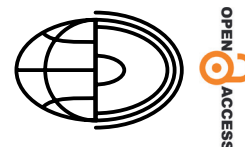


Flood susceptibility mapping using Fuzzy, Statistical enhancement and Machine Learning Models: a study on the Mayurakshi River Basin of eastern India



Sohini Neogy¹, Pradip K. Polley², Sreyashi Shib³, Sukanta Das⁴,
Konia Basu⁴, Ujwal Deep Saha⁵[✉]

¹Banwarilal Bhalotia College, Department of Geography, Asansol, West Bengal, India

²Vidyasagar College, Department of Geography, Kolkata, West Bengal, India

³Indian Institute of Science Education and Research (IISER) Bhopal, Earth and Environmental Sciences, Madhya Pradesh, India

⁴School of Civil and Environmental Engineering, Indian Institute of Technology Mandi, Himachal Pradesh, India

⁵JIS University, Kolkata, India

⁶Polish Academy of Sciences, Institute of Geography and Spatial Organisation, Krakow, Poland

*Correspondence e-mail: sahaujwal.geo@gmail.com

 [✉]<https://orcid.org/0000-0002-0225-7710>

Abstract. The Mayurakshi River Basin (MRB) is imperilled by flooding every year and is a major flood-prone area of eastern India. The present study highlights the crucial factors triggering flood inundation and identifies the flood-susceptible zones (FSZ) in the MRB using pertinent models and a statistical-model-based supervised enhancement. The FSZs of the MRB were determined using Fuzzy Analytical Hierarchy Process (FAHP), enhanced FAHP using Partial Least Square Regression-Variable Importance in Projection (PLS-VIP-FAHP) method and machine learning (ML) algorithms, Support Vector Machine (SVM) and Artificial Neural Network (ANN). The FAHP-, PLS-VIP-FAHP-, SVM- and ANN-based FSZ outcomes were validated using the Receiver Operating Characteristics (ROC) curve with a 70:30 training–testing ratio. The SVM model registered the highest degree of accuracy, with 0.96, while PLS-VIP-FAHP registered the lowest, with 0.89. Interestingly, the accuracy metric for FAHP surpassed the same for ANN, with FAHP registering 0.93 and the latter registering 0.92. While testing with a greater number of sample points, ANN registered an AUC of 0.90, and FAHP registered 0.89. Application of ground knowledge in selecting response parameters could help mathematical models to gain significant accuracy compared to the ML algorithm.

Key words:

ANN,
flood susceptibility,
Fuzzy-Analytical Hierarchy
Process,
ROC curve,
SVM,
Partial Least Squares

Introduction

Floods occurring in numerous low-lying deltaic areas can be deemed as advantageous, specifically for augmenting soil fertility on flood plains, but also as a hazard for imperilling human life, destroying property and sculpting the environment, whether induced by natural events or anthropogenic stresses or both (Ghosh and Kar 2018; Kaushik et al. 2025). Globally

regarded as one of the most pernicious hydrological and meteorological hazards, the occurrences of floods have amplified pre-eminently in the last few decades (Ghosh and Maiti 2021; Misra et al. 2025). Considering its negative consequences, flood susceptibility mapping has become inevitable to control future flooding (Tehrany et al. 2015). Nevertheless, flooding has multifaceted circumstances that reduce the certainty of predictions, and this demands a proper choice of factors that influence floods. Past studies

have analysed the flood problems as well as the frequency of occurrences concerning tropical India (Sen 2010), and flood occurrence has undoubtedly been stated as a by-product of human–nature interactions (Mukhopadhyay and Let 2014). The catastrophic outcome – life loss and infrastructural damage – of the Rarh Bengal rivers has been vividly depicted in the study conducted by Chatterjee and Jana (2022). On the Rarh plain, the riverine flood is a frequent natural hazard that triggers persistent fluctuations in landscape dynamics and societal alterations (Islam and Ghosh 2021). In this context, floods of the Mayurakshi River Basin (MRB) are especially noted to severely affect both the life and livelihood of the people (Islam and Sarkar 2020). It is since the 1860s that the floods of the MRB have been documented thoroughly. The reports obtained from O'Malley (1914), the *District Disaster Management Plan* (2016), and the *Flood Preparedness and Management Plan* (2014) put forth the fact that this region is a witness to severe flooding conditions. Additionally, it is evident from the works of Ghosh and Pal (2015) and Pal (2016) that the lower catchment of MRB has transformed from high-magnitude, low-frequency flood events to low-magnitude, high-frequency flood events owing to the construction of major dams and barrages (Islam and Sarkar 2020).

Since the behaviour of river flooding has changed drastically, it becomes crucial to comprehend and model the probable spatial extent of floodplain inundation (Jafarzadegan et al. 2023). However, modelling floods is a multi-dimensional, dynamic task wherein a multitude of factors are supposed to be taken into consideration, but the identification of parameters plays a crucial role as flood inundation is sensitive to these parameters locally (Islam and Barman 2020; Swain et al. 2020). Location-specific factors are primarily reckoned to entail accuracy in flood predictions (Islam et al. 2021a). The selection of these factors exceedingly depends on the spatial scale and the geographical attribute of the envisioned flood vulnerability investigation (Vojtek and Vojtekova 2019). Additionally, the factors influencing flood vulnerability also depend on the lack of forecasting extreme precipitation and flooding events, especially during the non-monsoon months in India (Mishra et al. 2024). Choosing an inadequate number of factors often increases the probability of receiving faulty results (Mahmoud and Gan 2018).

The application of remote sensing and statistical models has gained ground in solving such problems (Das 2018; Mudashiru et al. 2021). ML algorithms are highly effective and hold great potential in accurate and

efficient prediction modelling. The appropriate results obtained from the hybrid Machine Learning algorithms aid in predicting several natural hazard models (Islam et al. 2021b). In this context, both the Multi-Criterion Decision Making (MCDM) approach (Khosravi et al. 2019; Souissi et al. 2020), which implies statistical exercises, and ML algorithms solve such complexities in environmental modelling (Salvati et al. 2023; Yang et al. 2023). In MCDM, the relative and respective importance of the chosen factors can be prioritised by assigning weights, which helps in portraying how one alternative outranks the other subject on the specified criteria (Cao et al. 2016). It is a broad term in which a set of methods can be implemented to support the decision-making process by taking into account numerous and often conflicting parameters through a structured framework (Brito and Evers 2016). This method embodies the spatially continuous and collective nature of the inducing criteria on the flood-generating mechanisms (Tang et al. 2018). Similarly, ML provides a wide range of algorithms where Support Vector Machine (SVM) (Tehrany et al. 2015; Bammou et al. 2024) and Artificial Neural Network (ANN) (Priscillia et al. 2021; Lin et al. 2023) have been proven significant in flood susceptibility modelling. Many past ventures have been conducted using an ensemble modelling approach where these two algorithms were kept as the base (Islam et al. 2021b; Liu et al. 2022).

This study deals with comparing different models being tested for flood susceptibility zonation (FSZ) and the efficiency of factor-respective simulations (based on statistical models). The selection of parameters depending upon the previously gained ground knowledge, along with referencing literature on FSZ in rivers of similar traits (Das and Gupta 2021; Kundu et al. 2022), was the initial thrust of this study. Thereafter, a statistical model and its enhancement were used to map the spatial extent of flood susceptibility in the MRB. It was done to identify how a shallow pool of selected parameters (based on their relative contribution) enhances the outcome of the base model. The SVM and ANN, two of the most utilised ML algorithms for hazard mapping, were tested to compare their efficiency with MCDM and externally enhanced MCDM models. The results were tested twice using different amounts of testing points to see whether the test results change with considerations of greater spatial variability or not. So, the focus lies on preparing FSZ for this flood-prone river basin using a selected pool of criteria based on ground knowledge and how different methods of statistical and machine learning can improve

prediction accuracy. The accuracy of the models is also tested twice for different numbers of testing points to get a better resolution regarding the accuracy of the models. The novelty of the research lies in the study area, which is a geomorphologically and topographically complex region of the flood-prone Rarh Bengal, wherein such an integrated flood susceptibility analysis has never previously been conducted. Unlike previous studies, the present research combines a statistically enhanced MCDM framework (to better reflect the relative contribution of flood-inducing factors) with machine learning algorithms – specifically SVM and ANN – thus improving the precision of flood susceptibility mapping in the MRB. The methodology takes into account gaps in earlier conducted flood susceptibility zonation studies where prominent yet conventional ML algorithms for flood susceptibility mapping have been investigated (Islam et al. 2025; Sharma et al. 2025). The present study does not replicate standard models but puts forward a refined model wherein factor selection and weight calibration have been statistically validated. Incorporating statistical enhancement has significantly strengthened the scientific validity and predictive performance of the flood hazard model. Ultimately, this approach contributes to the creation of more stable and spatially representative flood hazard

mapping, which is critical for effective risk planning and mitigation. A general drawback of this study remains in averaging the flood inundation spatially, as average rainfall and a static land use property have been used. Thus, the modelled outcome is static to a particular set of environmental setups; rather, it fails to accommodate the dynamicity of these two most determining factors (Islam et al. 2021a).

Precisely, the major aims of this study are as follows. Firstly, to map the FSZ of the MRB using different models with varied nature of simulation to get a wide range of output with different levels of accuracy. Secondly, to simulate a base model with a shallow pool of criteria to test the relative importance of factors depending on the environmental character of the study area. Lastly, to judge all the models with different sets of testing points to see whether the relative nature of testing points affect the model accuracies.

Study area

MRB encompasses the region between the Chotanagpur Plateau and the western part of the Bengal basin of Eastern India and covers an area of 9,595 km² (Fig. 1). The basin is flanked by four rivers, namely, the Brahmani, Dwaraka, Mayurakshi and Kuye. This

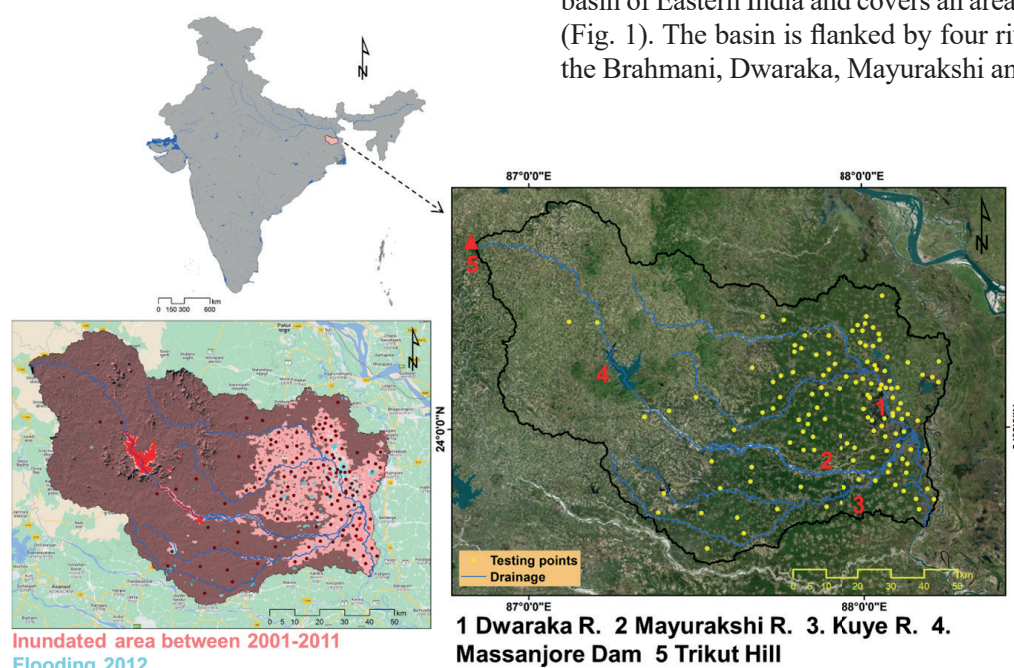


Fig. 1. Location of the study area. The MRB comes under the lower catchment of the Ganga River. (Bottom-left) Flood inundation map of the MRB (source: <https://floodobservatory.colorado.edu/>). Only the set of 100 training-testing points is mentioned in this figure, and the set of 200 training-testing points is mentioned in Fig. 2 for a clearer visualisation. Google maps (2021) has been used as base map in the flood inundation map and Google satellite image (2021) in the map showing training-testing points

