

Evaluating the Performance of MaxEnt and Frequency Ratio Models for Flood Susceptibility Mapping in Kerala State, India

M S Kendagannaswamy¹, C K Roopa² and B S Harish³

¹⁻³Department of Information Science and Engineering, JSS Science and Technology University, Mysuru, 570006, Karnataka, India.

Email: kswamy@jssstuniv.in, ckr@jssstuniv.in, bsharish@jssstuniv.in

Abstract—Floods are the greatest catastrophic natural disasters on a global scale. Identifying flood-prone areas is imperative for preventing floods, reducing risks, and making informed decisions. Flood susceptibility mapping plays a crucial role in flood risk assessment and management.

While machine learning shows promise in flood prediction, a significant research gap remains in effectively integrating advanced algorithms with Kerala's diverse regional flood data to develop more accurate prediction models.

The main objective of the proposed research is to develop, implement, and compare the performance of Maximum Entropy (MaxEnt) and Frequency Ratio (FR) models for flood susceptibility mapping of Kerala state using a spatial flood database integrated within the ArcGIS interface. The research employs two paired approaches, the MaxEnt and FR, to assess flood susceptibility levels within the study region. The results show that both models achieved strong predictive performance; AUC was 0.928 for MaxEnt and 0.946 for the FR model, indicating its enhanced capability in flood susceptibility evaluation. The MaxEnt model is a promising tool for assessing flood-prone areas and allows for proper planning and management of flood hazards.

Keywords— Flood Susceptibility, Frequency Ratio, Susceptibility, GIS, MaxEnt, Machine Learning Algorithm.

I. INTRODUCTION

Heavy floods that hit Kerala in 2018-2019 destroyed life and infrastructure across the state. Flooding is more dangerous to Kerala because the Western Ghats Mountains and 44 rivers intersect across its southwestern coastal area. The state's flood vulnerability arises from being a transition point for two regional winds, plus its dense population and active land usage patterns (Aju et al., 2024).

In 2018 the devastating floods in Kerala transformed state history by claiming 400 lives and causing \$3.8 billion worth of damage through effects on the economy. This major climate crisis hit 5.4 million residents and destroyed over 100,000 buildings on a scale never seen before in Kerala structures (Dhasmana et al., 2023). The flood showed weak spots in how we assess and manage flood risk when cities grow fast and weather patterns change over time (Lal et al., 2020). The 2018 flood disaster in Kerala state revealed several weaknesses in how we handle flood control today.

Research shows that disaster response failed to address important issues with environmental factors and changing land use patterns in flood assessments (Mathew et al., 2023). Furthermore, the occurrence of this event underscored the presence of shortcomings in early warning systems and evacuation planning protocols, thereby emphasizing the urgent need for enhanced accuracy and spatial specificity in flood susceptibility mapping (Ifejube et al., 2024).

Modern geographic and remote sensing technologies combined with machine learning methods enable better flood prediction systems according to (Fischer et al., 2023). Researchers are developing new methods combining advanced tech and traditional statistics to boost flood vulnerability studies and responses (Zeng et al., 2019).

With these developments, it is relevant to assess the recent events of the Maximum Entropy (MaxEnt) and Frequency Ratio (FR) models (Siahkamari et al., 2018). The techniques combine several environmental inputs to produce probabilistic maps that indicate the areas at risk of a flood. Accurate flood susceptibility maps are needed to assist flood risk management methods in urban planning (Hussain et al., 2023). These maps, as vital instruments being used in the enforcement of spatial planning and zoning regulations, safeguarding and development of infrastructure, and defining concurrent mass evacuation strategy, are an exhaustive and objective representation through which circumstances and hot spots of high risk can be pinpointed (Hofmann et al., 2021).

II. LITERATURE REVIEW

The two methodological approaches implemented in mapping flood susceptibility using machine learning methods are present. The first approach method uses classification algorithms (Khosravi et al., 2018) to generate flood maps; in the second approach, regression models (Lee et al., 2017) are developed to create flood prediction maps. As shown by (Morshed et al., 2022) study, logistic regression successfully created flood maps for Bangladesh with a 78 to 82 % precision. However, more research is needed to fully develop and validate the application of a regression-based approach to flood susceptibility mapping.

The research by (Menon et al., 2024), who showed that machine learning was up to 15-20% more accurate at predicting floods than historical forecasting techniques. One of the most popular methods for achieving more than 80% accuracy and ranking the key feature importance is the ensemble strategy and excellent handling of mixed type variables to choosing extensive computational resources and SVM with 85% accuracy but cannot scale up for large applications (Kumar et al., 2022; Singh, 2019).

The MaxEnt is considered a valuable tool for natural hazard mapping because of the scientific mapping due to MaxEnt mathematical statistics, a tool for analyzing the occurrence data and giving out strong results even with incomplete information about environmental connections (Kawami & Miura, 2022). While MaxEnt has been utilized in landslide susceptibility mapping, such an application to flood mapping for different geographical settings has been limited due to the successful applications in flood mapping, which have been found to outperform in data-poor settings. The Frequency Ratio (FR) model in susceptibility mapping has moved from a trivial but also effective approach based on the use of the probability theory ((Sharma et al., 2024) to a widely applied statistical method of natural hazard susceptibility assessment. It is demonstrated that the method has been used recently in landslide susceptibility mapping applications and is increasingly used for flood susceptibility applications, achieving consistently good performance across regions. However, FR has been increasingly used with other methods to further enhance the results ((Huangfu et al., 2024)).

