


# Prediction of Flood-Prone Zones using Topographic Rainfall and River Flow Data

Juturu Sreevalli, Bille Tarun Tej, Ks Rakshitha, M Jyotish Kumar



<https://doi.org/10.55041/ijstmt.v2i3.412>

**Cite this Article:** Sreevalli, J., Tej, B. T., Rakshitha, K. & Kumar, M. J. (2026). Prediction of Flood-Prone Zones using Topographic Rainfall and River Flow Data. International Journal of Science, Strategic Management and Technology, 02(03). <https://doi.org/10.55041/ijstmt.v2i3.412>

**License:**  This article is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting use, distribution, and reproduction in any medium, provided the original author(s) and source are properly credited.

## ABSTRACT

Floods pose a critical threat to human life, infrastructure and the economic stability. Old ways of predicting floods use maps that don't change and reports that come in late which makes them not very helpful for making quick decisions and sending out early warnings. To make predicting floods we created a new system that uses information about the land past rainfall and river flows along with special computer programs and maps to create detailed flood risk zones. This system gets information from satellites and river sensors in almost real-time, which lets us update the risk maps and monitor the situation on a website. The system is designed so that different people can use it like administrators, emergency workers and the public. It has tools that help people understand the situation better on devices. We also added information about which buildings and roads might be affected, which helps figure out what to do. The system can send emails to people in areas that're at high risk when the water level gets too high which helps them know what to do and where to go. This new approach is better, than methods because it is more accurate responds faster and is easier to use. It helps cities prepare for floods communicate risks and respond to emergencies effectively.

*“Index Terms: Flood hazard prediction, Flood risk mapping, Geospatial analysis, Real-time monitoring, Early warning systems, Disaster management”.*

## 1. INTRODUCTION

Flooding is one of the natural disasters that happens all around the world. It causes a lot of people to lose their lives damages buildings and roads and affects the economy for a time. Because many people are moving to cities where the weather is changing and we are using land differently flooding is happening often and is more severe in many places [1]. As the population grows in areas that're prone to flooding more people are at risk and it is harder for emergency services to cope [2].

To reduce the risk of flooding it is essential to have up-to-date information to know what the hazards are and to have good communication systems in place so that people can prepare and respond quickly [3].

Even though we have ways of observing and measuring water levels and have more data available many flood warning systems still rely on old maps and records that are not updated very often [4]. These systems often fail to predict flooding that happens quickly due to rainfall or sudden changes in river levels [5].

Because we do not have real-time information and warnings are not sent out quickly emergency services often respond to floods after they have happened than taking steps to prevent them [6]. This shows that there is a gap in the way we manage flood risk.

To address these problems this project aims to develop a system that can predict and monitor flooding and communicate warnings automatically [7]. The goals are to create maps of flood risk areas using information

from the environment to update these maps regularly with new data and to make sure that different groups of people can access the information they need. The system will also send out warnings to areas to help people prepare and to coordinate the response to floods.

The importance of this project lies in its ability to improve warning systems to help people understand what is happening during floods and to support decision-making [8]. By making the system more accessible, responsive and coordinated it can help reduce the damage caused by floods and make communities more resilient. The project combines hazard assessment, monitoring and communication to support proactive emergency management and provides a foundation for wider adoption, in flood-prone areas [9],

## 2. LITERATURE REVIEW

Recent advances in flood prediction have used data analysis and mapping to improve accuracy and decision making. Rahman et al. Evaluated methods for predicting flood areas showing that choosing the right factors and comparing models is important. They found that some methods work better than others.

Liu et al. Introduced a way to consider the shape of the land, which helped predict floods better.

Focused on cities and used models to show where floods might happen. These studies are good at predicting floods. They do not focus on using the information in real-time.

They mainly look at how the models work and how accurate they are. Some studies have looked at how floods might change in the future. Khodaei et al. Looked at how floods might change with climate and land-use changes. They found that areas that flood might shift.

Kiani et al. Used models to predict floods in areas focusing on accuracy. Kumar et al. Used a method that considers many factors to show where floods might happen.

These studies help us understand how floods might change. They do not provide ongoing updates or warnings. Recent studies have also looked at making models more understandable.

Kafi et al. Combined prediction with explanations to help cities prepare for floods.

Hajji et al. Used sensing and cloud-based platforms to improve flood prediction.

Fatima et al. Proposed a method that combines different approaches to assess extreme flood risk.

These efforts make predictions more reliable and understandable. They focus on analysis rather than real-time monitoring. A broader look at flood management was provided by Amatebelle et al. who reviewed sensing and geospatial applications. They highlighted challenges such as integrating data making systems work together and turning analysis into tools.