river system is constituted of five major civil structures, i.e., the Massanjore dam and Tilpara barrage on the Mayurakshi River, the Deucha barrage on the Dwaraka River, the Brahmani bar—rage on the Brahmani River, and the Bakreshwar weir on the Kuye River. The basin, in its entirety, is a fragment of the Rarh tract and is largely composed of Dharwar sediments followed by Hercynian orogeny in the upper portion, lateritic soil and hard clay deposition in the middle segment, and alternate layers of recent sand, silt and clay in the lower extensions (Mukherjee et al. 2009). Geomorphologically, rolling uplands and lateritic badlands can be viewed in the upper portion of the basin, and low-lying flat depressed lands can be seen in the middle and lower parts. It is this depressed region that is highly vulnerable to flooding and thus records recurrent flood incidents. This area is dominated by a sub-tropical monsoon climate with the average annual rainfall being ~1,500 mm. Around 80% of the rainfall is received in June to September under the influence of the south-west monsoons (Pal 2016). This high degree of seasonality of rainfall causes water stagnation in the lower catchment, leading to

frequent flood episodes. Historical records and several government reports from the past three decades specify the severity of the intensity and magnitude of the floods faced in MRB. Since 2000, a total of five flood incidents have been recorded – in 2000, 2006, 2007, 2009 and 2015 (Islam and Barman 2020). Flood depth and duration happen to be highly correlated in the lower region of MRB (Islam and Sarkar 2020).

Methods

Data procurement

This study heavily relies on a digital database. The DEM (SRTM 30m) data were collected freely from <https://earthexplorer.usgs.gov/>, while the LULC dataset and rainfall data were collected from various scientific web-based platforms. The procurement of the entire dataset has been illustrated in Table 1 and the spatial distribution can be visualised in Figure 2..

Table 1. Description of the flood conditioning factors

Parameters	Calculation	Source	Remarks
Elevation	-	SRTM 30m DEM	5 class classification using natural breaks
Slope	Topography module in ArcGIS	SRTM 30m DEM	5 class classification using natural breaks
Topographic Wetness Index (TWI)	$\ln\left(\frac{A}{\tan\beta}\right)$; A = area of the catchment; β = slope angle	SRTM 30m DEM	5 class classification using natural breaks; higher value = greater flood risk
Topographic Ruggedness Index (TRI)	Amount of elevation difference from adjacent cells	SRTM 30m DEM	5 class classification using natural breaks
Curvature	Topography module in ArcGIS	SRTM 30m DEM	5 class classification using natural breaks
Rainfall	Average of monsoon rainfall (June to September)	https://www.worldclim.org/	5 class classification using natural breaks
Flow Accumulation	Hydrology module in ArcGIS	SRTM 30m DEM	5 class classification using natural breaks
Soil		https://nbsslup.icar.gov.in/	Thematic
Groundwater level	Average monsoon water level (June to September)	http://cgwb.gov.in/	5 class classification using natural breaks
Distance from the river	Vector overlay operation	Vectors of river channels	Uniform
Drainage Density (DD)	Interpolation using the IDW method	Vectors of river channels	5 class classification using natural breaks
Land use land cover (LULC)	-	https://livingatlas.arcgis.com/landcover/	Thematic
Geology	-	https://bhukosh.gsi.gov.in/Bhukosh/Public	Thematic

Flood conditioning factors

Elevation

Elevation of an area is deemed one of the triggering factors that control the flood vulnerability of a region. With greater flow accumulation, a lower elevation is more susceptible to flood occurrence. However, it can also determine the topographic character and vegetation distribution (Khosravi et al. 2016).

Slope

The degree of slope is another unswerving inducing factor because it regulates surface water flow. Slope regulates the surface run-off and vertical percolation of water, and there exists a negative relationship between slope angle and infiltration rate (Youssef et al. 2011). Thus, a higher slope induces more run-off while water stagnancy pertains on lower slopes, thereby leading to flood-like conditions on lower slopes.

Topographic Wetness Index (TWI)

TWI aids in the quantification of the degree of water accumulation in a certain landscape. It is calculated as $TWI = \ln(\alpha / \tan \beta)$, where α is the flow accumulation and β refers to the slope in radians (Moore et al. 1991). Higher TWI values correspond to higher chances of flooding, and vice versa, because it is the degree of downstream trends of discharge (Gokceoglu et al. 2005).

Topographic Ruggedness Index (TRI)

TRI can be defined as the homogeneity in the altitudinal distribution of the land surface. Areas with low TRI values generally exhibit a high possibility of flooding owing to their flat nature (Khodaei et al. 2025). Consequently, these values can be located in and around floodplains. High values, on the other hand, indicate rugged topographic conditions. The TRI raster was classified into five classes using the natural breaks method. Since the study area is composed of plateaus and plains with transitions in between, TRI is an essential component for flood mapping in such terrain.

Curvature

Defined as the rate of change of slope in a specified direction, curvature is a central factor in the flooding of a region. It is often considered an essential controlling factor of flood, as it controls runoffs; its acceleration and deceleration are based on the nature of the slope (Shahabi et al. 2021). The concave shape of the landscape is represented by negative curvature values, whereas the convexity of the topography is showcased by near-zero curvature and positive values. Areas with curvature in negative values are more susceptible to flooding (Nandi et al. 2016), thus representing flat surfaces with high occurrences of flooding.

Rainfall

Rainfall has a profound bearing on flood occurrences, as it has a direct relationship with river water discharge (Wu et al. 2016). Flash flood-like conditions can also exist in case of sudden downpours in a short span of time. Here, the rainfall factor has been quantified based on 30 years of precipitation data obtained from 1970 to 2000.

Flow Accumulation

Flow accumulation is an obvious factor in influencing the chances of flooding. High flooding can be detected in areas of high accumulation, and vice versa. This can be attributed to the augmented runoff in low topographic regions leading to increased flood susceptibility.

Soil

Infiltration capacity (and, consequently, runoff) is directly linked with soil texture. Regions with fine soil texture have a greater possibility of flooding than regions with coarse soil texture. In this study, the plain is mostly composed of clayey soil or clay-loam texture, which has a low capacity for infiltration. Owing to low slope and low ruggedness, these regions become more susceptible to waterlogging, since vertical penetration is low.

Groundwater level

The groundwater level is an imperative factor in the determination of flood occurrences. High groundwater levels lead to higher chances of flooding, whereas

lower levels of groundwater, on the other hand, lead to lower chances of flooding events. Since these areas often experience late monsoon rainfall, the groundwater level becomes a pertinent factor here (Bandyopadhyay et al. 2014). A higher groundwater level restricts vertical penetration of surface water into the soil.

Distance from the river

Distance from the river happens to be a major factor in determining flood occurrences in a region. Generally, a flood occurs near the riverbanks, further inundating the adjacent low-elevated regions. As the distance keeps increasing, the slope and elevation become higher, reducing flood risk. Thus, areas close to the river are more susceptible to flooding if we put the topographic consideration aside (Shahabi et al. 2021). It was calculated using buffer distance (Euclidean) in the ArcGIS software.

Drainage Density (DD)

Drainage density refers to the stream length per unit area in the basin area. It is directly proportional to the likelihood of flood events. High DD entails high runoff, high flow accumulation, and low infiltration rate, thus inducing high chances of flooding, and vice versa.

Land use land cover (LULC)

Both the natural and anthropogenic features on the surface of the Earth affect the magnitude of floods, as they modify the topographic condition and degree of infiltration (Pahari et al. 2024). Vegetated areas enhance infiltration capacity, thus lowering the risk of floods. Subsequently, the possibility of floods is much higher over impermeable surfaces, due to reduced infiltration rates. Additionally, modifications to the utilisation of the land trigger chances of flooding. Here, the LULC map was downloaded from the ESRI server, prepared based on the 2020 satellite dataset at 10-m resolution. This region is an agriculturally dominated region; thus, the economic loss due to flooding is an issue of significant concern for the government.

Geology

Geological structures with impermeable formations (i.e., clay and silt) augment the rate of runoff, thereby amplifying the occurrence of floods. Alternatively, permeable structures, like sand, allow infiltration,

leading to decreased chances of flooding. The hard rock concentration on the upper part of the catchment contributes to greater surface runoff unless restricted in fracture zones. This degree of runoff contributes to more water accumulation downstream, where, on the plains, the concentration of hard clay and a higher groundwater level restrict vertical percolation.

Collinearity testing

The variance inflation factor (VIF) is a common method for detecting the presence of linear relationships involving two or more independent variables (i.e., collinearity) in the multiple linear regression model (Gomez et al. 2016). VIF assesses how much the correlation between a predictor variable and the other predictor variables in the model increases the variance of the estimated coefficient for that variable (Kariminejad et al. 2019). Ranging from one to positive infinity, a VIF of one specifies that the variable is not correlated with other predictors, while higher VIF values suggest increasing levels of multicollinearity (Yilmaz 2007). VIF is generally calculated as:

$$VIF = \frac{1}{1 - R_j^2}$$

where R_j^2 is the unadjusted coefficient of determination for the regression model.

In this study, VIF and tolerance (TOL) have been used to quantify the importance of flood conditioning factors. $VIF > 9$ and $TOL < 0.1$ are considered threshold limits, and they indicate the problem of multicollinearity in the chosen factors (Islam et al. 2021b). Therefore, it is suggested not to include those factors for modelling that has a $VIF > 9$.

Fuzzy Analytical Hierarchy Process (FAHP)

The size of pair comparisons between criteria is constrained to single values in the conventional AHP method following the designated scaling system. As a result, AHP is criticised for having an imbalanced scale of judgment and failing to adequately account for the inherent ambiguity and vagueness in conducting comparisons between pairs (Pourghasemi et al. 2012; Yang et al. 2013). Fuzzy theory helps in reducing the fuzziness and uncertainty in data distribution (Lyu et al. 2020). Integration of Fuzzy theory and AHP models is a practical method widely used in environmental risk assessment (Peng et al. 2021). The Fuzzy AHP (FAHP) method is used to determine

the weights of the selected criteria in this study. We have used the traditional trapezoidal fuzzy number to determine the fuzzy weights. This type of fuzzy is widely used for hazard susceptibility mapping as it can handle the vagueness and uncertainty in the used dataset. It helps in achieving better accuracy in determining weights (Wang et al. 2021). Here, a total of four external opinions were considered to prepare a pairwise comparison matrix.