This literature research reviews the availability of the MaxEnt and FR models, standard evaluation metrics, environment with multiple environmental factors, regional-specific model optimization, and more advanced model validation frameworks. This research uses both MaxEnt and FR models to develop an improved evaluation framework for flood map susceptibility mapping of Kerala state that provides a methodological approach for a similar geographical setting.

III. STUDY AREA AND DATA

A. Study Area

Kerala is a state along India's southwestern coast at 8° 18'N and 12° 48'N latitude and 74° 52'E and 77° 22'E longitude. The state occupies an area of about 38,863 km² and contains three separate physiographic divisions running parallel to the coastline.

B. Data Collection

Multiple datasets were integrated to ensure a comprehensive analysis of flood-prone areas detailed in Table 1. Terrain characteristics, including slope and aspect, and flow accumulation patterns, were analyzed using Digital

Elevation Model (DEM) data from the Shuttle Radar Topography Mission (SRTM) at 30-meter resolution. Multi-temporal optical imagery of Landsat (30m) and Sentinel-2 (10m) satellites is processed to detect and classify the surface feature changes, urban growth patterns and natural drainage networks. Precipitation analysis included Tropical Rainfall Measuring Method (TRMM) and Global Precipitation Measurement (GPM) satellite rainfall measurements to uncover rainfall patterns and rates. Synthetic Aperture Radar (SAR) data from Sentinel-1 satellites, all-weather monitoring capabilities were particularly beneficial during monsoon seasons when optical imagery as its limited due to coverage by clouds. All datasets were projected to UTM Zone 43N coordinate system and resampled to match spatial resolutions.

TABLE I: INFLUENCING FACTORS DESCRIPTION

#	Parameters	Data source
1	DEM	NASA (Shuttle Radar Topography Mission (SRTM))
2	Slope	DEM
3	Profile and Plan Curvature	DEM
4	Aspect	DEM
5	TWI	DEM
6	SPI	DEM
7	TRI	DEM
8	STI	DEM
9	LULC	Esri Sentinel-2 Land Cover
10	NDVI	Landsat 8 from GEE
11	Rainfall	CRU TS monthly high-resolution gridded multivariate climate dataset
12	Geology	Bhukosh
13	Geomorphology	Bhukosh
14	Soil	National Bureau of Soil Survey and Land Use Planning
15	Distance from Road Network	OpenStreetMap
16	Distance from Stream Network	OpenStreetMap
17	Drainage Density	Hydro SHEDS

C. Flood Conditioning Factors

Understanding and predicting flood susceptibility, along with identifying and analyzing flood conditioning factors, represent essential processes in flood hazard assessment.

The integration of flood conditioning factors derived from multiple satellite platforms and remote sensing technologies provides crucial input data.

The analysis encompasses eighteen key flood conditioning factors, examining their theoretical foundations and remote sensing acquisition methodologies.

D. Topographic Factors

➤ *Profile and Plan Curvature* represents the horizontal component of flow convergence and divergence across a landscape. This factor is derived from high-resolution Digital Elevation Models (DEMs), primarily SRTM data at 30m resolution. The mathematical expression for plan curvature is:

$$\text{Plan Curvature} = -\frac{\delta^2 z}{\delta x^2} \cdot \left(\frac{\delta y}{\delta x}\right)^2 + 2 \frac{\delta^2 z}{\delta x \delta y} \cdot \frac{\delta x}{\delta s} \cdot \frac{\delta y}{\delta s} - \frac{\delta^2 z}{\delta y^2} \cdot \left(\frac{\delta x}{\delta s}\right)^2 \quad (\text{eq. 1})$$

$$\text{Profile Curvature} = -\frac{\delta^2 z}{\delta x^2} \cdot \left(\frac{\delta y}{\delta x}\right)^2 - 2 \frac{\delta^2 z}{\delta x \delta y} \cdot \frac{\delta x}{\delta s} \cdot \frac{\delta y}{\delta s} - \frac{\delta^2 z}{\delta y^2} \cdot \left(\frac{\delta x}{\delta s}\right)^2 \quad (\text{eq. 2})$$

Positive values indicate divergent flow (ridges), while negative values suggest convergent flow (valleys). The coordinate system comprises three dimensions: z represents elevation (height) above a reference datum, while x and y define the horizontal plane, representing east-west and north-south distances, respectively (Florinsky, 1998).

➤ *Slope* is fundamental topographic parameter, directly influences runoff velocity and infiltration rates. Derived from SRTM DEM data, the slope is calculated as:

$$\text{Slope}(\%) = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \times 100 \quad (\text{eq. 3})$$

In topographic analysis, elevation (z) represents the vertical height above a reference plane, while x and y coordinates denote horizontal distances in the east-west and north-south directions, respectively. The partial

derivatives $\partial z/\partial x$ and $\partial z/\partial y$ characterize the terrain's rate of change in elevation with respect to these horizontal directions (Horn, 1981).

