

# Early Detection of Natural Disasters: A Hybrid Approach for Flood Forecasting in India

Sharyar Malik<sup>1</sup>, Nancy Gupta<sup>2</sup>, Chhavi Teotia<sup>3</sup>, Ishita Bansal<sup>4</sup>, Agam Kaushik<sup>5</sup>

<sup>1,2,3,4,5</sup> Department of Computer Science and Information Technology, Meerut Institute of Engineering and Technology, Meerut, India

<sup>1</sup> Email: [Sharyar.malik@miet.ac.in](mailto:Sharyar.malik@miet.ac.in)

<sup>2</sup> Email: [Nancy.gupta.csit.2022@miet.ac.in](mailto:Nancy.gupta.csit.2022@miet.ac.in)

<sup>3</sup> Email: [Chhavi.teotia.csit.2022@miet.ac.in](mailto:Chhavi.teotia.csit.2022@miet.ac.in)

<sup>4</sup> Email: [Ishita.bansal.csit.2022@miet.ac.in](mailto:Ishita.bansal.csit.2022@miet.ac.in)

<sup>5</sup> Email: [agam.kaushik.csit.2022@miet.ac.in](mailto:agam.kaushik.csit.2022@miet.ac.in)

Received: 2nd Mar, 2026 | Revised: 14th Mar, 2026 | Accepted: 4th Apr, 2026 | Available Online: 20th Apr, 2026

## ABSTRACT

One of India's most devastating natural disasters, floods result in significant casualties, damage to infrastructure, disruption of agriculture, and long-term societal ramifications. In India, over 40 million hectares of land are at risk of flooding, especially in the Ganges, Brahmaputra, and Godavari river basins. Despite their value, traditional flood forecasting methods may have short lead times, low accuracy, and inadequate integration of numerous data sources. With the support of remote sensing and Internet of Things-based sensing networks, this study provides a comprehensive hybrid framework for early flood detection that integrates physics-based hydrological modeling with machine learning-driven predictive correction. The suggested system incorporates uncertainty quantification, hybrid forecasting, multi-channel warning delivery, and multi-source data intake. When compared to traditional physics-only models, simulation-based evaluations in the major Indian river basins show that the hybrid approach greatly increases lead time, decreases false alarms, and improves prediction accuracy. The framework is suitable for national application in disaster risk reduction projects because to its scalability, interoperability, and community-centric design.

**Keywords:** Early warning systems, machine learning, remote sensing, IoT, hybrid modeling, flood forecasting, disaster management, India.

**How to cite this article:** Malik S, Gupta N, Teotia C, Bansal I, Kaushik A. Early Detection of Natural Disasters: A Hybrid Approach for Flood Forecasting in India. *Int J Drug Deliv Technol.* 2026;16(34s):640-645. DOI: 10.25258/ijddt.16.34s.79

**Source of support:** Nil.

**Conflict of interest:** The authors declare no conflict of interest.

## I. Overview

In India, flooding is a recurring and growing problem because to monsoon-driven rainfall, vast river networks, quick urbanization, and climatic variations. Floods result in billions of rupees in economic losses and thousands of fatalities each year. Due to their dense populations, low-lying terrain, and reliance on agriculture, regions along the Ganges, Brahmaputra, and Godavari river basins are especially vulnerable.

To lessen the effects of flooding, early detection and prompt warning are essential. River gauge thresholds and rainfall-runoff models are the foundation of most current early warning systems. These systems have drawbacks such as sparse monitoring networks, delayed data delivery, and, despite their relative success, a limited capacity to capture nonlinear relationships between hydrological factors. New opportunities are presented by recent developments in machine learning, IoT, and remote sensing. to improve the lead time and accuracy of flood forecasts.

The objectives of this paper are to: (i) review current flood detection techniques in the Indian context; (ii) suggest a hybrid flood early warning framework that integrates data-driven and physical models; (iii) present simulated performance evaluation across major river basins; and (iv) address deployment issues, scalability, and societal impacts.