Among those, the judgment with the lowest error was selected. These opinions were taken from researchers with prior knowledge of applying AHP in hazard-related studies. The overall calculation process of the Fuzzy AHP model is summed up briefly in four steps which are summarised as: (a) representation of judgments for pairwise comparison for creating the decision matrix; (b) aggregation of Fuzzy sets for group decisions and criteria

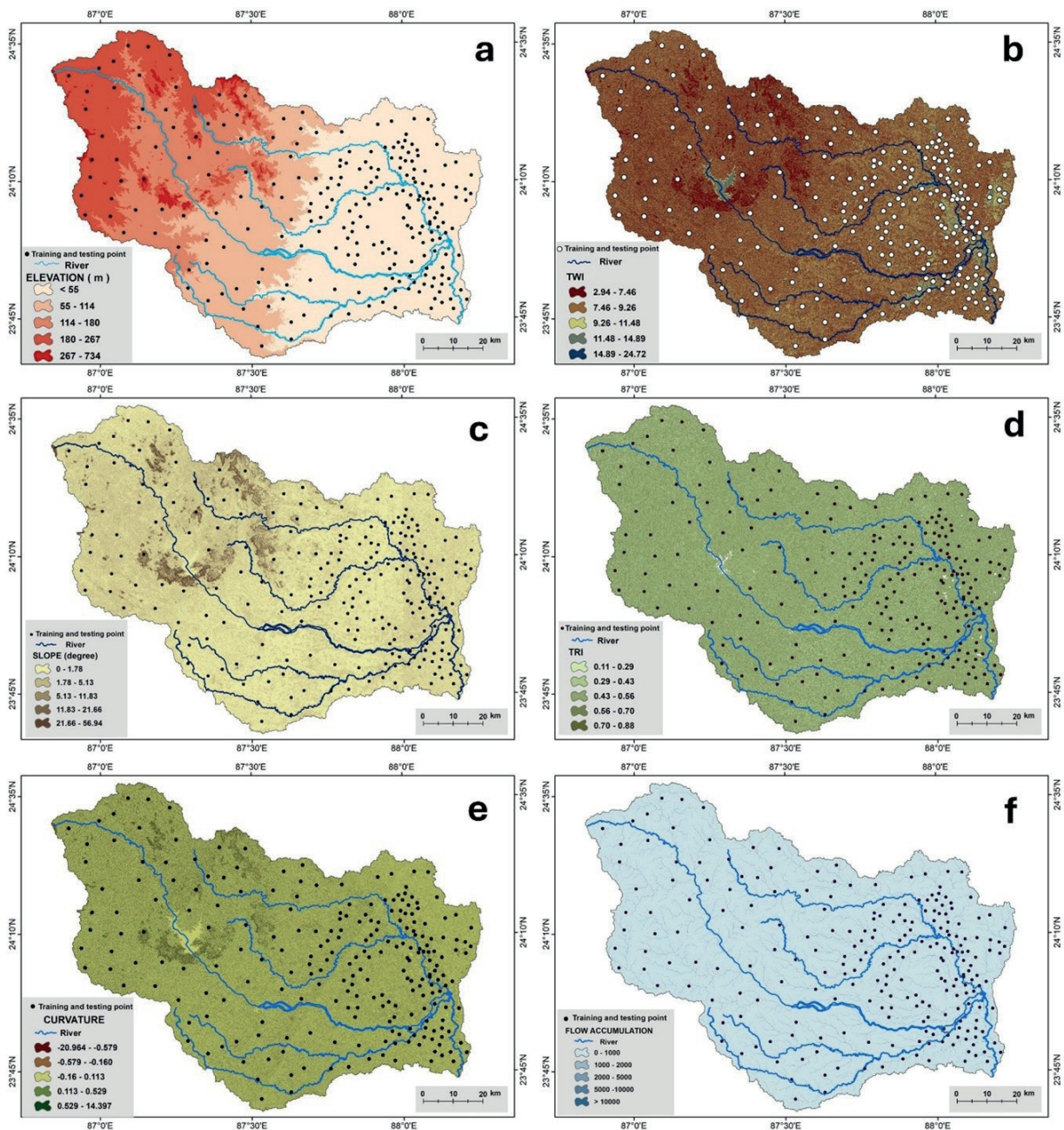


Fig. 2. Parameters selected for modelling flood susceptibility zones: a) elevation, b) TWI, c) slope, d) TRI, e) curvature, f) flow accumulation, g) rainfall, h) drainage density, i) ground water level, j) LULC, k) geology, l) soil texture, m) distance from river (continuation on the next page)

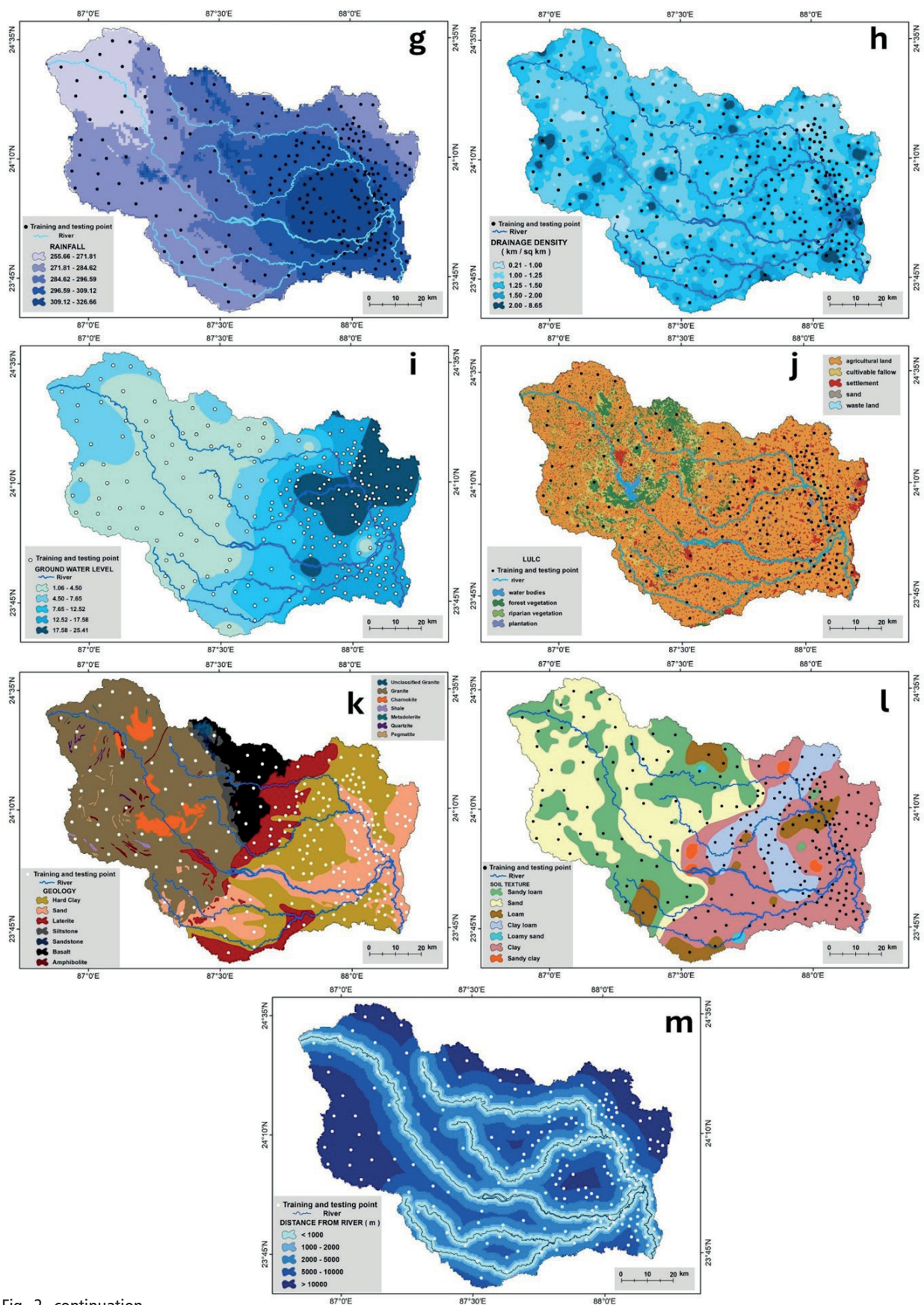


Fig. 2. continuation

weights; (c) de-fuzzification of a fuzzy set for further comparison and, lastly, (d) consistency measurement (consistency level <0.1) for limited contradiction which will be addressed in turn (Baral et al. 2023). The geometric mean method was used to determine the fuzzy weights (Table 2). A detailed description of the FAHP process can be found in the works of Aliyev et al. (2020) and Mahad et al. (2019). Here, apart from the computation of the weightage of the selected parameters, the class rank of each of the parameters was also derived using the AHP method (Table 3). The weighted linear combination model was used to calculate the composite scores, the product of the weightage of the parameters and the respective class rank for each pixel.

Partial Least Squares Regression-Variable Importance in Projection (PLS-VIP) method

The data were statistically analysed using the PLS-VIP method to determine the relevance of variables toward flood susceptibility. The primary purpose of

this technique is to reduce the number of predictors to a smaller set of variables to identify the best-optimised model (Wold et al. 1993; Mukherjee et al. 2015). VIP scores are used in the PLS model to summarise the impact of individual factors, i.e., the independent variables. The amount of explained variance for the dependent variable in each extracted latent dimension is taken into consideration while calculating VIP scores, which are calculated as the weighted sum of squares of the PLS weights (Farres et al. 2015; Das and Gupta 2021). The research works of Wold et al. (1993) and Mukherjee et al. (2015) comprise a detailed methodology and equations for the PLS-VIP approach. This approach was applied because it scores the independent variables quantitatively according to their relative weight. XLSTAT software was used to calculate PLS-VIP values. The PLS-VIP scores were generated for all the flood conditioning factors (Fig. 3). However, those factors (soil, distance from the river, LULC, and flow accumulation) whose scores were less than the threshold value of 0.6 were excluded (Saha et al. 2022) (Table 4).

Table 2. Decision matrix and respective weights of the criteria as derived from the FAHP model

Parameters	Elevation	Slope	Rainfall	Flow Accumulation	Distance from the river	TWI	DD	TRI	LULC	Curvature	Groundwater level	Soil	Geology	AHP Weight	FAHP Weight
Elevation	1	2	2	2	3	3	4	5	5	6	6	6	8	0.193	0.189
Slope	1/2	1	2	2	4	4	3	3	4	5	5	5	7	0.162	0.159
Rainfall	1/2	1/2	1	2	2	3	4	3	4	4	5	5	7	0.132	0.135
Flow Accumulation	1/2	1/2	1/2	1	2	3	3	3	4	5	5	5	7	0.118	0.123
Distance from the river	1/3	1/4	1/2	1/2	1	2	3	3	3	4	4	4	6	0.089	0.090
TWI	1/3	1/4	1/3	1/3	1/2	1	2	2	2	3	4	4	6	0.068	0.068
DD	1/4	1/3	1/4	1/3	1/3	1/2	1	2	2	3	4	5	5	0.060	0.059
TRI	1/5	1/3	1/3	1/3	1/3	1/2	1/2	1	2	3	3	4	5	0.052	0.052
LULC	1/5	1/4	1/4	1/4	1/3	1/2	1/2	1/2	1	2	3	4	4	0.042	0.041
Curvature	1/6	1/5	1/4	1/5	1/4	1/3	1/3	1/3	1/2	1	2	2	3	0.028	0.028
Groundwater level	1/6	1/5	1/5	1/5	1/4	1/4	1/3	1/3	1/3	1/2	1	2	2	0.023	0.023
Soil	1/6	1/5	1/5	1/5	1/4	1/5	1/4	1/4	1/4	1/2	1/2	1	2	0.019	0.019
Geology	1/8	1/7	1/7	1/7	1/6	1/5	1/5	1/5	1/4	1/3	1/2	1/2	1	0.014	0.014

Lambda = 14.12 CI = 0.0937 CI/RI = 0.060

Table 3. Scores for respective classes of the parameters, which were only classified into 5 classes

Class	Class I	Class II	Class III	Class IV	Class V	AHP Weight
Class I	1	2	4	7	9	0.454
Class II	1/2	1	3	5	8	0.295
Class III	1/4	1/3	1	4	5	0.151
Class IV	1/7	1/5	1/4	1	3	0.066
Class V	1/9	1/8	1/5	1/3	1	0.034