E. Hydrological Indices

- *Topographic Wetness Index (TWI)* derived from DEM data, predicts zones of saturation and runoff generation:

$$TWI = \ln \left(\frac{A_s}{\tan \beta} \right) \quad (\text{eq. 4})$$

Where 'A_s' represents the specific catchment area and β is the local slope angle (Sørensen et al., 2006).

- *Topographic Ruggedness Index (TRI)* is a measure of terrain heterogeneity. equation (5) is the standard mathematical expression:

$$TRI = \sqrt{\sum_{i=1}^n (z_0 - z_i)^2} \quad (\text{eq. 5})$$

The mean elevation difference between a central cell (z_0) and its adjacent cells (z_i), where i ranges from 1 to n . In standard implementations, n represents the eight neighbouring cells in a 3×3 moving window, forming the fundamental spatial unit for ruggedness assessment (Sorensen et al., 2006).

- *Stream Power Index (SPI)* measures the amount of erosion power of flowing water:

$$SPI = A_s \times \tan \beta \quad (\text{eq. 6})$$

The specific catchment area tells us how much land releases water upstream toward a 1-meter-long contour length. This index shows water erosion strength by uniting A_s and β values in one metric (Moore et al., 1991).

- *The Sediment Transport Index (STI)*, which is also known as the Modified Stream Power Index, can be expressed mathematically as:

$$STI = \left(\frac{A_s}{22.13} \right)^{0.6} \times \left(\frac{\sin \beta}{0.0896} \right)^{1.3} \quad (\text{eq.7})$$

The catchment area (A_s), measured in square meters per meter (m^2/m), and the local slope angle (β) in degrees. The equation includes two standardization factors derived from the Universal Soil Loss Equation (USLE): 22.13 and 0.0896 (Moore et al., 1991).

F. Vegetation and Land Cover Parameters

- *Normalized Difference Vegetation Index (NDVI)* is calculated using multispectral data from Landsat-8 OLI satellites, indicating the density of vegetation:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (\text{er. 8})$$

The NDVI uses spectral reflectance measurements from the Near-Infrared (NIR) and red bands. For the Landsat-8 Operational Land Imager (OLI), these correspond to Band 5 (0.85-0.88 μm) and Band 4 (0.64-0.67 μm), respectively. Higher NDVI values are generally associated with lower susceptibility to flooding due to increased infiltration and reduced surface runoff (Rouse Jr. et al., 1974).

- *Land use/Land Cover (LULC)* is a complex remote sensing technique that uses supervised classification of multispectral satellite imagery. The images were retrieved from Sentinel-2 stars through a classification process. This classification process is based on spectral signatures and patterns of the Earth. It organizes textures into different categories. The main classification categories include built-up areas (including urban and infrastructure development), agricultural areas (including arable land and agricultural fields), forest cover areas (including various types of canopies and forest), water bodies (rivers, lakes and reservoirs) and bare land (unprotected land).

G. Geological and Soil Factors

- *Geological* features are characterized by an integrated mapping method that combines multiple data sources and monitoring methods. The process begins with the analysis of ASTER satellite images, which are specifically designed for mineral mapping through special spectral bands that can detect unique mineral signatures on the Earth's surface.
- *Soil* characteristics assessment involves a comprehensive approach that combines advanced remote sensing techniques with traditional landscape mapping techniques. It starts with the analysis of Sentinel 1 SAR (Synthetic Aperture Radar), which can analyze soil moisture through its ability to penetrate cloud cover and

vegetation. Spectral indices derived from live satellite imagery complement this information, e.g. identifying broad areas viz., particle content, shape, or mineral interactions using the selected bandwidth and selected counting algorithm.

H. Hydrological Network Parameters

➤ *Drainage Density* is calculated using DEM-derived stream networks:

$$DD = \frac{\sum L}{A} \quad (\text{eq. 9})$$

Where L represents the total stream length and A is the watershed area.

➤ *Distance to Drainage* is derived from the DEM using hydrological analysis capabilities in ArcGIS software. Eliminating any depressions in the DEM to guarantee an uninterrupted water flow. Designating flow direction for each cell in the Digital Elevation Model to ascertain the movement of water throughout the terrain (Costabile et al., 2022).

I. Meteorological Factors

Annual rainfall forecast uses a multi-data system. It combines satellite rainfall measurements with ground-based observations. The TRMM (Tropical Rainfall Measuring Mission) satellite and its successor, the GPM (Global Precipitation Measuring) satellite, serve as the starting point for the project, providing fixed spatial data about Rainfall conditions. Rain is measured from the sky using unique rain gauges.

IV. MATERIALS AND METHODS

The methodology for flood susceptibility assessment is outlined in Figure 1. The workflow demonstrates the integration of 18 flood conditioning factors with flood inventory points, which are partitioned into training (80%) and validation (20%) sets. These are then processed using both MaxEnt and FR models to generate flood susceptibility mapping. The model validation ensures performance exceeds 80% before the final flood susceptibility map is created.