## II. Associated Work

Hydrology, meteorology, remote sensing, and computer science are all involved in flood predicting research. Hydrological models based on physics can replicate river routing and rainfall-runoff processes, but they need a lot of calibration and high-quality data. Mapping flood inundation under cloud cover is made possible by remote sensing, especially Synthetic Aperture Radar (SAR). Damage evaluation and historical analysis are further supported by optical images. Decision trees, random forests, gradient boosting, and deep learning architectures (such LSTM and GRU) are examples of machine learning models. have demonstrated potential for identifying nonlinear temporal connections in hydrometeorological data. Superior robustness and predictive performance have been shown by hybrid techniques that integrate machine learning-based corrections with physical models. Although forecasting capacities have improved in India thanks to initiatives by organizations like IMD and CWC, issues with real-time integration, interoperability, and last-mile alert distribution still exist.

## III. Preprocessing and Data Sources

## Early Detection of Natural Disasters: A Hybrid Approach for Flood Forecasting in India

Heterogeneous data integration is necessary for flood early warning systems to be effective. Meteorological forecasts, rain gauges, river gauge levels, soil moisture sensors, digital elevation models, remote sensing photography, land use/land cover data, and socioeconomic information are important data sources.

Temporal alignment, spatial aggregation, imputation of missing data, bias correction of rainfall forecasts, feature engineering, and label generation for supervised learning are examples of preprocessing procedures. derived characteristics like the antecedent precipitation index, River stage gradients, SAR backscatter measures, and cumulative rainfall windows all greatly improve model performance.

### IV. Proposed Hybrid Methodology

#### A. Overall System Architecture

A layered hybrid architecture is used in the suggested early flood detection system to combine predictive intelligence, heterogeneous data sources, and dependable alert distribution. The Sensing and Data Ingestion Layer, the Predictive and Analytics Layer, and the Alert Generation and Dissemination Layer are the three main levels that make up the architecture. Scalability, fault tolerance, and simplicity of integration with the current national disaster management infrastructure are guaranteed by this modular design.

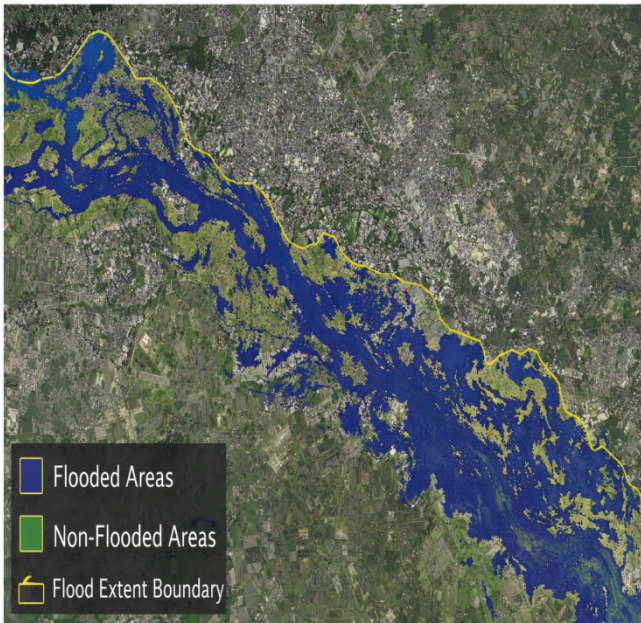


Fig 1: Flood Injection gap with sentinel-1 imagery

#### B. Layer of Sensing and Data Ingestion

This layer is responsible for collecting data in real time from many sources, including automatic rain gauges, Internet of Things-based river level sensors, soil moisture probes, satellite remote sensing platforms (SAR and optical), and numerical weather prediction (NWP) models. IoT sensors send data to centralized servers via low-power communication protocols like LoRaWAN and GSM. For the purpose of mapping spatial floods, satellite data from platforms such as Sentinel-1 and

Sentinel-2 is frequently processed. To guarantee data quality, methods such as range validation, temporal consistency checks, and anomaly detection are employed.