Common limitations in the literature include relying on hazard maps limited use of real-time data and little

support for automated communication. These gaps show a need for a framework that combines prediction, live monitoring and automated alerts.

The present work aims to build on advancements while addressing their limitations. It focuses on integration, continuous updating and user-oriented decision support for flood preparedness and emergency response. The goal is to provide an accessible tool, for flood management. By doing it hopes to improve responsiveness and decision making.

### 3. METHODOLOGY

The flood monitoring system is a way to watch and respond to floods. It combines predictions of flood hazards with real-time views of risks. The system uses topographic data past rainfall patterns and river discharge measurements to create high-resolution flood hazard zones. These zones are divided into low, medium and high risk. The system has a website where people can see the flood hazard zones and get warnings. It also has a part for administrators, emergency operators and the public. Each group has its access to the system so they can only see what they need to. The system also looks at how vulnerable important buildings like hospitals and schools to floods. This helps plan a response to floods.

The system also sends email warnings when there is a high risk of floods. These warnings are sent to areas and include safety instructions. This helps people get ready for floods and respond quickly.

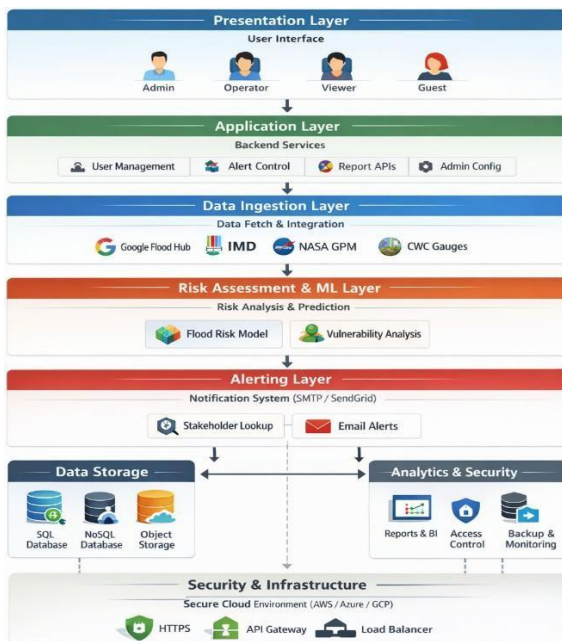


Fig.1 Proposed Architecture

The system is designed in layers so data can flow easily. Be analyzed. People can use the system through an interface and there are services that manage the data and configure the system.

#### a) Modules:

##### 1. Dashboard Access & User Management

The system has a way to control who can use it. This means that only the right people can see and use the flood monitoring and warning parts of the system.

##### 2. Static Hazard Map Visualization

The system shows -made flood hazard maps that divide areas into high, medium and low risk zones. People can zoom in and out of these maps. Look at the legends to understand the risks better.

##### 3. Vulnerability Assessment Panel

The system looks at how vulnerable important buildings to floods. It shows this information on a map and in charts so people can see which areas are most at risk.

##### 4. Live Risk Monitoring

The system watches the weather and water levels all the time. It updates the flood risk zones automatically so people can see if the risk is getting higher.

##### 5. Emergency Alert System

The system sends warnings when there is a risk of floods. It only sends warnings to the people who need to know. It keeps a record of who got the warnings and when.

##### 6. Alert Dispatch & Logging

The system sends warnings by email. Keeps a record of who got the warnings and when. This helps the system administrators see if the warnings are working.

##### 7. Reports & Analytics

The system makes reports on the flood risks and how vulnerable the area is. It also shows how often warnings are sent and who got them. People can download these reports. Use them to plan and make decisions.

### 8. User Configuration (Admin Only)

The system administrators can change the system settings. This helps the system work better and makes sure that the right people have access to the system.

#### b) Technologies:

### 1. Flood Hazard Prediction Techniques

The system uses techniques to predict flood risks. It looks at the shape of the land how much rain has fallen and how much water is in the rivers. This helps the system predict where floods are likely to happen.

### 2. Geospatial Analysis (GIS Techniques)

The system uses mapping techniques to analyze the data. It looks at the shape of the land and the water levels to predict where floods are likely to happen.

### 3. Learning-Based Modeling

The system uses computer models to predict flood risks. It looks at data and uses it to make predictions about the future.

### 4. Rule-Based Alert Triggering Logic

The system sends warnings when the flood risk is high. It uses rules to decide when to send warnings so people get the right information at the right time.