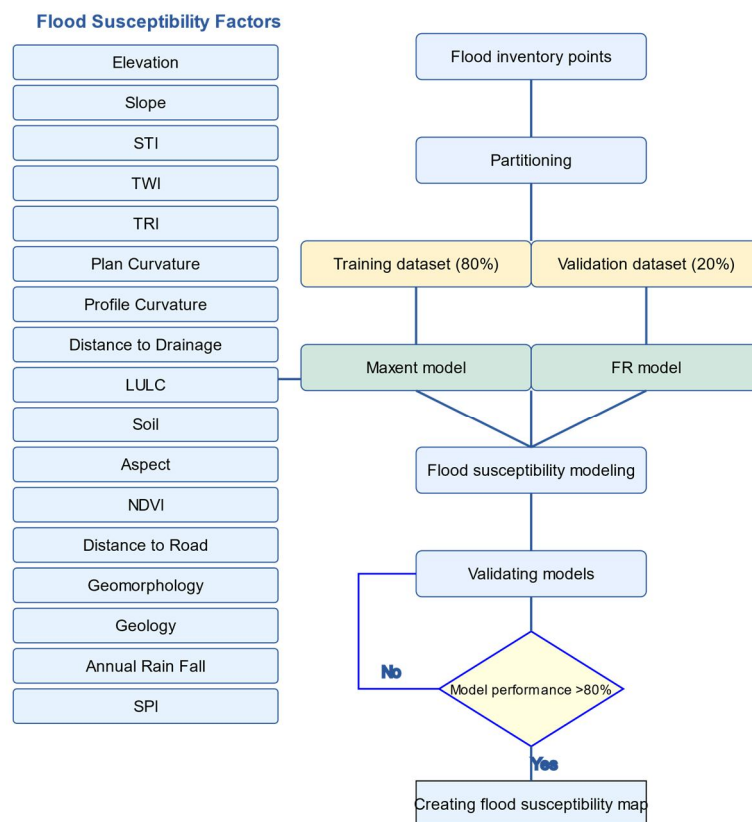


Figure 1. Workflow of flood susceptibility modeling using MaxEnt and FR approaches.

V. DATA PREPROCESSING AND FEATURE SELECTION

A. Data Preprocessing

The data preprocessing involved a complete series of steps to get data to a common high standard of quality and consistency. Data cleaning procedures include discarding outlying values and outliers by using the systematic procedure of removal and error correction mechanisms to maintain data integrity. Standardization of all datasets to 30m resolution and establishing a uniform coordinate system (UTM Zone 43N) were the focus of spatial data processing. Boundary alignment, clipping, and DEM hydro correction are part of this process that would ensure spatial consistency. The task of time alignment on multi-temporal datasets synchronized patterns and historical data integration is achieved.

B. Feature Selection and Feature Importance

A multilevel feature selection method, as shown in Figure 2, is employed to classify the key flood susceptibility factors. Comprehensive statistical significance testing is performed, combined with a variance threshold analysis with a 0.01 threshold to remove constant and quasi-constant features. To identify and deal with correlated features ($r > 0.5$), Correlation analysis, as measured through Pearson correlation coefficients, specifically addresses multicollinearity and feature redundancy elimination. This study evaluated feature importance through the integration of three complementary methods: Random Forest-based importance scoring, information gain analysis, and chi-square feature ranking, all of which were cross-validated to ensure robustness. Random Forest-based scoring was selected for its ability to capture non-linear relationships and handle both numerical and categorical variables while being robust to outliers and missing data.

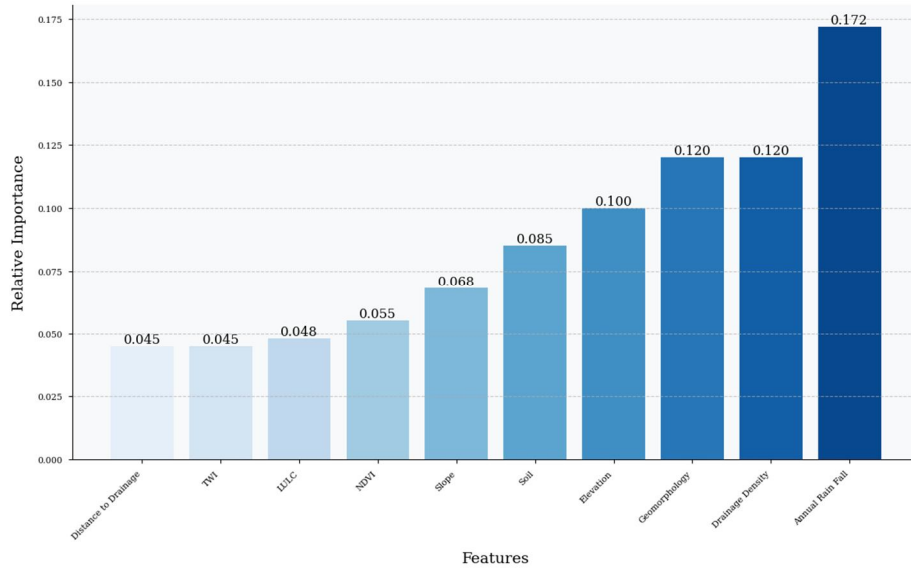


Figure 2 Top 10 Feature Importance Analysis by Random Forest

C. Data Standardization

All datasets received standardized values through z-score normalization for continuous variables, min-max scaling for bounded attributes, and appropriate encoding for categorical features.