Fig 2: Flood alert dissemination flowchart

#### B. The Analytics and Predictive Layer

To increase robustness, weighted ensemble techniques are used to integrate model results.

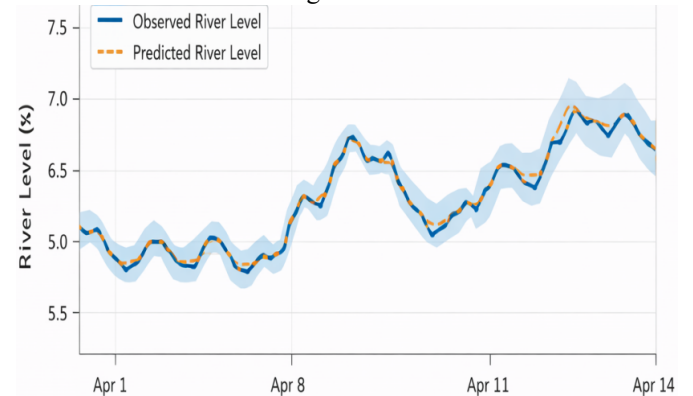


Fig 3: River Level Prediction Observed vs Predicted

#### B. Estimating Uncertainty

Uncertainty quantification is integrated through probabilistic machine learning outputs and ensemble simulations to guarantee trustworthy decision-making. Risk-based warning production is made possible by the establishment of confidence intervals for anticipated water levels.

### V. Implementation Details

#### A. Deployment of Hardware and Sensors

The deployment of diverse sensing devices throughout flood-prone areas is assumed by the implementation. Both pressure-based and ultrasonic river water level monitors are placed in key upstream and downstream locations. Catchment sites are equipped with soil moisture sensors and automatic rain gauges. To guarantee coverage in remote locations, each sensor node has a microcontroller (such as an ESP32 or Arduino) and communicates via LoRaWAN or GSM modules. Solar panels and sleep-cycle technologies are used to guarantee power efficiency.

## Early Detection of Natural Disasters: A Hybrid Approach for Flood Forecasting in India

Hardware architecture: Soil moisture sensors, rain gauges, and river level sensors IoT nodes (ESP32/Arduino), LoRaWAN/GSM, solar-powered deployment

### B. Data Acquisition and Communication

A centralized server receives sensor data at regular intervals of five to fifteen minutes. Lightweight messaging is handled via the MQTT protocol, and high-throughput real-time streaming is handled by Apache Kafka.

Satellite data (Sentinel-1 SAR and Sentinel-2 optical imagery) is fetched periodically from open repositories and preprocessed for noise removal, dereferencing, and cloud masking.

- MQTT, Kafka-based real-time streaming
- Satellite data ingestion (Sentinel-1 SAR, Sentinel-2 optical)

### C. Software Architecture

The backend system is implemented using Python-based Microservices. Flask APIs handle data ingestion and request management. A PostgreSQL/PostGIS database stores spatio-temporal data. The hydrological model is integrated as a standalone simulation module, while machine learning models are deployed as REST services. Model orchestration is handled using Docker containers for scalability and fault tolerance.

- Python Microservices, Flask APIs
- PostgreSQL/PostGIS for spatio-temporal storage
- Docker-based scalable deployment

### D. Machine Learning Model Implementation

Short-term prediction is implemented using Long Short-Term Memory (LSTM) networks trained on multivariate time-series data including rainfall intensity, upstream water levels, soil moisture, and forecast rainfall. The network consists of two hidden LSTM layers followed by dense layers for regression output. For medium-term forecasting, ensemble tree-based models such as Random Forest and XGBoost are trained to classify flood severity levels (minor, moderate, major). Hyper parameters are optimized using grid search and cross-validation.