Lambda = 5.296 CI = 0.074 CI/RI = 0.066

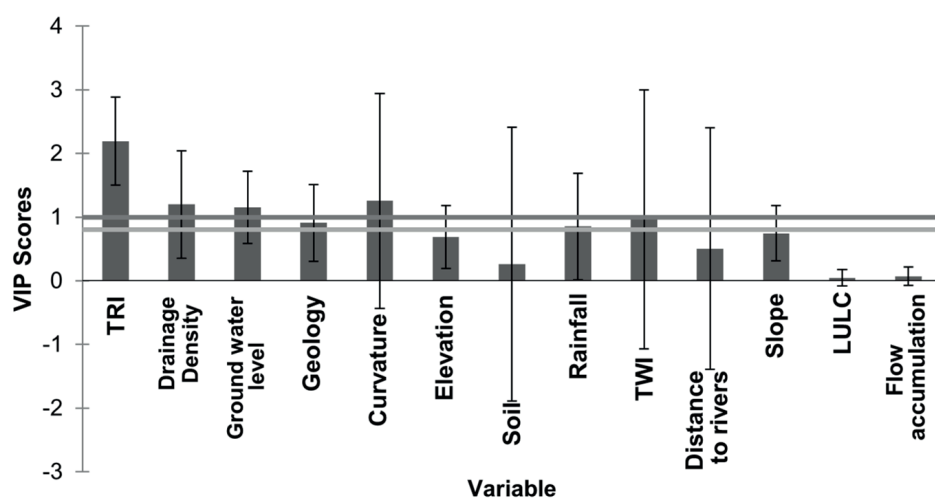


Fig. 3. PLS-VIP values of the used flood conditioning criteria in the FAHP model

Table 4. Determination of PLS-VIP enhanced FAHP weights

Factors	FAHP Weight	VIP	Product	Normalised	PLS-VIP Enhanced Weight
Elevation	0.189	0.687	0.132	0.191	19.131
Slope	0.159	0.745	0.120	0.174	17.414
TWI	0.068	0.962	0.065	0.094	9.438
TRI	0.052	2.193	0.114	0.164	16.453
Curvature	0.028	1.253	0.035	0.050	5.062
Rainfall	0.135	0.855	0.112	0.162	16.284
Groundwater level	0.023	1.154	0.026	0.038	3.829
DD	0.059	1.199	0.071	0.103	10.380
Geology	0.014	0.993	0.013	0.020	2.005

Support Vector Machine (SVM)

SVM is a widely used prediction model for hazard susceptibility analysis. It is present with multiple linear indicating functions (Vapnik 2013). Here, the data transformation in the SVM model was done using the Kernel mathematical system. The training dataset was used for preparing a hyperplane while converting the SVM dataset to high high-dimensional data space. Since the suitability of the Kernel function determines the efficiency of the SVM model output, the radial basis function (RBF) is usually of high regard among other kernel functions (Islam et al. 2021a). The RBF is flexible in dealing with multiple dimensions, and it has a better generalisation ability. These qualities make this model highly utilised for flood susceptibility modelling (Yang and Cervone 2019). The selection of variables is very important in this model because the aptness of the controlling factors reduces the dimensionality of the model space and noise within the

model. Thus, it increases the model accuracy (Shahabi et al. 2020).

Artificial Neural Network (ANN)

ANNs are intricate ML models that contain networks of connected nodes and neurons. A vector of feature values is the input for every neuron. It multiplies this input vector by a column vector of weights and adds a weight called “bias”. Ultimately, the product of this multiplication is fed into an activation function, which yields an output (Atkinson and Tatnall 1997). This study models flood susceptibility using an ANN model with input, hidden and output layers using a Back-Propagation (BP) and error-correction learning technique. Back-propagation is actually used while training the dataset to determine each neuron's weight (Gurney 1997). During training, the magnitude of the update to neuron weights is

controlled by the learning rate, a hyperparameter of the ANN. To avoid overfitting to the training set, the α regularisation term limits the size of the weights. The BP helps in analysing the non-linear relationships with the set of criteria and flood events (occurrence points) (Dodangeh et al. 2020).

Here, all the FSZ maps were divided into five classes. Selection of five classes is ideal for these kinds of studies, as it provides a certain range for rasters, and selecting a greater number of classes will not change the understanding significantly if the method of classification remains the same (Yang et al. 2023). Since the rasters do not have a significant range, reducing the issue of having high outliers, we opted for a quantile method to classify FSZ rasters. Dividing raster values into five classes is quite convenient as it gives us extra gradation between low-to-medium and medium-to-high intensity. Dividing into a greater number of classes could lead to insignificant dilution of values into classes, which are visually distorted but also inappropriate for preparing any planning schemes based on such gradations. More than the number of classes, what matters most is the model that is actually being used. With the presence of outliers, we could have used other models rather than the Quartile method (ESRI 2024).

Validation

A thorough quantitative technique for accuracy assessment is the receiver operating characteristic (ROC) curve (Al-Juaidi et al. 2018). It is a globally recognised and widely used method for accuracy testing because of its comprehensive and easy-to-understand outlook (Chen et al. 2021). Using the area under the curve (AUC) method, the ROC method authenticated the model by comparing known data with likely flood points. The AUC ranges from 0.5 (the lowest accuracy) to 1 (the highest accuracy) (Al-Abadi and Al-Najar 2020). Particularly, there is a poor (0.5–0.6), medium (0.6–0.7), good (0.7–0.8), very good (0.8–0.9), and excellent (0.9–1) correlation between AUC and the accuracy of spatial model prediction (Yesilnacar and Topal 2005). In this study, the ROC process was done in a two-step process. One hundred and 200 ground control points (GCPs) were selected separately from the flood extent map prepared by Dartmouth Flood Observatory (<https://floodobservatory.colorado.edu/>) in a non-biased way (randomly). Firstly, the 100 points

were taken only from the inundated area only, and, for the 200 points procedure, points were randomly selected from both the inundated and non-inundated areas (Fig. 1). A 70:30 ratio of training and testing points was considered for testing the FSZ maps from the first set of GCPs, and later, the entire 200 points were selected to test the accuracy of the FSZ. Here, only the two highest flood susceptibility classes (high, moderately high) were considered as the probable flood inundation extent.

Results

The class respective areal coverage for each of the factors has been provided in the supplement section (TS1), and the visualisation of it is illustrated in Fig. 4. Here, except for the thematic rasters (Geology, soil texture and LULC), variables were classified using the natural break models because of the presence of outliers in their respective distribution.

Collinearity among flood conditioning factors

The results of the collinearity tests indicate that geology has the highest VIF, followed by elevation, groundwater level, soil, rainfall, distance from the river, slope, TWI, LULC, DD, flow accumulation, curvature, and TRI, respectively (Table 5). Here, VIF and TOL values among flood conditioning factors vary from 6.10 to 1.04 and 0.96 to 0.16. Consequently, there is no severe multi-collinearity problem among the chosen flood contributing factors. Thus, not a single criterion was excluded from the selected set of criteria for flood susceptibility modelling. The FSZ prepared using FAHP, SVM and ANN has used all the selected criteria except the model, where PLS-VIP based statistical enhancement has been done.

Flood susceptibility mapping

Fuzzy AHP (FAHP)

Taking into consideration all 13 flood conditioning factors, all five distinct flood susceptibility classes (i.e., low, low-to-moderate, moderate, moderate-to-high,

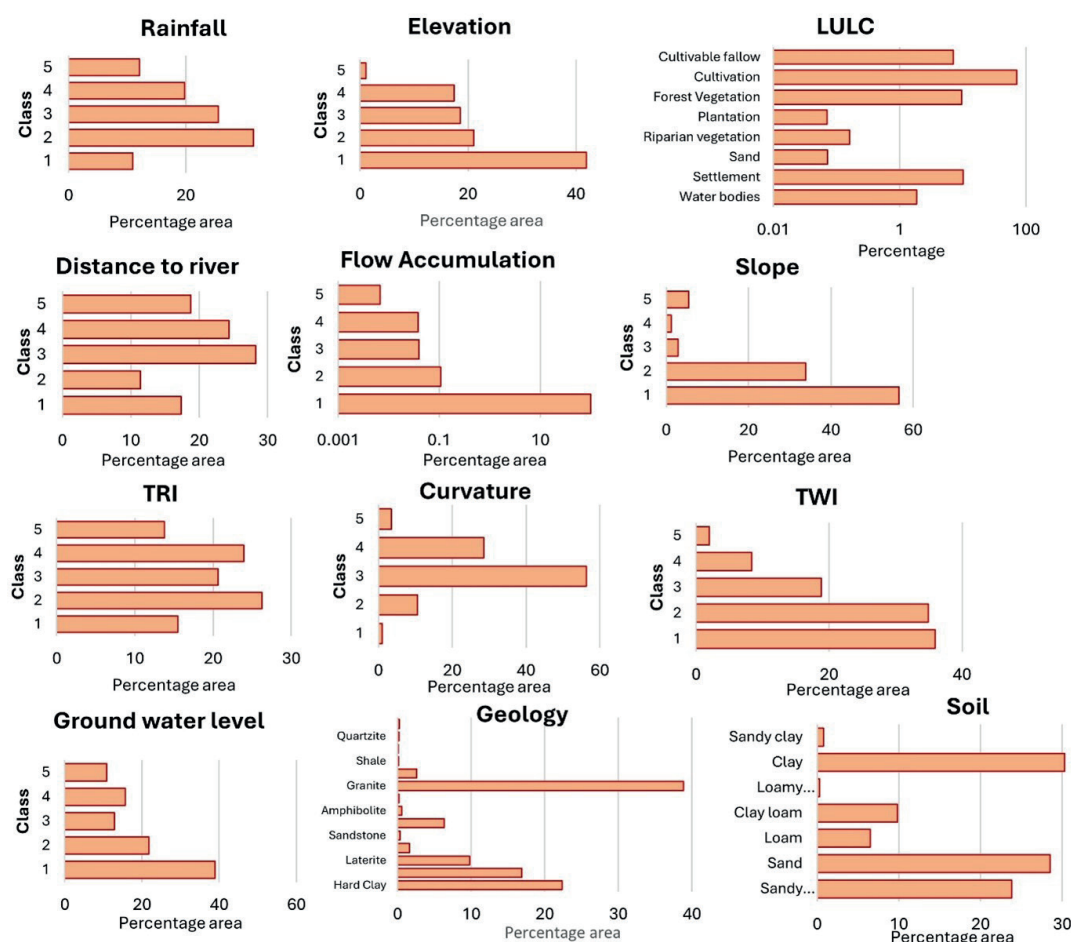


Fig. 4. Area concentration under different classes of the selected rasters. For LULC and Flow accumulation, the x-axis was taken in logarithmic progression for a better representation

Table 5. Collinearity statistics of the selected flood conditioning factors

Statistic	R ²	Tolerance	VIF
Elevation	0.81	0.19	5.22
Slope	0.22	0.78	1.28
TWI	0.19	0.81	1.23
TRI	0.04	0.96	1.04
Curvature	0.10	0.90	1.11
Rainfall	0.54	0.46	2.18
Flow accumulation	0.10	0.90	1.11
Soil	0.57	0.43	2.34
Groundwater level	0.67	0.33	3.04
Distance from the river	0.29	0.71	1.41
DD	0.16	0.84	1.19
LULC	0.18	0.82	1.22
Geology	0.84	0.16	6.10

Supplementary Table 1. Area-wise distribution of Fuzzy AHP based susceptibility class

Susceptibility class (Fuzzy AHP)	Area (km ²)
Low	1,076
Moderately Low	2,745
Moderate	3,176
Moderate to High	1,960
High	635

and high) were identified in the study area by the FAHP model (Fig. 5). The classes covering the least area are high flood susceptibility (635 km²) followed by low (1,076 km²). The moderate flood susceptibility class covers the largest area (3,176 km²). The low-to-moderate flood susceptibility class and the moderate-to-high flood susceptibility class stretch over 2,745 km² and 1,960 km², respectively.