VI. FREQUENCY RATIO MODEL

Theoretical Framework and Implementation

The sensitivity of areas to flood is determined based on two factors using the major statistical method called the Frequency Ratio (FR) model. It is of value to researchers on how natural disasters strike (Wang et al., 2021). Samanta et al., 2018 successfully employed the FR model to find weights by finding the relationship between flooding locations and their influencing conditions.

The FR model is expressed as:

$$FR = \left[\frac{N_{pix}(S, X_i)}{\sum_{i=1}^n N_{pix}(S, X_i)} \right] / \left[\frac{N_{pix}(X_j)}{\sum_{j=1}^n N_{pix}(X_j)} \right] \quad (\text{eq.10})$$

The percentages of all possible parameters in X when $N_{pix}(X_j)$ which denotes the total pixel count thrown in the parameter X category and $N_{pix}(X_j)$. To refine the model further, the Local Frequency Ratio (LFR) normalizes. The Prediction Rate (PR) is calculated due to the reciprocal interaction of flooding conditioning parameters.

$$PR = \frac{(RF_{max} - RF_{min})}{\min(RF_{max} - RF_{min})} \quad (eq.11)$$

Where $(RF_{max} - RF_{min})$ represent the highest and lowest RF values respectively, and the $\min(RF_{max} - RF_{min})$ indicates the minimum value among all selected components (Khosravi et al., 2018).

The final Flood Susceptibility Index (FSI) is derived through the following equation:

$$FSI = \sum_{i=1}^n PR_i \times RF_i \quad (eq.12)$$

In this equation, PR_i represents the weight assigned to each conditioning variable, RF_i denotes the number of subclasses of conditioning factors, and n indicates the total number of variables under consideration (Shahabi et al., 2020).

VII. MAXIMUM ENTROPY MODEL

Model Implementation

Several modifications are made to address the problem of flood susceptibility assessment using the Maximum Entropy (MaxEnt) model that Shannon first proposed (1948).

The principle of MaxEnt is based on the concept of information entropy, defined as:

$$H = -C \sum_{i=1}^n p_i \ln p_i \quad (eq.13)$$

where H represents the data entropy, p_i denotes the likelihood of occurrence, and C is an affirmative constant. The model seeks to maximize entropy H under experimental conditions, subject to known constraints.

The model processes the data through the following key steps:

- Feature Processing: The input features X are standardized using the Standard Scaler transformation:

$$X_{scaled} = \frac{X - \mu}{\sigma} \quad (eq.14)$$

- The model creates probability distribution estimates through binned histograms for both flood and non-flood circumstances regarding each feature.:

$$P(flood|X) = \frac{hist_{flood}(X)}{\sum hist_{flood}(X)} \quad (eq.15)$$

$$P(non - flood|X) = \frac{hist_{non-flood}(X)}{\sum hist_{non-flood}(X)} \quad (eq.16)$$

Equation 16 is the conditional probability of flooding given X , labeled as $P(flood|X)$. It simply takes the histogram to count for that feature value under flood conditions $hist_{flood}(X)$ and divides it by the total sum of histogram counts for flood conditions $\sum hist_{flood}(X)$, to give us this probability, which normalizes the distribution. Like this, equation 17 also determines $P(non - flood|X)$ the conditional probability of non-flooding, conditioned by the same feature value.

VIII. RESULTS AND DISCUSSION

A. Model Performance Analysis

It is shown that most of the model's advantages and disadvantages can be determined from the Frequency Ratio (FR) and Maximum Entropy (MaxEnt). Ultimately, the models' advantages and weaknesses are evaluated, discussed, and analysed under several perspectives of outcomes.

The models were evaluated in terms of the discrimination capacity over the flood-prone/non-flood-prone areas using the Area Under the Curve value of the ROC curve, which measures the models' ability to discriminate flood-prone from non-flood-prone areas, varying all the possible rating thresholds. We assessed these methods using our statistical evaluation of the prediction results that we obtained.

The AUC score for the Frequency Ratio model that we predicted floods was 0.946, and the maxent model was 0.928. The model performance results are better than random predictions 0.5. The strong discriminative

capability of the two models was indicated by the ROC curves (Figure 3), with the FR model marginally performing better.

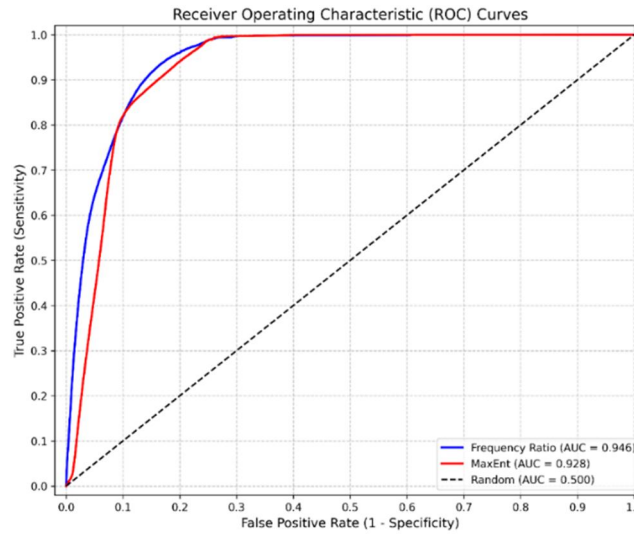


Figure 3 ROC Curves of FR and MaxEnt

B. Calibration and Threshold Selection

Analysis of the frequency ratio (FR) Model is consistent with a cautious forecasting system, which has performed exceptionally well in predicting safe areas with specificity of 0.938. It achieves a sensitivity of 0.689 for flood prone areas while retaining excellent predictive value of safe zones (negative predictive value = 0.999). The results indicate that the FR model demonstrates utility under conditions requiring high precision.