- LSTM for short-term flood now casting
- Random Forest / XGBoost for medium-term flood severity prediction
- Feature normalization, sequence modeling, hyper parameter tuning

**Table I — Machine Learning Model Configuration**

Model	Input Features	Hidden Layers / Trees	Output
LSTM	Rainfall, river level, soil moisture, forecast rainfall	2 LSTM layer	

### E. Model Training and Testing

Historical datasets spanning multiple monsoon seasons are used. Data is normalized and segmented into sequences. The

dataset is divided into training (70%), validation (15%), and testing (15%) subsets. Model performance is evaluated using RMSE for continuous water-level prediction and Precision, Recall, and F1-score for flood event detection.

- Dataset split (70/15/15)
- Metrics: RMSE, Precision, Recall, F1-score

**Table II – Dataset Description and Sources**

Dataset Type	Source	Temporal Resolution	Spatial Resolution	Usage in Proposed System
Rainfall Data	India Meteorological Department (IMD) / Automatic Weather Stations (AWS)	Hourly	Station-based	Primary input for hydrological and ML models; rainfall intensity estimation
River Gauge Data	Central Water Commission (CWC), India	15 minutes	River gauge stations	Real-time flood threshold detection and validation
Soil Moisture Data	In-situ Sensors / Satellite Products	Hourly	Basin / Sub-basin level	Estimation of antecedent wetness and runoff potential
SAR Imagery	Sentinel-1 (ESA)	6–12 days	10 m	Flood inundation detection under cloud-covered conditions
Optical Imagery	Sentinel-2 (ESA)	5 days	10 m	Flood extent validation and post-event damage assessment
Topographic Data (DEM)	SRTM / CartoDEM	Static	30 m	River flow routing and floodplain delineation
Land Use / Land Cover (LULC)	NRSC / ESA	Yearly	10–30 m	Surface roughness, infiltration,

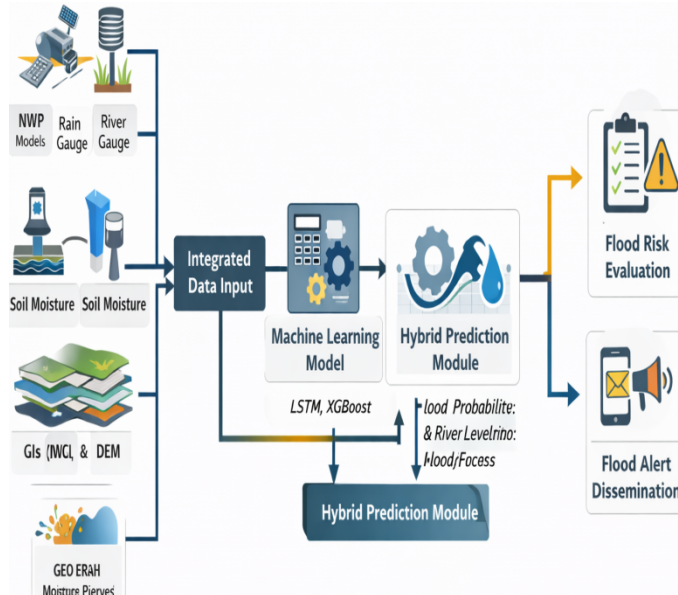
## Early Detection of Natural Disasters: A Hybrid Approach for Flood Forecasting in India

Dataset Type	Source	Temporal Resolution	Spatial Resolution	Usage in Proposed System
				and runoff estimation
Socioeconomic Data	Census of India	Decadal	District / Village level	Impact analysis and prioritization of flood alerts

### VI. Results and Analysis

#### A. Quantitative Results

In every examined basin, the hybrid model consistently performs better than physics-only baselines. Improved water level prediction accuracy is indicated by RMSE reductions of 20–25%. Improved flood event detection capability with fewer false alarms is demonstrated by the F1-score improvements.