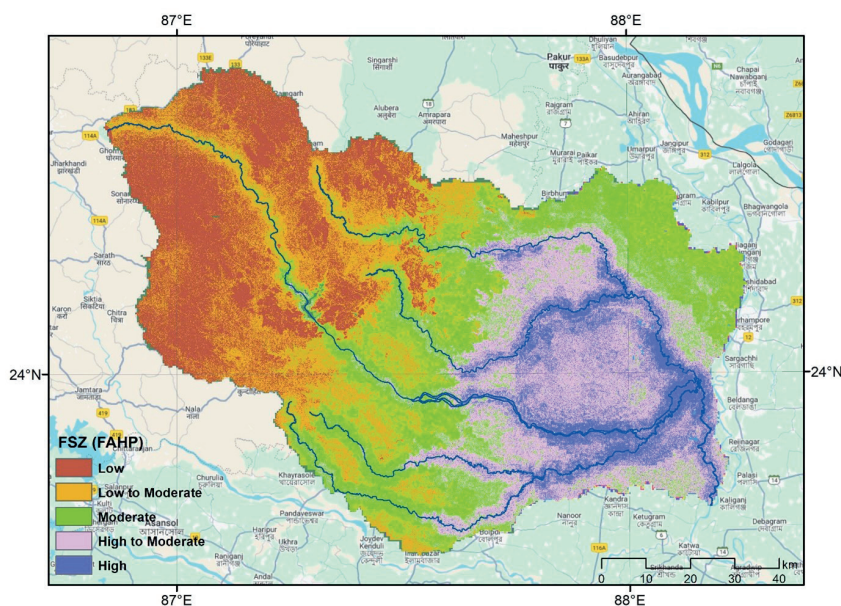


Fig. 5. Flood Susceptibility Zone in MRB, determined by FAHP

Induction of PLS-VIP scores

While applying the PLS-VIP enhanced FAHP method, we excluded soil texture, flow accumulation, distance to rivers, and LULC parameters from the set of criteria used in the FAHP model. These factors have registered PLS-VIP scores below 0.6. Focusing on the excluded factors, it might be justified if the regional geology registered a low VIP score. But it did not – probably due to the testing points being concentrated only within the previously marked flood susceptible zones. Moreover, as the testing points were mostly obtained from a similar kind of lithological unit, and, as the flat floodplains of the downstream reach are composed of fine sand only, the lithological composition within the model's environment develops its significance. One of the shortcomings of this model lies in the selection of the GCPS. Since all the GCPS were taken from the flat plains, which are susceptible to floods, important parameters like distance to the river were excluded due to a low VIP score. However, this is solely based on the primary raster, which is FAHP based FSZ here, and the factors being selected based on their VIP scores are somehow dependent on the weightage primarily applied to the variables during the FAHP modelling. Based on the generated PLS-VIP enhanced FAHP, five susceptibility classes have been identified (Fig. 6). Owing to the exclusion of certain factors, the results obtained completely varied with respect to the FAHP classes. In this case, the minimum area was covered by

the low susceptibility class (587 km²), and the maximum area of 2,659 km² was covered by the moderate susceptibility class, while in FAHP based FSZ, the highest flood susceptible region registered the lowest areal expansion. The low-to-moderate and moderate-to-high susceptibility classes occupied 2,514 and 2,601 km² area, respectively. The high susceptibility class, however, covered 1,190 km², which is nearly double that of the same class in the FAHP-based FSZ.

SVM

Concerning the SVM model, all five distinct flood susceptibility classes have been identified, taking into consideration all flood conditioning factors (Fig. 7). The least area of 566 km² was covered by the moderate flood susceptibility class, followed by high susceptibility (1,052 km²). The moderate-to-high flood susceptibility class covers the largest area (2,640 km²). The low-to-moderate flood susceptibility class and the low flood susceptibility class stretch over 1,322 km² and 4,012 km², respectively.

ANN

The ANN with the back-propagation method for assigning weights to each vector predicted the flood susceptibility map of the MRB (Fig. 8). Keeping the class divisions similar, in this case, the minimum area was covered by the low-to-moderate susceptibility