The MaxEnt Model is a comprehensive early warning system, with a high sensitivity of 0.880, and can capture a major proportion of potential flood areas. It has a perfect negative predictive value of 1.000 and a specificity of 0.857, thus convincingly displaying reliability while identifying safe zones. The study highlights the characteristics of the MaxEnt Model as being especially applicable to applications where comprehensive risk identification is required and provides robust performance in comprehensive flood risk assessment.

C. Spatial Distribution of Flood Susceptibility Classes

Results from spatial distribution analysis of flood susceptibility classes showed clear pattern between the Frequency Ratio (FR) (Table 4) and Maximum Entropy (MaxEnt) (Table 5) models. The FR model assessment indicated that 15,627.67 km² (86.06%) fell within low/no susceptibility zones. A total of 526.68 km² (2.90%) was high susceptibility and 2005.53 km² (11.04 %) moderate susceptibility zones. The Maxent model showed distinctive pattern of distribution with 13,608.95 km² (74.94%) of study area defined as low/no susceptibility zones. Maxent model suggested high susceptibility zones in 3,704.97 km² (20.40%) of the area and moderate susceptibility zones in 845.97 km² (4.66%).

Characteristic differences between the two models are also shown in the distribution patterns in the classification of high-risk areas. Results from the Maxent model revealed the larger areas to be highly susceptible to flooding than that of FR model suggesting practical implications for the assessment and management of flood risks.

TABLE II: STATISTICS OF FR MODEL

#	Flood susceptibility	Area(km ²)	Percentage (%)
1	High	526.6835	2.9002
2	Moderate	2005.5283	11.0437
3	Low/No	15627.6745	86.0560

TABLE III: STATISTICS OF MAXENT MODEL

#	Flood susceptibility	Area(km ²)	Percentage (%)
1	High	3704.9677	20.4019
2	Moderate	845.9714	4.658
3	Low/No	13608.9472	74.9396

IX. ENVIRONMENTAL FACTOR CONTRIBUTION

The relative contribution analysis of environmental factors revealed varying importance patterns between the Frequency Ratio (FR) and Maximum Entropy (MaxEnt) models (Table 6). In the MaxEnt model, geological characteristics emerged as the primary contributor with 18.40% importance, followed by land use/land cover (LULC 2018) at 15.20%. Topographic wetness index (TWI), NDVI, and drainage density showed comparable contributions of 12.80%, 12.90%, and 12.60%, respectively. Aspect contributed 9.00%, while distance to drainage networks accounted for 5.70% of the model's predictive power.

The FR model demonstrated different variable importance patterns (Table 6), with annual rainfall showing the highest contribution at 17.10%, followed by drainage density (12.70%) and geology (11.90%). Elevation parameters contributed 11.40%, while soil characteristics and slope accounted for 8.30% and 7.60%, respectively.

TABLE IV: ENVIRONMENTAL FACTOR CONTRIBUTIONS (%)

Environmental Factor	MaxEnt Model	FR Model
Geology	18.40%	11.90%
Annual Rainfall	7.10%	17.10%
LULC 2018	15.20%	4.70%
NDVI	12.90%	5.00%
TWI	12.80%	4.00%
Drainage Density	12.60%	12.70%
Elevation	0.10%	11.40%
Aspect	9.00%	2.90%
Soil	2.20%	8.30%
Slope	1.40%	7.60%
Distance to Drainage	5.70%	3.90%
TRI	0.20%	2.70%
Profile Curvature	0.00%	2.60%
Plan Curvature	0.00%	2.30%
SPI	2.20%	1.30%
STI	0.00%	0.80%
Distance to Road	0.20%	0.60%

X. FLOOD SUSCEPTIBILITY MAPS

The formulation and examination of flood susceptibility maps are essential to find and reduce flood risks in many places. These flood maps mix data about the environment with terrain information and water features to show which areas might flood.