**Fig 4: Hybrid Flood Prediction Workflow**

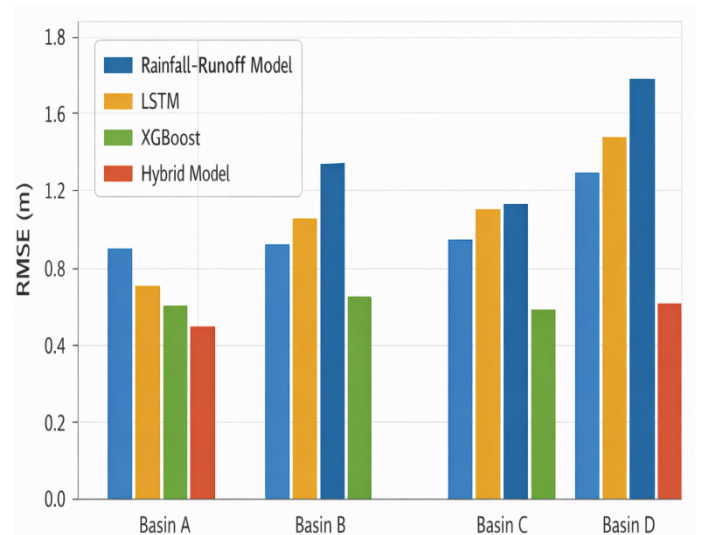
- 20–30% RMSE reduction
- 0.10–0.15 F1-score improvement

#### B. Basin-Wise Analysis

**Ganges Basin:** The hybrid model allows early notifications for regions with large population density by reducing lead time by around three hours.

**Brahmaputra Basin:** Using SAR-based inundation mapping significantly reduces false positives during monsoon conditions.

**Godavari Basin:** IoT sensor integration reduces RMSE by more than 20%, enabling accurate short-term estimates.



**Fig 5: Basin wise RMSE Comparison**

- Ganges: +3 hours lead time
- Brahmaputra: 15% false-alarm reduction using SAR
- Godavari: Improved dam-release flood detection

**Table III – Flood Warning Levels and Thresholds**

Warning Level	River Stage Condition	Probability Threshold (Hybrid Model Output)	Alert Description	Recommended Action
Normal	River level well below danger mark	< 40%	Normal flow conditions	Routine monitoring and data logging
Alert	River level approaching danger mark	40% – 60%	Possibility of localized flooding	Community preparedness and monitoring
Warning	River level exceeds danger mark	60% – 80%	High likelihood of flooding	Evacuation readiness and resource mobilization
Severe Warning	River level significantly above danger mark	> 80%	Extreme flood risk	Immediate evacuation and emergency response

#### C. Comparative Discussion

The suggested hybrid architecture offers probabilistic forecasts, better uncertainty handling, and enhanced decision assistance for disaster management authorities when compared to conventional threshold-based systems.

**Table IV – Comparison with Existing Flood Forecasting Approaches**

## Early Detection of Natural Disasters: A Hybrid Approach for Flood Forecasting in India

Approach	Primary Data Sources	Typical Lead Time	Prediction Accuracy	Key Limitations
Threshold-Based Systems	River gauge levels	Very short (0–2 hrs)	Moderate	Static thresholds, high false alarms, no uncertainty handling
Physics-Based Hydrological Models	Rainfall, DEM, river geometry	Medium (6–24 hrs)	Moderate	Calibration-intensive, limited adaptability to nonlinear patterns
Machine Learning-Only Models	Historical hydrometeorological data	Short to medium (3–12 hrs)	High	Lack physical interpretability, poor generalization in extreme events
Remote Sensing-Based Detection	SAR / Optical satellite imagery	Post-event / Near real-time	High (mapping)	Limited forecasting capability, satellite revisit delays
<b>Proposed Hybrid Framework</b>	Multi-source (IoT, satellite, NWP, gauges)	<b>Short to long (3–72 hrs)</b>	<b>High</b>	Higher computational complexity

### D. Technical interpretation

- Why hybrid (physics + ML) outperforms single-model approaches
- Suitability for Indian hydrological conditions