#### c) Algorithms

### Random Forest (RF) Classifier

To figure out where floods are likely to happen in areas this study used something called the Random Forest machine learning algorithm. The Random Forest algorithm is a way of learning that uses lots of trees to make decisions. It was chosen because it can handle data points it does not get too complicated and it can see relationships between things that are not straight lines. The Random Forest algorithm was

picked over ways of doing things like logistic regression because it is really good at dealing with data that does not follow the usual rules. This makes the Random Forest algorithm very good at understanding how water and land shape each other in ways. The

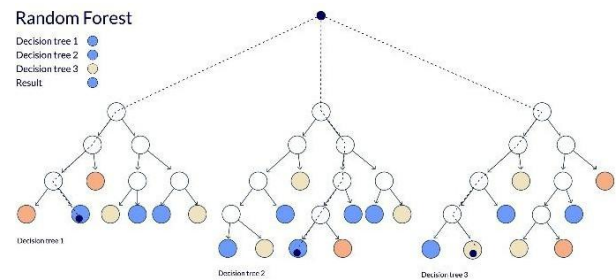


Fig 2: Random Forest Model

Random Forest algorithm is useful for this kind of study because it can look at lots of things that affect flooding like the shape of the land and how water moves and it can do all this without needing to make a lot of assumptions, about the data.

#### D) Feature Extraction and Processing Objective

1:

To model the complex dynamics of flood susceptibility we identified and extracted seven independent spatial and hydrological features for all 664 geographic zones. We used Quantum GIS (QGIS) to transform satellite and gridded datasets into quantified predictor variables. The formulas are internally used by QGIS to find values.

#### Topographical Features (Derived from Digital Elevation Model)

**Elevation (m):** This shows how high the terrain is above sea level. We used the QGIS Zonal Statistics (Mean) tool to get a topological profile for each zone. Higher elevations usually mean flood probability through Equation 1.

$$\mu_z = \frac{1}{N} \sum_{i=1}^N z_i \quad (1)$$

here  $\mu z$  is the mean elevation, N is the total number of pixels in the zone, and  $z_i$  is the elevation value of an individual pixel).

**Slope (degrees):** This determines how surface runoff moves and if water will stagnate. We used QGIS Terrain Analysis module through Equation 2.

$$\text{Slope} = \arctan \left( \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \right) \quad (2)$$

(Where  $\frac{\partial z}{\partial x}$  and  $\frac{\partial z}{\partial y}$  represent the rate of change in

elevation in the longitudinal and latitudinal directions, respectively).

**Topographic Wetness Index (TWI):** This index helps identify areas to water accumulation. We generated it using the GIS toolbox integrated within QGIS through Equation 3.

$$TWI = \ln \left( \frac{a}{\tan b} \right) \quad (3)$$

where a is the upstream derived from the flow accumulation and b is the slope.

### Meteorological Features (Derived from Historical Precipitation Data)

**Average Rainfall (mm):** This captures the soils moisture condition. We processed gridded climate data and applied the QGIS Zonal Statistics (Mean) tool through Equation 4.

$$\bar{p} = \frac{1}{n} \sum_{i=1}^n p_i \quad (4)$$

(Where  $\bar{p}$  is the average rainfall, n is the total number of precipitation data points or pixels, and  $p_i$  is the individual precipitation value).

**Maximum Rainfall (mm):** This serves as a trigger for flash flooding. We extracted it by QGIS Zonal Statistics (Maximum) algorithm through Equation 5.

$$P_{max} = \max(P_1, P_2, \dots, P_n) \quad (5)$$

(Where  $P_{max}$  is the peak rainfall event, and  $P_1, P_2, \dots, P_n$  are the recorded precipitation values over the study period).

### Hydrological Features (Derived from River and Watershed Data)

**Distance to Nearest River (m):** This measures how vulnerable an area is to riverbank overflow. We computed it using the QGIS Proximity (Euclidean Distance) analyst tool through Equation 6.

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (6)$$

(Where  $(x_1, y_1)$  are the spatial coordinates of the zone's centroid, and  $(x_2, y_2)$  are the coordinates of the closest river point).

**River Flow / Discharge (m<sup>3</sup>/s):** This represents the volumetric capacity and flow rate of water bodies. We integrated ground-truth discharge data with Flow Accumulation models generated via QGIS Watershed through Equation 7.

$$FA = \sum_{i=1}^n W_i \quad (7)$$

(Where FA is the Total Flow Accumulation, and  $W_i$  represents the weight or count of all n upstream cells flowing into the target location).