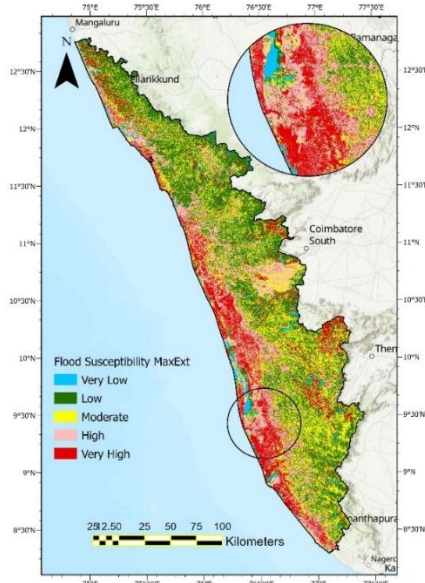


Figure 4. Flood Susceptibility Map MaxEnt

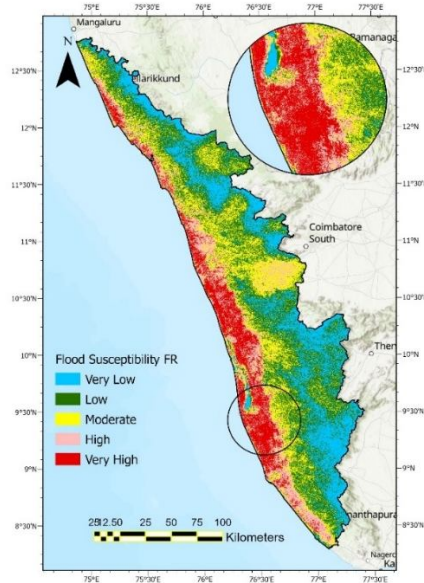


Figure 5. Flood Susceptibility Map FR

Figure 4 and Figure 5 illustrate flood susceptibility mapping for Kerala using MaxEnt and Frequency Ratio (FR) methods. The maps display susceptibility levels from very low (blue) to very high (red), with circular insets highlighting critical zones. Both visualizations integrate environmental, topographical, and hydrological factors to identify flood-prone regions.

XI. CONCLUSIONS AND LIMITATIONS

The flood susceptibility analysis in Kerala revealed how well different modeling approaches for flood risk assessment. The comparative analysis of Frequency Ratio (FR) and Maximum Entropy (MaxEnt) showed us how environmental variables team up to shape flood risk. Annual rainfall with drainage density and geology showed highest impact on flood risk according to both modeling techniques.

The outcome of the present study showed that both models have a very high predictive ability with unique features of the models driving their performances. The Frequency Ratio (FR) model achieved an AUC of 0.946, while the Maximum Entropy (MaxEnt) model performed better with an AUC of 0.928.

The flood assessment results showed that Frequency Ratio and Maximum Entropy models work uniquely when predicting flood risk. The FR model proved reliable at finding non-flooded locations with 94.6% efficiency. The MaxEnt model offers a wide range of flood risk detection capabilities that make it useful for comprehensive flood analysis to locate areas at higher risk.

The study's limitations include temporal constraints in rainfall data (2011-2020) and land use data (2018), the static nature of the models, and scale-dependent analysis that may overlook local flooding events. While these limitations affect result interpretation, the study provides valuable insights for flood risk assessment in Kerala and establishes a foundation for future research in flood susceptibility mapping.

BIBLIOGRAPHY

- [1] Aju, C. D., Achu, A. L., Prakash, P., Raicy, M. C., & Reghunath, R. (2024). An integrated statistical-geospatial approach for the delineation of flood-vulnerable sub-basins and identification of suitable areas for flood shelters in a tropical river basin, Kerala. *Geosystems and Geoenvironment*, 3(2), 100251. <https://doi.org/10.1016/j.geogeo.2024.100251>
- [2] Costabile, P., Costanzo, C., Gandolfi, C., Gangi, F., & Masseroni, D. (2022). Effects of DEM Depression Filling on River Drainage Patterns and Surface Runoff Generated by 2D Rain-on-Grid Scenarios. *Water*, 14(7), 997. <https://doi.org/10.3390/w14070997>
- [3] Dhasmana, M. K., Mondal, A., & Zachariah, M. (2023). On the role of climate change in the 2018 flooding event in Kerala. *Environmental Research Letters*, 18(8), 084016. <https://doi.org/10.1088/1748-9326/ace6c0>
- [4] Fischer, S., Lun, D., Schumann, A. H., & Blöschl, G. (2023). Detecting flood-type-specific flood-rich and flood-poor periods in peaks-over-threshold series with application to Bavaria (Germany). *Stochastic Environmental Research and Risk Assessment*, 37(4), 1395–1413. <https://doi.org/10.1007/s00477-022-02350-8>
- [5] Florinsky, I. V. (1998). Combined analysis of digital terrain models and remotely sensed data in landscape investigations. *Progress in Physical Geography: Earth and Environment*, 22(1), 33–60. <https://doi.org/10.1177/030913339802200102>
- [6] Hofmann, J., & Schütttrumpf, H. (2021). floodGAN: Using Deep Adversarial Learning to Predict Pluvial Flooding in Real Time. *Water*, 13(16), 2255. <https://doi.org/10.3390/w13162255>
- [7] Horn, B. K. P. (1981). Hill shading and the reflectance map. *Proceedings of the IEEE*, 69(1), 14–47. <https://doi.org/10.1109/PROC.1981.11918>
- [8] Huangfu, W., Qiu, H., Wu, W., Qin, Y., Zhou, X., Zhang, Y., Ullah, M., & He, Y. (2024). Enhancing the Performance of Landslide Susceptibility Mapping with Frequency Ratio and Gaussian Mixture Model. *Land*, 13(7), 1039. <https://doi.org/10.3390/land13071039>
- [9] Ifejube, O. J., Kuriakose, S. L., Anish, T. S., van Westen, C., & Blanford, J. I. (2024). Analysing the outbreaks of leptospirosis after floods in Kerala, India. *International Journal of Health Geographics*, 23(1), 11. <https://doi.org/10.1186/s12942-024-00372-9>
- [10] Kawami, J., & Miura, T. (2022). Improving Maximum Entropy Model by GIS. <https://doi.org/10.3233/FAIA210489>
- [11] Khosravi, K., Pham, B. T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., Prakash, I., & Tien Bui, D. (2018). A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Science of The Total Environment*, 627, 744–755. <https://doi.org/10.1016/j.scitotenv.2018.01.266>
- [12] Kumar, P., Kumari, N., & Sahu, N. C. (2022). Floods and economic growth in India: role of FDI inflows and foreign aid. *Management of Environmental Quality: An International Journal*, 33(5), 1114–1131. <https://doi.org/10.1108/MEQ-10-2021-0244>
- [13] Lal, P., Prakash, A., Kumar, A., Srivastava, P. K., Saikia, P., Pandey, A. C., Srivastava, P., & Khan, M. L. (2020). Evaluating the 2018 extreme flood hazard events in Kerala, India. *Remote Sensing Letters*, 11(5), 436–445. <https://doi.org/10.1080/2150704X.2020.1730468>