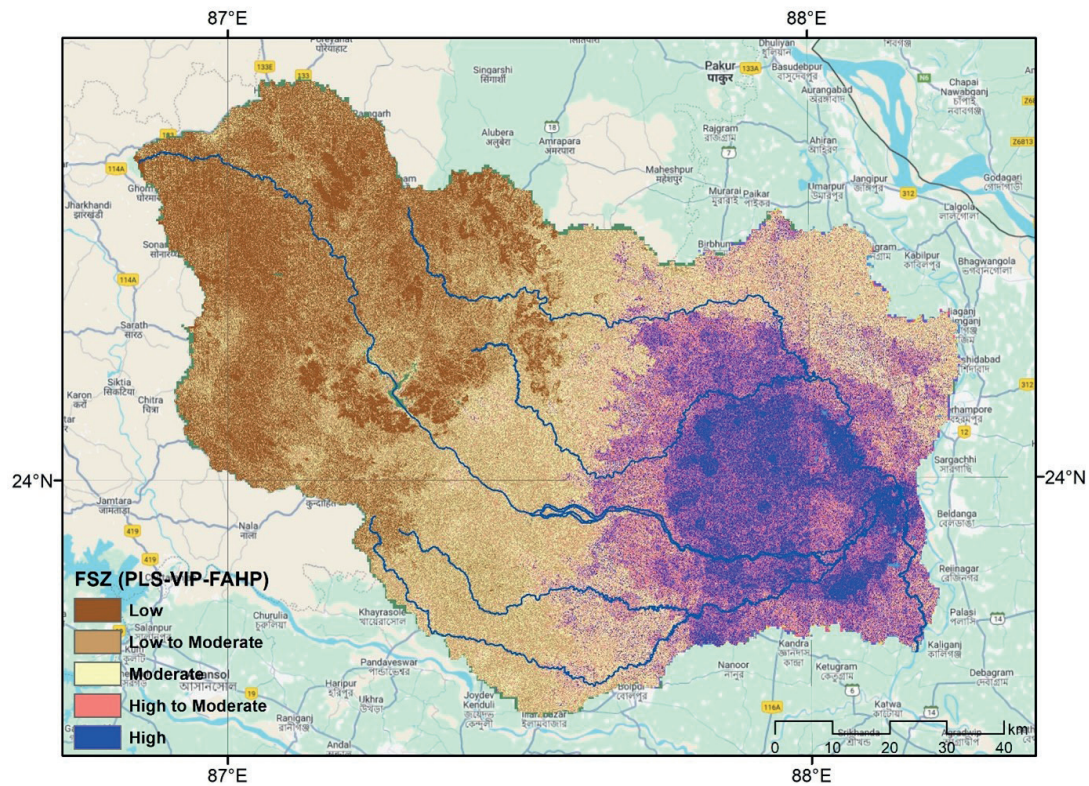


Fig. 6. Flood Susceptibility Zone in the MRB, determined by PLS-VIP-FAHP

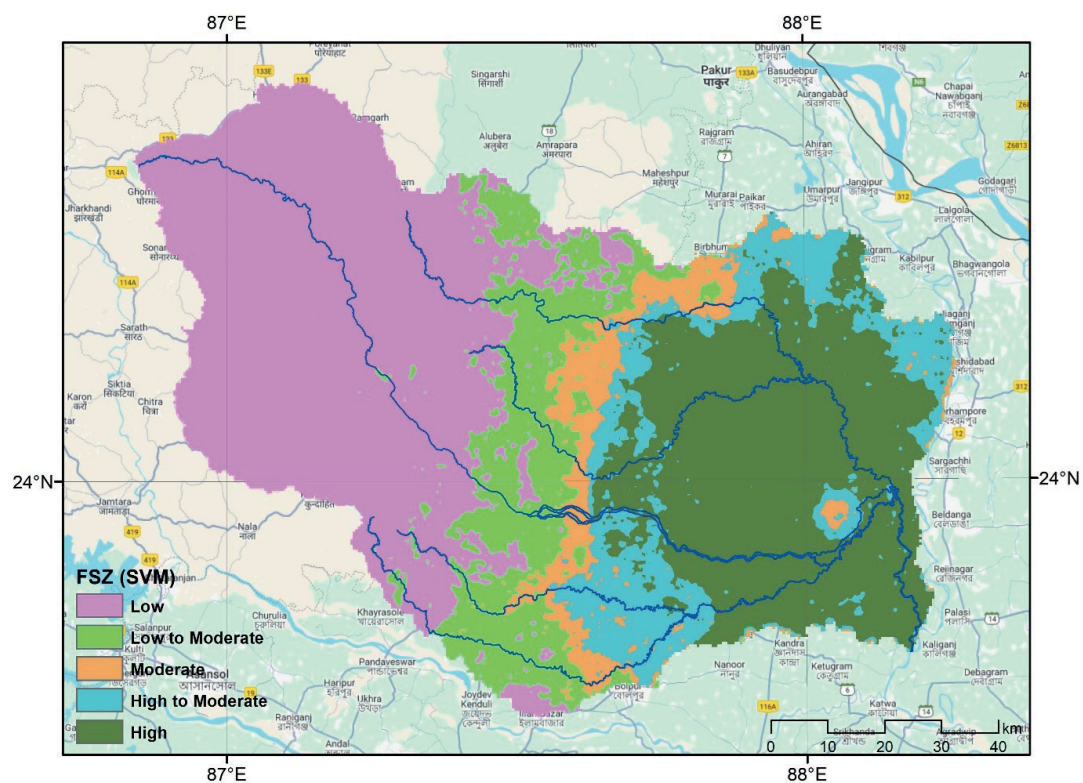


Fig. 7. Flood Susceptibility Zone in the MRB, determined by SVM

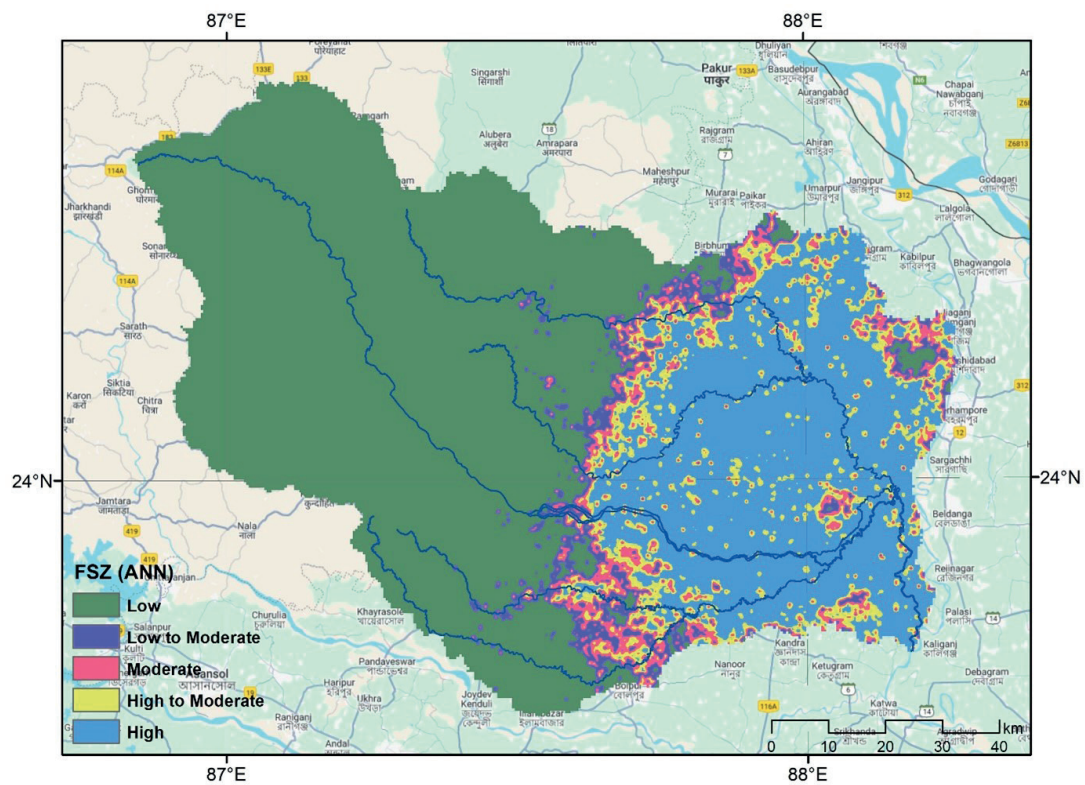


Fig. 8. Flood Susceptibility Zone in the MRB, determined by ANN

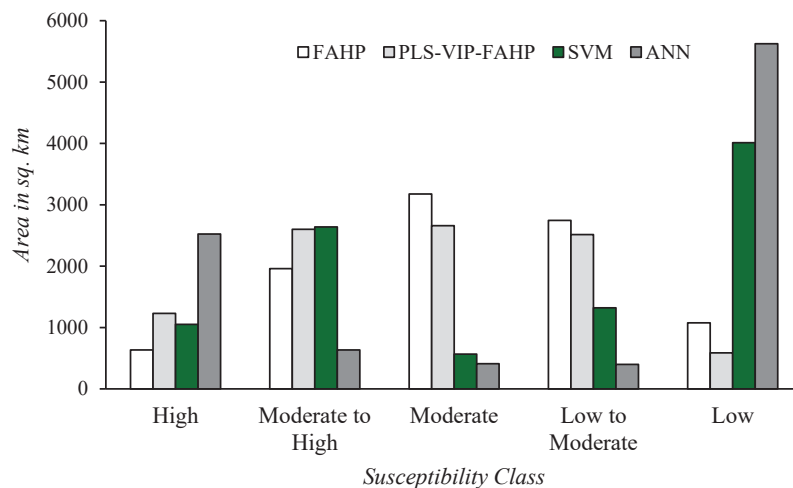


Fig. 9. Areal distribution of different susceptibility classes in the MRB, by model

Validation

class (398 km²), and the largest area (5,624 km²) was covered by the low susceptibility class. The moderate and moderate-to-high susceptibility classes occupied 411 and 635 km², respectively. The high susceptibility class, however, covered 2,523 km² (Fig. 9).

In the present study, the FAHP yields accuracy on the order of ~93% (AUC of 0.93). PLS-VIP-FAHP has shown model accuracy to be ~89% (Fig. 10). SVM has displayed the highest model accuracy of 96% (AUC of 0.96). Lastly, an accuracy of 92% has been shown by

ANN. Hence, all the models have performance levels in the category of “excellent”, except PLS-VIP-FAHP, which performed slightly lower (“very good”). Higher RMSE values by FAHP (0.880) and PLS-VIP-FAHP (0.930) models implied less accurate predictions. On the other hand, the lower RMSE values for the SVM and ANN models make accurate predictions and fit the data well. The SVM model outperformed all the models and emerged as the best-fit model (Table 6). However, in the case of testing the models with 200 points, the degree of accuracy changes for each of the models. PLS-VIP-FAHP registered the lowest value, with 0.80, followed by FAHP, with 0.89, while both the SVM and ANN registered an AUC of 0.90. Apart from testing with a greater number of points, the randomness of the points could have played a significant role in reducing the accuracy levels of the models (Table 7). Apart from these exercises, the SVM model was separately applied using the same dataset except the rainfall data. Here, we tested with a more recent rainfall dataset, that of

CHIRPS (Fig. S1). We prepared the rainfall raster by averaging the annual rainfall from 1991 to 2020 and simulated the model by replacing the previously mentioned Worldclim dataset to verify whether the recent data can affect the flood susceptibility scenario (Fig. S2). This simulation has registered a marginal change in the AUC value, 0.86 (Fig. S3), which was previously 0.90 using the Worldclim data. For validation of this simulation we used the second set of GCPs (200 points with a 70:30 training-to-testing ratio).

Discussion

Researchers of the global scientific fraternity have used MCDM techniques to predict flood susceptibility in different river basins located in different climatic regions (Pradhan 2010; Rahmati et al. 2016). This

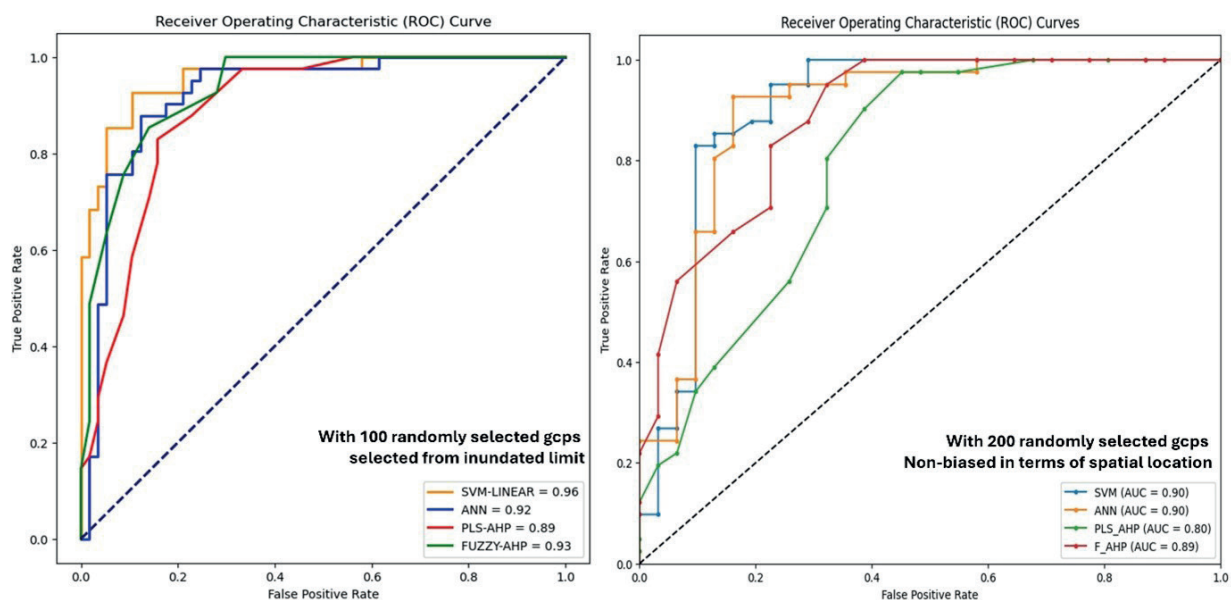


Fig. 10. ROC curves for FAHP, PLS-VIP enhanced FAHP model, ANN and SVM used in predicting FSZ in the MRB with a training-testing ratio of 70:30

Table 6. Model accuracy statements, tested with 100 control points divided at a 70:30 ratio

Model	AUC	Std. Error	RMSE
Fuzzy AHP	0.930	0.068	0.880
PLS-VIP Enhanced FAHP	0.890	0.072	0.930
SVM	0.960	0.020	0.336
ANN	0.920	0.020	0.323

Table 7. Model accuracy statements for testing with 200 control points

Model	AUC	Std. Error	RMSE
Fuzzy AHP	0.89	0.04	0.5
PLS-VIP Enhanced FAHP	0.80	0.07	0.55
SVM	0.90	0.05	0.45
ANN	0.90	0.03	0.39

rainfed tropical river, the Mayurakshi, has a significant amplitude of annual discharges during the monsoon period, and this is evident in frequent flood occurrences (Islam and Sarkar 2020). In this study, the selected flood conditioning criteria and the preferred models delivered a significant result with an excellent fit to the previously observed scenario of flood susceptibility in MRB. Multiple studies have dealt with the flooding nature in this river basin (Islam and Sarkar 2020; Islam and Barman 2020; Islam and Ghosh 2021; Islam et al. 2021b) and tried showcasing the controlling factors at specific sub-catchment levels (Islam and Barman 2020). The overall results suggest the dominance of the topographic factors in determining the extent of flood inundation. Similar results have also been obtained in the works of Tehrany et al. (2015) and Zhu et al. (2024) in the inundated lands of Kuala Terengganu and Malaysia, and in the Yangtze River Delta, respectively. Since the region has varied terrain characteristics, the assumption that increasing distance from the river moderates the chances of flood inundation does not fit here. Apart from that, topographic dominance reduces the role of finer soil texture and LULC properties. Khosravi et al. (2019) highlighted that soil permeability and changes in land use and land cover patterns were the most notable factors, rather than topography, in Northern Iran, contrasting with the Mayurakshi case, where terrain controls dominate. The congestion of drainage, flat floodplains, and convergence of multiple rivers within a small distance also reduces the impact of in-channel flow accumulation. However, excluding the criteria whose impacts are somehow suppressed by the topographic controls and drainage orientation does not improve the significance of the model built-up, as the model accuracy of PLS-VIP-FAHP was 4% lower than that of FAHP. Most likely, on the flat plains, the converging drainage arteries have developed low inter-channel distances in the lower reach of the Mayurakshi River. This increases the degree of flow accumulation. In each of the models, this region was found to have with highest flood susceptibility value. Lack of slope downstream and inundation of the river mouth region during flood situations in the Bhagirathi-Hooghly River restricts flow passage from the basin (Sen et al. 2020). This is quite common for eastwards-flowing rivers joining the Bhagirathi-Hooghly channel in the lower deltaic part of the Ganges. Mostly, the lower reach of the basin is found to be highly susceptible to flood, as is also verified from the flood observation raster. The confluence region of the Dwaraka, Mayurakshi and Kuye Rivers creates

a converging area of river discharge, and the junction area becomes very narrow to flush the entire discharge out (Islam et al. 2021a). These considerations helped in the selection of the flood conditioning factors. Since the topographic control is dominant, the terrain property was set at the highest tier while assigning weights to the criteria. Similar causal considerations have been used in several studies that registered a greater degree of model fitting (Saha et al. 2022).

This study has raised the issue of pertinent selection of conditioning factors based on the study area's geomorphic character and assessed even how reducing the number of factors using relevant statistical techniques can lower the model's accuracy. This issue does not hamper the replicability of the methodology in other, different regions because our selection of relevant factors improves accuracy in determining flood susceptibility rather than testing a model using a generalised suite of globally applied conditioning factors (Yang et al. 2023). The emphasis should be on following the core methodology rather than on replicating the factors employed for the particular study area of the present study. A similar claim has also been made by Wang et al. (2021), where factors were chosen based on their significance concerning the nature of the study area.

Conclusions

This study identifies the potential sites for flood inundation in the MRB, where floods threaten the overall infrastructure and cause rampant damage to agriculture, livelihoods and human settlements. Here, a selection of relevant flood conditioning factors for this study area prepared the baseline for higher accuracy of the predictive models, which is significant enough to be considered appropriate for use in developing an efficient flood preparedness strategy and sustainable resource management in the MRB.

- This study compares MCDMs and ML models and proves that a careful selection of parameters and prior field knowledge applied in assigning relative weights of the criteria can help an MCDM model to outperform ML algorithms.
- Here, the model accuracy of FAHP (AUC=0.93) was higher than the ANN (AUC=0.92), but a lower RMSE for ANN compared to the FAHP indicates a significantly greater association of the ANN model with selected environmental conditions.