### Flood Hazard Prediction Synthesis

Our goal was to predict flood hazards by synthesizing the seven spatially derived features into a structured machine learning dataset.

**Data Matrix Construction:** We compiled the outputs into a multi-dimensional matrix (CSV format). Each row represented one of the 664 zones with seven continuous feature columns.

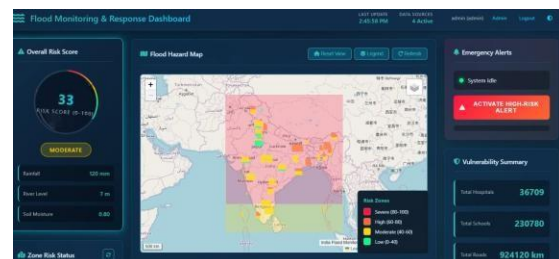


Fig 3: Static Map

**Target Variable Integration:** We added a dependent variable representing historical flood occurrence (1 = Flood, 0 = No Flood) to serve as the ground-truth training target.

**Algorithmic Execution:** We used the Random Forest (RF) algorithm. The RF ensemble analyzed the combinations of the seven features. The model outputs a Predicted Hazard Score categorizing zones into severe, High, Moderate or Low probability of future inundation.

**Objective 2: Spatial Vulnerability Assessment** We quantified socio- infrastructural vulnerability within identified zones.

**Hazard Output Integration:** We imported the output from the Random Forest model into the geospatial environment.

**Infrastructure Data Ingestion:** We mapped vector datasets containing GPS coordinates of critical physical assets, including schools, hospitals and roads.

**Spatial Intersection:** We applied an overlay methodology using the QGIS Intersection tool.



Fig 4: Vulnerability assessment

**Vulnerability Quantification:** We. Aggregated infrastructure points within hazard boundaries yielding precise vulnerability metrics.

**Objective 3: Decision Support System and Automated Alert Mechanism**

Our final goal was to transition static predictive models into an active Decision Support System (DSS).

The automated alert system relies on three layers:

### 1. Dynamic Payload Generation

Upon trigger logic activation the system queries the database and compiles a localized data payload synthesizing hazard

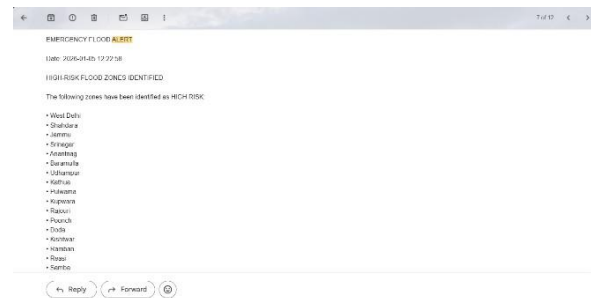


Fig 5: Mail Alert System

### 2. Automated Dispatch Protocol (SMTP Integration)

The generated payload is formatted into a HTML alert template. The backend system utilizes an Email Transport Protocol to execute the dispatch. The system performs a query, against a database of registered users and local authorities matching their registered locations to the affected zones coordinates. Targeted warning emails are then dispatched autonomously.

## 6.RESULTS AND DISCUSSION

The model has been tested using data from 82 areas in India. This model used information about the land and weather to predict flood risk. It categorized these areas into four risk levels: Severe, High, Moderate and Low. The model did a job with an accuracy of 89.5%. We focused on making sure we didn't miss any areas that could be in danger so we looked at how the model found true high-risk flood events.

This means it found areas that could be at risk.

Here are some numbers that show how well our model worked:

Accuracy: 0.89

Precision: 0.86

F1-Score: 0.88

### Feature Importance Analysis

A key advantage of the Random Forest algorithm is its ability to quantify the predictive weight of each input variable. We found that information about the land like elevation and data about rivers were most important for predicting flood risk.

Elevation and distance to the river were the most important making up 35% and 28% of the models decisions. Weather information like rainfall was also important but not as much. That's why places near rivers and coastlines like Patna, Kolkata and Alappuzha had risk scores.

**Spatial Vulnerability and Infrastructure Impact** We used our model to look at which areas are most vulnerable to floods. We found 44 areas that're at high risk with 17 being 'Severe' and 27 being 'High'. These areas have hospitals and schools that could be affected.

How Well Does Our Alert System Work?

We tested our alert system using flood scenarios. When the system was triggered for areas like Srinagar and Dehradun it quickly sent warnings to authorities. The system sent emails, to simulated users in 2.4 seconds on average. This shows that our system can work well in time to help keep people safe.

