environmental Research and Public Health*, Vol. 18, No. 3, p.1072.

Luu, C., Jason, V.M. and Sittimont, K. (2018) 'Assessing flood hazard using flood marks and analytic hierarchy process approach: a case study for the 2013 flood event in Quang Nam, Vietnam', *Natural Hazards*, Vol. 90, No. 3, pp.1031–1050.

Nachappa, T.G. et al. (2020) 'Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory', *Journal of Hydrology*, Vol. 590, No. 1, pp.125–275.

Ngo, P.T. et al. (2021) 'A novel hybrid quantum-PSO and credal decision tree ensemble for tropical cyclone induced flash flood susceptibility mapping with geospatial data', *Journal of Hydrology*, Vol. 96, No. 1, pp.125–682.

Rafiei, S.E. et al. (2021) 'Evaluating urban flood risk using hybrid method of TOPSIS and machine learning', *International Journal of Disaster Risk Reduction*, Vol. 66, No. 1, pp.102–614.

Saravanan, S., Abijith, D., Reddy, N.M., Parthasarathy, K.S.S., Janardhanam, N., Sathiyamurthi, S. and Sivakumar, V. (2023) 'Flood susceptibility mapping using machine learning boosting algorithms techniques in Idukki district of Kerala India', *Urban Climate*, Vol. 49, No. 1, p.101503.

Seejata, K. et al. (2018) 'Assessment of flood hazard areas using analytical hierarchy process over the Lower Yom Basin, Sukhothai Province', *Procedia Engineering*, Vol. 212, pp.340–347.

Shafapour, T.M. et al. (2019) 'Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods', *Geomatics, Natural Hazards and Risk*, Vol. 10, No. 1, pp.79–101.

Shatnawi, N., Matouq, M., Khasawneh, A. and Eslamian, S. (2020) 'Derivation of digital terrain models and morphological parameters from very high resolution satellite images', *International Journal of Hydrology Science and Technology*, Vol. 10, No. 6, pp.515–526.

Singh, P. et al. (2018) 'Vulnerability assessment of urban road network from urban flood', *International Journal of Disaster Risk Reduction*, Vol. 28, No. 1, pp.237–250.

Tang, X. et al. (2020) 'Flood susceptibility assessment based on a novel random Naïve Bayes method: a comparison between different factor discretization methods', *Catena*, Vol. 190, No. 1, pp.104–536.

Toosi, A.S. et al. (2019) 'River basin-scale flood hazard assessment using a modified multi-criteria decision analysis approach: a case study', *Journal of Hydrology*, Vol. 574, No. 1, pp.660–671.

Waghwala, R.K. and Agnihotri, P.G. (2019) 'Flood risk assessment and resilience strategies for flood risk management: a case study of Surat City', *International Journal of Disaster Risk Reduction*, Vol. 40, No. 1, pp.101–155.

Wang, Y. et al. (2020) 'Flood susceptibility mapping using convolutional neural network frameworks', *Journal of Hydrology*, Vol. 582, No. 1, pp.124–482.

Yang, Z., Lu, H., Zhang, Z., Liu, C., Nie , R., Zhang, W., Fan, G., Chen, C., Ma, L., Dai, X. and Zhang, M. (2023) 'Visualization analysis of rainfall-induced landslides hazards based on remote sensing and geographic information system-an overview', *International Journal of Digital Earth*, Vol. 16, No. 1, pp.2374–2402.

Youssef, A.M, Mahdi, A.M. and Pourghasemi, H.R. (2023) 'Optimal flood susceptibility model based on performance comparisons of LR, EGB, and RF algorithms', *Natural Hazards*, Vol. 115, No. 2, pp.1071–1096.

Zhang, Y., Liu, P., Chen, L., Xu, M., Guo, X. and Zhao, L. (2023) 'A new multi-source remote sensing image sample dataset with high resolution for flood area extraction: GF-FloodNet', *International Journal of Digital Earth*, Vol. 16, No. 1, pp.2522–255.

Zhao, G. et al. (2019) 'Assessment of urban flood susceptibility using semi-supervised machine learning model', *Science of the Total Environment*, Vol. 659, No. 1, pp.940–949.

Zhu, K., Lily, D.L. and Michael, L. (2021) 'School timetabling optimization using artificial bee colony algorithm based on a virtual searching space method', *Mathematics*, Vol. 10, No. 1, p.73.