- Increasing the number of testing points decreased model accuracy. This is probably due to greater spatiality in terms of distribution of points and induced greater uncertainty in the randomness of distribution.
- The study ascertains that the high flood susceptibility zone covers ~10.97% of the MRB using the SVM model, the best-fit one.
- The occurrence of floods in this basin is the outcome of a combination of natural factors, like low slope gradient, lower capacity of stream channels, higher soil moisture content, and poor infiltration capacity of soils.

The current study put forward an efficacious model that can be employed to demarcate flood-susceptible zones, thereby aiding planners, stakeholders and decision-makers in properly supervising flood-prone areas in the MRB. Furthermore, this will aid in mitigating forthcoming flood vulnerability and damage and in ensuring appropriate and sustainable socio-economic development. The coupling of certain machine learning models and the incorporation of ensemble algorithms using non-stationary models with AHP could be considered for future work.

Acknowledgment

We are thankful to the anonymous reviewers for their valuable comments. No external funding has been used to conduct this study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Author contributions

Study design: SN, US; data collection: SD; statistical analysis: SN, PKP, SS, SD, KB, US; result interpretation: SN, PKP, SS, KB; manuscript preparation: SN, PKP, SS; literature review: PKP.

References

- AL-ABADI AM and AL-NAJAR NA, 2020, Comparative assessment of bivariate, multivariate and machine learning models for mapping flood proneness. *Natural Hazards* 100(2): 461-491. DOI: <https://doi.org/10.1007/s11069-019-03820-z>.
- ALIYEV R, TEMIZKAN H and ALIYEV R, 2020, Fuzzy analytic hierarchy process-based multi-criteria decision making for universities ranking. *Symmetry* 12(8): 1351. DOI: <https://doi.org/10.3390/sym12081351>.
- AL-JUAIDI AE, NASSAR AM and AL-JUAIDI OE, 2018, Evaluation of flood susceptibility mapping using logistic regression and GIS conditioning factors. *Arabian Journal of Geosciences* 11: 1-10. DOI: <https://doi.org/10.1007/s12517-018-4095-0>.
- ATKINSON PM and TATNALL AR, 1997, Introduction neural networks in remote sensing. *International Journal of Remote Sensing* 18(4): 699-709. DOI: <https://doi.org/10.1080/014311697218700>.
- BAMMOU Y, BENZOUAGH B, IGMOLLAN B, OUALALI A, KADER S, SPALEVIC V, PAUL S, PAOLO B and MARKOVIĆ SB, 2024, Optimizing flood susceptibility assessment in semi-arid regions using ensemble algorithms: a case study of Moroccan High Atlas. *Natural Hazards* 120(8): 7787-7816. DOI: <https://doi.org/10.1007/s11069-024-06547-8>.
- BANDYOPADHYAY S, KAR NS, DAS S and SEN J, 2014, River systems and water resources of West Bengal: a review. *Geological Society of India Special Publication* 3: 63-84.
- BARAL U, SAHA UD, MUKHOPADHYAY U and SINGH D, 2023, Drought risk assessment on the eastern part of Indian peninsula—a study on Purulia district, West Bengal. *Environmental Monitoring and Assessment* 195(11): 1364. DOI: <https://doi.org/10.1007/s10661-023-11986-0>.
- BRITO DE MM and EVERS M, 2016, Multi-criteria decision-making for flood risk management: a survey of the current state of the art. *Natural Hazards and Earth System Sciences* 16(4): 1019-1033. DOI: <https://doi.org/10.5194/nhess-16-1019-2016>.
- CAO C, XU P, WANG Y, CHEN J, ZHENG L and NIU C, 2016, Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas. *Sustainability* 8(9): 948. DOI: <https://doi.org/10.3390/su8090948>.
- CHATTERJEE S and JANA NC, 2022, Hydro-Meteorological Analysis of 2015 Rarh Bengal Flood in the Lower Gangetic Plain of India: Exceptional, Fast and Furious. In: *Climate, Environment and Disaster in Developing Countries*: 401-424. Springer Nature Singapore. DOI: https://doi.org/10.1007/978-981-16-6966-8_20.

- CHEN J, HUANG G and CHEN W, 2021, Towards better flood risk management: Assessing flood risk and investigating the potential mechanism based on machine learning models. *Journal of Environmental Management* 293: 112810. DOI: <https://doi.org/10.1016/j.jenvman.2021.112810>.
- DAS S, 2018, Geographic information system and AHP-based flood hazard zonation of Vaitarna basin, Maharashtra, India. *Arabian Journal of Geosciences* 11(19): 576. DOI: <https://doi.org/10.1007/s12517-018-3933-4>.
- DAS S and GUPTA A, 2021, Multi-criteria decision based geospatial mapping of flood susceptibility and temporal hydro-geomorphic changes in the Subarnarekha basin, India. *Geoscience Frontiers* 12(5): 101206. DOI: <https://doi.org/10.1016/j.gsf.2021.101206>.
- DODANGEH E, CHOUBIN B, EIGDIR AN, NABIPOUR N, PANAHI M, SHAMSHIRBAND S and MOSAVI A, 2020, Integrated machine learning methods with resampling algorithms for flood susceptibility prediction. *Science of the Total Environment* 705: 135983. DOI: <https://doi.org/10.1016/j.scitotenv.2019.135983>.
- FARRÉS M, PLATIKANOV S, TSAKOVSKI S and TAULER R, 2015, Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *Journal of Chemometrics* 29(10): 528-536. DOI: <https://doi.org/10.1002/cem.2736>.
- GHOSH A and KAR SK, 2018, Application of analytical hierarchy process (AHP) for flood risk assessment: a case study in Malda district of West Bengal, India. *Natural Hazards* 94: 349-368. DOI: <https://doi.org/10.1007/s11069-018-3392-y>.
- GHOSH A and MAITI R, 2021, Soil erosion susceptibility assessment using logistic regression, decision tree and random forest: study on the Mayurakshi river basin of Eastern India. *Environmental Earth Sciences* 80: 1-16. DOI: <https://doi.org/10.1007/s12665-021-09669-5>.
- GHOSH KG and PAL S, 2015, Impact of dam and barrage on flood trend of lower catchment of Mayurakshi river basin, eastern India. *European Water* 50: 3-23.
- GOKCEOGLU C, SONMEZ H, NEFESLIOGLU HA, DUMAN TY and CAN T, 2005, The 17 March 2005 Kuzulu landslide (Sivas, Turkey) and landslide-susceptibility map of its near vicinity. *Engineering Geology* 81(1): 65-83. DOI: <https://doi.org/10.1016/j.enggeo.2005.07.011>.
- GURNEY K, 1997, *Multilayer nets and backpropagation*. Dept. Human Sciences, Brunel University, Uxbridge, Middx. UK.
- ISLAM ARMT, JIBON MJN, SIDDIK MA, MALLICK J, ALQADHI S, PAL SC, PUROHIT S and SENAPATHI V, 2025, Predicting flood risks using advanced machine learning algorithms with a focus on Bangladesh: influencing factors, gaps and future challenges. *Earth Science Informatics* 18(3): 300. DOI: <https://doi.org/10.1007/s12145-025-01719-x>.
- ISLAM A and BARMAN SD, 2020, Drainage basin morphometry and evaluating its role on flood-inducing capacity of tributary basins of Mayurakshi River, India. *SN Applied Sciences* 2(6): 1087. DOI: <https://doi.org/10.1007/s42452-020-2900-3>.
- ISLAM A and GHOSH S, 2021, Economic transformation in the wake of flood: a case of the lower stretch of the Mayurakshi River Basin, India. *Environment, Development and Sustainability* 23(10): 15550-15590. DOI: <https://doi.org/10.1007/s10668-021-01311-5>.
- ISLAM A and SARKAR B, 2020, Analysing flood history and simulating the nature of future floods using Gumbel method and log Pearson Type III: The case of the Mayurakshi River Basin, India. *Bulletin of Geography. Physical Geography Series* 19(1): 43-69. DOI: <https://doi.org/10.2478/bgeo-2020-0010>.
- ISLAM A, PAL SC, CHOWDHURI I, SALAM R, ISLAM MS, RAHMAN MM, ZAHID A and IDRIS AM, 2021a, Application of novel framework approach for prediction of nitrate concentration susceptibility in coastal multi-aquifers, Bangladesh. *Science of The Total Environment* 801: 149811. DOI: <https://doi.org/10.1016/j.scitotenv.2021.149811>.
- ISLAM A, TALUKDAR S, MAHATO S, KUNDU S, EIBEK KU, PHAM QB, KURIQI A and LINH N, 2021b, Flood susceptibility modelling using advanced ensemble machine learning models. *Geoscience Frontiers* 12(3): 101075. DOI: <https://doi.org/10.1016/j.gsf.2020.09.006>.
- JAFARZADEGAN K, MORADKHANI H, PAPPENBERGER F, MOFTAKHARI H, BATES P, ABBASZADEH P, MARSOOLI R, FERREIRA C, CLOKE HL, OGDEN F and DUAN Q, 2023, Recent advances and new frontiers in riverine and coastal flood modeling. *Reviews of Geophysics* 61(2): e2022RG000788. DOI: <https://doi.org/10.1029/2022RG000788>.
- KARIMINEJAD N, HOSSEINALIZADEH M, POURGHASEMI HR, BERNATEK-JAKIEL A and ALINEJAD M, 2019, GIS-based susceptibility assessment of the occurrence of gully headcuts and pipe collapses in a semi-arid environment: Golestan Province, NE Iran. *Land Degradation & Development* 30(18): 2211-2225. DOI: <https://doi.org/10.1002/ldr.3396>.
- KAUSHIK K, PANDEY AC, PARIDA BR and DWIVEDI CS, 2025, Geoinformatics-based watershed-level flood hazard analysis using fuzzy analytic hierarchy process across the Ganga-Brahmaputra basin. *Journal of Earth System Science* 134(3): 1-22. DOI: <https://doi.org/10.1007/s12040-025-02593-3>.
- KHODAEI H, NASIRI SALEH F, NOBAKHT DALIR A and ZAREI E, 2025, Future flood susceptibility