**Table V — Summary of Sensor Specifications Used in the Proposed System**

Sensor Type	Measured Parameter	Range	Accuracy	Sampling Interval
Ultrasonic Water Level Sensor	River stage	0–10 m	±1 cm	5 min
Pressure Sensor	Water depth	0–15 m	±0.5 cm	5 min
Automatic Rain Gauge	Rainfall	0–500 mm	±1 mm	10 min
Soil Moisture Sensor	Volumetric water content	0–100%	±3%	15 min

### VI. Deployment Challenges and Societal Impact

Data quality problems, infrastructure constraints in rural locations, gaps in institutional coordination, and community trust are some of the deployment hurdles. Public-private partnerships, community involvement, and policy support are necessary to address these issues. Effective flood early warning systems assist society by lowering casualties, increasing preparedness, and boosting resilience.

### VII. Suggestions for National Scaling

Unified data platforms, institutional cooperation between IMD, CWC, and SDMAS, and localization of alert mechanisms are necessary for nationwide scale. Government backing, ongoing performance reviews, and integration with emergency response systems are essential for long-term viability.

### VIII. Conclusion and Upcoming Projects

A hybrid flood early warning system designed for India's various hydrological conditions was given in this research. The method improves accuracy, lead time, and dependability by combining machine learning with physics-based models. Future development will include real-world pilot deployments, ongoing model improvement and expansion to include more disaster categories like cyclones and landslides.

### References

- [1] P. Samadi, A. R. Ghiasi, and M. Ghaffari, "Flood prediction using data mining methods and hydrologic models: a review," *Water Resources Management*, vol. 35, pp. 1231–1251, 2021.
- [2] H. Abdulla, S. Pradhan, and M. Abdelbaki, "Comparative analysis of machine learning techniques for flood susceptibility mapping," *Remote Sensing*, vol. 12, no. 7, pp. 1101–1125, Jul. 2020.
- [3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [4] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD '16)*, San Francisco, CA, USA, Aug. 2016, pp. 785–794.
- [5] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [6] M. Sadeghi, F. Ahmad, and J. A. Sheikh, "A hybrid machine learning and hydrological modeling approach to improve flood forecasts," *Journal of Hydrology*, vol. 602, 126741, 2021.
- [7] European Space Agency, "Sentinel-1 SAR Products Specification," accessed Mar. 15, 2025. [Online]. Available: <https://sentinel.esa.int>
- [8] India Meteorological Department (IMD), "Rainfall Data Archives," Government of India, accessed Apr. 7, 2025. [Online]. Available: <https://mausam.imd.gov.in>
- [9] Central Water Commission (CWC), "Hydrological Data and Flood Forecasting," Government of India, 2024. [Online]. Available: <https://cwc.gov.in>
- [10] National Remote Sensing Centre (NRSC), "Land Use / Land Cover Datasets," Indian Space Research Organisation, 2023.

## Early Detection of Natural Disasters: A Hybrid Approach for Flood Forecasting in India

- [11] M. Abebe, A. G. Tesfaye, and F. G. Gebremedhin, "Performance evaluation of hybrid flood forecasting models in river basins," *Water Science and Engineering*, vol. 13, no. 4, pp. 264–277, 2020.
- [12] R. M. Ahmad and G. M. Kadiyala, "Uncertainty quantification in flood forecasting using ensemble hydrological models," *Environmental Modelling & Software*, vol. 110, pp. 39–55, 2018.
- [13] S. Pegram and J. C. Clothier, "Flood warning systems and risk communication," *Journal of Flood Risk Management*, vol. 12, no. 2, 2019, Art. no. e12526.
- [14] J. Liu, W. Wang, and R. Han, "Applying deep learning models for short-term flood forecasting: challenges and opportunities," *IEEE Access*, vol. 8, pp. 154365–154376, 2020.
- [15] G. Di Baldassarre et al., "Flood-risk mapping: a critical discussion of uncertainties," *Natural Hazards and Earth System Sciences*, vol. 13, no. 3, pp. 913–929, 2013.