- [14] Mathew, G., Varghese, A., Kumar, S. S., & Chacko, A. (2023). "After the rains" – A qualitative study exploring flood-related experiences, community response, coping mechanisms, and strategies from Kerala. *Journal of Datta Meghe Institute of Medical Sciences University*, 18(1), 18. https://doi.org/10.4103/jdmimsu.jdmimsu_314_22
- [15] Menon, N., Parastoo, S., & Adriaensen, R. (2024). Flood Inundation Mapping of the 2018 Kerala Floods: A Comparative Study of Traditional Remote Sensing, Machine Learning, and Deep Learning Methods. *Egusphere*, 24, 22400. <https://doi.org/10.5194/egusphere-egu24-22400>
- [16] Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5(1), 3–30. <https://doi.org/10.1002/hyp.3360050103>
- [17] Morshed, S., Rahman, Md. T., Rokonzaman, S., & Hossain, A. (2022). The Economic Impact of Monsoon Flood and Its Spillover on the Households of Bangladesh. *Journal of Sustainable Development*, 15(3), 23. <https://doi.org/10.5539/jsd.v15n3p23>
- [18] Rouse Jr., J. ~W., Haas, R. ~H., Schell, J. ~A., & Deering, D. ~W. (1974). Monitoring Vegetation Systems in the Great Plains with ERTS. In S. C. Freden, E. P. Mercanti, & M. A. Becker (Eds.), *NASA Special Publication* (Vol. 351, p. 309).
- [19] Samanta, S., Pal, D. K., & Palsamanta, B. (2018). Flood susceptibility analysis through remote sensing, GIS and frequency ratio model. *Applied Water Science*, 8(2), 66. <https://doi.org/10.1007/s13201-018-0710-1>
- [20] Shahabi, H., Shirzadi, A., Ghaderi, K., Omidvar, E., Al-Ansari, N., Clague, J. J., Geertsema, M., Khosravi, K., Amini, A., Bahrani, S., Rahmati, O., Habibi, K., Mohammadi, A., Nguyen, H., Melesse, A. M., Ahmad, B. Bin, & Ahmad, A. (2020). Flood Detection and Susceptibility Mapping Using Sentinel-1 Remote Sensing Data and a Machine Learning Approach: Hybrid Intelligence of Bagging Ensemble Based on K-Nearest Neighbor Classifier. *Remote Sensing*, 12(2), 266. <https://doi.org/10.3390/rs12020266>
- [21] Sharma, A., Poonia, M., Rai, A., Biniwale, R. B., Tügel, F., Holzbecher, E., & Hinkelmann, R. (2024). Flood Susceptibility Mapping Using GIS-Based Frequency Ratio and Shannon's Entropy Index Bivariate Statistical Models: A Case Study of Chandrapur District, India. *ISPRS International Journal of Geo-Information*, 13(8), 297. <https://doi.org/10.3390/ijgi13080297>
- [22] Siahkamari, S., Haghizadeh, A., Zeinivand, H., Tahmasebipour, N., & Rahmati, O. (2018). Spatial prediction of flood-susceptible areas using frequency ratio and maximum entropy models. *Geocarto International*, 33(9), 927–941. <https://doi.org/10.1080/10106049.2017.1316780>
- [23] Singh, sheilja. (2019). Analytical Study of Kerala Floods. *International Journal for Research in Applied Science and Engineering Technology*, 7(6), 1851–1855. <https://doi.org/10.22214/ijraset.2019.6311>
- [24] Sørensen, R., Zinko, U., & Seibert, J. (2006). On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences*, 10(1), 101–112. <https://doi.org/10.5194/hess-10-101-2006>
- [25] Wang, Y., Fang, Z., Hong, H., Costache, R., & Tang, X. (2021). Flood susceptibility mapping by integrating frequency ratio and index of entropy with multilayer perceptron and classification and regression tree. *Journal of Environmental Management*, 289, 112449. <https://doi.org/10.1016/j.jenvman.2021.112449>
- [26] Zeng, Z., Guan, D., Steenge, A. E., Xia, Y., & Mendoza-Tinoco, D. (2019). Flood footprint assessment: a new approach for flood-induced indirect economic impact measurement and post-flood recovery. *Journal of Hydrology*, 579, 124204. <https://doi.org/10.1016/j.jhydrol.2019.124204>