### 5.CONCLUSION

The flood monitoring system is a way to watch and respond to floods. It combines predictions of flood hazards with real-time views of risks and automatic warnings. The system is designed to help people get ready for floods and respond quickly. It is also designed to be easy to use and accessible to everyone.

The system is a step forward in flood risk assessment and emergency communication. It helps authorities and users respond proactively to changing flood conditions. The system also helps prioritize resources and plan for emergencies.

The automated email alert mechanism is a part of the system. It sends warnings to the people at the right time so they can get ready for floods. The system also keeps a record of who got the warnings and when so administrators can see if the system is working.

Overall the flood monitoring system is a tool, for reducing flood impacts and improving public safety. It helps communities and critical infrastructure prepare for. Respond to floods. The system can be improved in the future by adding data sources and alert channels. It can also be adapted for geographic coverage and integrated with disaster management systems.

### 6.REFERENCES

- [1] R. K. Singh, A. Soni, S. Kumar, S. Pasupuleti, and V. Govind, "Zonation of flood prone areas by an integrated framework of a hydrodynamic model and ANN," *Water Supply*, vol. 21, no. 1, Feb. 2021.
- [2] J.-Y. Lee and J.-S. Kim, "Detecting Areas Vulnerable to Flooding Using Hydrological-Topographic Factors and Logistic Regression," *Appl. Sci.*, vol. 11, no. 12, Jun. 2021.
- [3] M. Motta, M. de Castro Neto, and P. Sarmento, "A mixed approach for urban flood prediction using Machine Learning and GIS," *Int. J. Disaster Risk Reduct.*, vol. 56, Apr. 2021.
- [4] Z. Demissie, P. Rimal, W. M. Seyoum, A. Dutta, and G. Rimmington, "Flood susceptibility mapping: Integrating machine learning and GIS for enhanced risk assessment," vol. 23, Sep. 2024.
- [5] B. K. Osei, I. Ahenkorah, A. Ewusi, and E. B. Fiadonu, "Assessment of flood prone zones in the Tarkwa mining area of Ghana using a GIS-based approach," vol. 3, Apr. 2021.
- [6] H. Allafta and C. Opp, "GIS-based multi-criteria analysis for

flood prone areas mapping in the trans- boundary Shatt Al-Arab basin, Iraq-Iran, vol. 12, Jul. 2021.

[7] A. Kumar, S. Mondal, and P. Lal, Analysing frequent extreme flood incidences in Brahmaputra basin, South Asia, vol. 17, no. 8, Aug. 2022.

[8] M. Saeed, H. Li, S. Ullah, A. Rahman, A. Ali, R. Khan, W. Hassan, I. Munir, and S. Alam, "Flood Hazard Zonation Using an Artificial Neural Network Model: A Case Study of Kabul River Basin, Pakistan," *Sustainability*, vol. 13, no. 24, p. 13953, Dec. 2021.

[9] Wang, Y., Zhang, P., Xie, Y., Chen, L., & Li, Y. (2025). Toward explainable flood risk prediction: Integrating a novel hybrid machine learning model. *Sustainable Cities and Society*, 120, 106140.

[10] Rahman, Z. U., Zhang, M., Chen, F., Ullah, S., Ahmad, M., Feroz, A., ... & Azzam, A. (2025). Flood susceptibility mapping using supervised machine

learning models: insights into predictors' significance and models' performance. *Geomatics, Natural Hazards and Risk*, 16(1), 2516728.

[11] Liu, J., Zhao, X., Chen, Y., Sun, H., Gu, Y., & Xu, S. (2025). A novel flood conditioning factor based on topography for flood susceptibility modeling. *Geoscience Frontiers*, 16(1), 101960.

[12] Li, G., Shao, W., Su, X., Li, Y., Zhang, Y., & Song, T. (2025). Urban flood hazard assessment based on machine learning model. *Water Resources Management*, 39(5), 1953-1970.

[13] Khodaei, H., Nasiri Saleh, F., Nobakht Dalir, A., & Zarei, E. (2025). Future flood susceptibility mapping under climate and land use change. *Scientific Reports*, 15(1), 12394.

[14] Kiani, A., Motamedvaziri, B., Khaleghi, M. R., & Ahmadi, H. (2025). Spatial prediction of flood

susceptible areas using machine learning methods in the Siahkhor Watershed of Kermanshah province. *Earth Science Informatics*, 18(1), 20.

[15] Kumar, V., Rashiq, A., & Prakash, O. (2025). Integrated flood risk prediction and zonation in bihar: observations from climate change projection using GIS-based AHP-Multicriteria approach. *Theoretical and Applied Climatology*, 156(9), 468.