- mapping under climate and land use change. *Scientific Reports* 15(1): 12394. DOI: <https://doi.org/10.1038/s41598-025-96713-8>.
- KHOSRAVI K, SHAHABI H, PHAM BT, ADAMOWSKI J, SHIRZADI A, PRADHAN B, DOU J, LY HB, GRÓF G, HO HL and HONG H, 2019, A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. *Journal of Hydrology* 573: 311-323. DOI: <https://doi.org/10.1016/j.jhydrol.2019.03.073>.
- LIN J, SRENG C, OARE E and BATARSEH FA, 2023, NeuralFlood: an AI-driven flood susceptibility index. *Frontiers in Water* 5: 1291305. DOI: <https://doi.org/10.3389/frwa.2023.1291305>.
- LYU HM, SHEN SL, YANG J and ZHOU AN, 2020, Risk Assessment of earthquake-triggered geo-hazards surrounding Wenchuan, China. *Natural Hazards Review* 21: 05020007. DOI: [https://doi.org/10.1061/\(ASCE\)NH.1527-6996.0000387](https://doi.org/10.1061/(ASCE)NH.1527-6996.0000387).
- MAHAD NF, YOSUF NM and ISMAIL NF, 2019, The application of fuzzy analytic hierarchy process (FAHP) approach to solve multi-criteria decision making (MCDM) problems. *Journal of Physics: Conference Series*. DOI: <https://doi.org/10.1088/1742-6596/1366/1/012058>.
- MAHMOUD SH and GAN TY, 2018, Multi-criteria approach to develop flood susceptibility maps in arid regions of Middle East. *Journal of Cleaner Production* 196: 216-229. DOI: <https://doi.org/10.1016/j.jclepro.2018.06.047>.
- MISHRA V, NANDITHA JS, DANGAR S, CHUPHAL DS and VEGAD U, 2024, Drivers, changes, and impacts of hydrological extremes in India: A review. *Wiley Interdisciplinary Reviews: Water* 11(5): e1742. DOI: <https://doi.org/10.1002/wat2.1742>.
- MISRA A, WHITE K, NSUTEZO SF, STRAKA III W and LAVISTA J, 2025, Mapping global floods with 10 years of satellite radar data. *Nature Communications* 16(1): 5762. DOI: <https://doi.org/10.1038/s41467-025-60853-4>.
- MOORE ID, GRAYSON RB and LADSON AR, 1991, Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes* 5(1): 3-30. DOI: <https://doi.org/10.1002/hyp.3360050103>.
- MUDASHIRU RB, SABTU N, ABUSTAN I and BALOGUN W, 2021, Flood hazard mapping methods: A review. *Journal of Hydrology* 603: 126846. DOI: <https://doi.org/10.1016/j.jhydrol.2021.126846>.
- MUKHERJEE A, FRYAR AE and THOMAS WA, 2009, Geologic, geomorphic and hydrologic framework and evolution of the Bengal basin, India and Bangladesh. *Journal of Asian Earth Sciences* 34(3): 227-244. DOI: <https://doi.org/10.1016/j.jseaes.2008.05.011>.
- MUKHERJEE R, SENGUPTA D and SIKDAR SK, 2015, Selection of sustainable processes using sustainability footprint method: A case study of methanol production from carbon dioxide. In: *Computer Aided Chemical Engineering* 36: 311-329. Elsevier. DOI: <https://doi.org/10.1016/B978-0-444-63472-6.00012-4>.
- MUKHOPADHYAY S and LET S, 2014, Changing flood intensity zone of Dwarka river basin in eastern India. *Transactions* 36(1).
- NANDI A, MANDAL A, WILSON M and SMITH D, 2016, Flood hazard mapping in Jamaica using principal component analysis and logistic regression. *Environmental Earth Sciences* 75(6): 465. DOI: <https://doi.org/10.1007/s12665-016-5323-0>.
- OFFICE OF THE DISTRICT MAGISTRATE, 2014, Flood preparedness and Management Plan, Murshidabad District Govt. of West Bengal.
- OFFICE OF THE DISTRICT MAGISTRATE, 2016, District Disaster Management Plan, Murshidabad District. Govt. of West Bengal.
- O'MALLEY LSS, 1914, *Bengal District Gazetteers Murshidabad*. Calcutta, Bengal Sekretariat Book Depot.
- PAHARI S, DAS CHATTERJEE N and BARMAN NK, 2024, Noise vulnerability assessment using an integrated approach of multi-criteria decision-making model and geospatial techniques. *International Journal of Environmental Science and Technology* 21(6): 5421-5436. DOI: <https://doi.org/10.1007/s13762-023-05394-7>.
- PAL S, 2016, Impact of Massanjore dam on hydro-geomorphological modification of Mayurakshi river, Eastern India. *Environment, Development and Sustainability* 18: 921-944. DOI: <https://doi.org/10.1007/s10668-015-9676-4>.
- PENG G, HAN L, LIU Z, GUO Y, YAN J and JIA X, 2021, An application of fuzzy analytic hierarchy process in risk evaluation model. *Frontiers in Psychology* 12: 715003. DOI: <https://doi.org/10.3389/fpsyg.2021.715003>.
- POURGHASEMI HR, PRADHAN B and GOKCEOGLU C, 2012, Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Natural Hazards* 63: 965-996. DOI: <https://doi.org/10.1007/s11069-012-0217-2>.
- PRADHAN B, 2010, Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. *Journal of Spatial Hydrology* 9(2).
- PRISCILLIA S, SCHILLACI C and LIPANI A, 2021, Flood susceptibility assessment using artificial neural networks in Indonesia. *Artificial Intelligence in Geosciences* 2: 215-222. DOI: <https://doi.org/10.1016/j.aiig.2022.02.001>.

- RAHMATI O, POURGHASEMI HR and ZEINIVAND H, 2016, Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golistan Province, Iran. *Geocarto International* 31(1): 42-70. DOI: <https://doi.org/10.1080/10106049.2015.1041559>.
- SAHA UD, MITRA S, MONDAL M and BISWAS N, 2022, Assessment of the impact of anthropogenic stress on wetlands of the upper Ganga Delta, India: a multi criteria decision making approach. *Geocarto International* 37(26): 12580-12601. DOI: <https://doi.org/10.1080/10106049.2022.2068670>.
- SALMERÓN GÓMEZ R, GARCÍA PÉREZ J, LÓPEZ MARTÍN MDM and GARCÍA CG, 2016, Collinearity diagnostic applied in ridge estimation through the variance inflation factor. *Journal of Applied Statistics* 43(10): 1831-1849. DOI: <https://doi.org/10.1080/02664763.2015.1120712>.
- SALVATI A, NIA AM, SALAJEGHEH A, GHADERI K, ASL DT, AL-ANSARI N, SOLAIMANI F and CLAGUE JJ, 2023, Flood susceptibility mapping using support vector regression and hyper-parameter optimization. *Journal of Flood Risk Management* 16(4): e12920. DOI: <https://doi.org/10.1111/jfr3.12920>.
- SEN A, SAHA UD, MAJUMDER A and BISWAS N, 2020, Facets of Anthropogenic Encroachment within the Palaeo-Fluvial Regime of Kana-Ghia-Kunti System, Damodar Fan Region. In: *Anthropogeomorphology of Bhagirathi-Hooghly River System in India*: 491-520. CRC Press.
- SEN D, 2010, Flood hazards in India and management strategies. In: *Natural and Anthropogenic Disasters: Vulnerability, Preparedness and Mitigation*: 126-146. Dordrecht: Springer Netherlands. DOI: https://doi.org/10.1007/978-90-481-2498-5_7.
- SHAHABI H, SHIRZADI A, GHADERI K, OMIDVAR E, AL-ANSARI N, CLAGUE JJ, GEERTSEMA M, KHOSRAVI K, AMINI A, BAHRAMI S and 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. DOI: <https://doi.org/10.3390/rs12020266>.
- SHAHABI H, SHIRZADI A, RONOUD S, ASADI S, PHAM BT, MANSOURIPOUR F, GEERTSEMA M, CLAGUE JJ and BUI DT, 2021, Flash flood susceptibility mapping using a novel deep learning model based on deep belief network, back propagation and genetic algorithm. *Geoscience Frontiers* 12(3): 101100. DOI: <https://doi.org/10.1016/j.gsf.2020.10.007>.
- SHARMA VK, SHUKLA AK, CHOWDARY VM, SARAN S and SRIVASTAV SK, 2025, Integration of recurrent neural network models and flood inundation extent libraries for flood forecasting. *Remote Sensing Applications: Society and Environment* 39: 101649. DOI: <https://doi.org/10.1016/j.rsase.2025.101649>.
- SOUISSI D, ZOUHRI L, HAMMAMI S, MSADDEK MH, ZGHIBI A and DLALA M, 2020, GIS-based MCDM-AHP modeling for flood susceptibility mapping of arid areas, southeastern Tunisia. *Geocarto International* 35(9): 991-1017. DOI: <https://doi.org/10.1080/10106049.2019.1570346>.
- SWAIN KC, SINGHA C and NAYAK L, 2020, Flood susceptibility mapping through the GIS-AHP technique using the cloud. *ISPRS International Journal of Geo-Information* 9(12): 720. DOI: <https://doi.org/10.3390/ijgi9120720>.
- TANG Z, ZHANG H, YI S and XIAO Y, 2018, Assessment of flood susceptible areas using spatially explicit, probabilistic multi-criteria decision analysis. *Journal of Hydrology* 558: 144-158. DOI: <https://doi.org/10.1016/j.jhydrol.2018.01.033>.
- TEHRANY MS, PRADHAN B, MANSOR S and AHMAD N, 2015, Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. *Catena* 125: 91-101. DOI: <https://doi.org/10.1016/j.catena.2014.10.017>.
- VAPNIK VN, 2013, Inductive principles of statistics and learning theory. In: *Mathematical Perspectives on Neural Networks*: 777-841. Psychology Press.
- VOJTEK M and VOJTEKOVÁ J, 2019, Flood susceptibility mapping on a national scale in Slovakia using the analytical hierarchy process. *Water* 11(2): 364. DOI: <https://doi.org/10.3390/w11020364>.
- WOLD S, JOHANSSON E and COCCHI M, 1993, PLS: partial least squares projections to latent structures. In: *3D QSAR in Drug Design: Theory, Methods and Applications*: 523-550. Kluwer ESCOM Science Publisher.
- WU T, LI HC, WEI SP, CHEN WB, CHEN YM, SU YF, LIU JJ and SHIH HJ, 2016, A comprehensive disaster impact assessment of extreme rainfall events under climate change: A case study in Zheng-wen river basin, Taiwan. *Environmental Earth Sciences* 75(7): 597. DOI: <https://doi.org/10.1007/s12665-016-5392-0>.
- YANG D, ZHANG T, ARABAMERI A, SANTOSH M, SAHA UD and ISLAM A, 2023, Flash-flood susceptibility mapping: a novel credal decision tree-based ensemble approaches. *Earth Science Informatics* 16(4): 3143-3161. DOI: <https://doi.org/10.1007/s12145-023-01088-3>.
- YANG L and CERVONE G, 2019, Analysis of remote sensing imagery for disaster assessment using deep learning: a case study of flooding event. *Soft Computing* 23(24): 13393-13408. DOI: <https://doi.org/10.1007/s00500-019-03879-7>.
- YANG XL, DING JH and HOU H, 2013, Application of a triangular fuzzy AHP approach for flood risk evaluation and response measures analysis. *Natural*

- Hazards* 68: 657-674. DOI: <https://doi.org/10.1007/s11069-013-0642-x>.
- YESILNACAR E and TOPAL T, 2005, Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Engineering Geology* 79(3-4): 251-266. DOI: <https://doi.org/10.1016/j.enggeo.2005.02.002>.
- YILMAZ NG, YURDAKUL M and GOKTAN RM, 2007, Prediction of radial bit cutting force in high-strength rocks using multiple linear regression analysis. *International Journal of Rock Mechanics and Mining Sciences* 44(6): 962-970. DOI: <https://doi.org/10.1016/j.ijrmms.2007.02.003>.
- YOUSSEF AM, PRADHAN B and HASSAN AM, 2011, Flash flood risk estimation along the St. Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery. *Environmental Earth Sciences* 62(3): 611-623. DOI: <https://doi.org/10.1007/s12665-010-0551-1>.
- ZHU K, WANG Z, LAI C, LI S, ZENG Z and CHEN X, 2024, Evaluating Factors Affecting Flood Susceptibility in the Yangtze River Delta Using Machine Learning Methods. *International Journal of Disaster Risk Science* 15: 1-16. DOI: <https://doi.org/10.1007/s13753-024-00590-6>.
- ESRI, 2024, Data classification methods. Available at: <https://pro.arcgis.com/en/proapp/latest/help/mapping/layer-properties/data-classification-methods.htm>.
- WANG G, LIU L, SHI P, ZHANG G and LIU J, 2021, Flood Risk Assessment of Metro System Using Improved Trapezoidal Fuzzy AHP: A Case Study of Guangzhou. *Remote Sensing* 13: 5154. DOI: <https://doi.org/10.3390/rs13245154>.

Received 23 August 2024

Accepted 5 November 2025