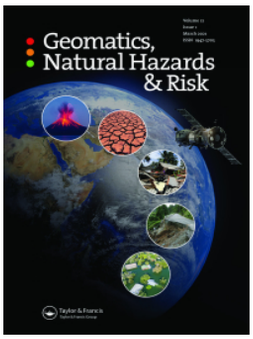

Geomatics, Natural Hazards and Risk

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgnh20

SAR-driven flood inventory and multi-factor ensemble susceptibility modelling using machine learning frameworks

Krishnagopal Halder, Anitabha Ghosh, Amit Kumar Srivastava, Subodh Chandra Pal, Uday Chatterjee, Dipak Bisai, Frank Ewert, Thomas Gaiser, Abu Reza Md. Towfiqul Islam, Edris Alam & Md Kamrul Islam

To cite this article: Krishnagopal Halder, Anitabha Ghosh, Amit Kumar Srivastava, Subodh Chandra Pal, Uday Chatterjee, Dipak Bisai, Frank Ewert, Thomas Gaiser, Abu Reza Md. Towfiqul Islam, Edris Alam & Md Kamrul Islam (2024) SAR-driven flood inventory and multi-factor ensemble susceptibility modelling using machine learning frameworks, Geomatics, Natural Hazards and Risk, 15:1, 2409202, DOI: 10.1080/19475705.2024.2409202

To link to this article: https://doi.org/10.1080/19475705.2024.2409202

Published online: 16 Oct 2024.

GEOMATICS, NATURAL HAZARDS AND RISK
2024, VOL. 15, NO. 1, 2409202
https://doi.org/10.1080/19475705.2024.2409202

SAR-driven flood inventory and multi-factor ensemble susceptibility modelling using machine learning frameworks

Krishnagopal Halder[a,b], Anitabha Ghosh[c], Amit Kumar Srivastava[d,e], Subodh Chandra Pal[f], Uday Chatterjee[c,g], Dipak Bisai[c], Frank Ewert[d,e], Thomas Gaiser[d], Abu Reza Md. Towfiqul Islam[h,i], Edris Alam[j,k] and Md Kamrul Islam[l]

[a]Department of Remote Sensing and GIS, Vidyasagar University, Vidyasagar University, Midnapore, West Bengal, India; [b]Centre of Excellence in Disaster Mitigation and Management (CoEDMM), Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, India; [c]Coastal Environmental Studies Research Centre, Egra S.S.B. College, (Affiliated to Vidyasagar University), Kharagpur, West Bengal, India; [d]Institute of Crop Science and Resource Conservation, University of Bonn, Bonn, Germany; [e]Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany; [f]Department of Geography, The University of Burdwan, Purba Bardhaman, West Bengal, India; [g]Department of Geography, Bhatter College, Kharagpur, West Bengal, India; [h]Department of Disaster Management, Begum Rokeya University, Rangpur, Bangladesh; [i]Department of Development Studies, Daffodil International University, Dhaka, Bangladesh; [j]Faculty of Resilience, Rabdan Academy, Abu Dhabi, United Arab Emirates; [k]Department of Geography and Environmental Studies, University of Chittagong, Chittagong, Bangladesh; [l]Department of Civil and Environmental Engineering College of Engineering, King Faisal University, Saudi Arabia

ABSTRACT

Climate change has substantially increased both the occurrence and intensity of flood events, particularly in the Indian subcontinent, exacerbating threats to human populations and economic infrastructure. The present research employed novel ML models—LR, SVM, RF, XGBoost, DNN, and Stacking Ensemble—developed in the Python environment and leveraged 18 flood-influencing factors to delineate flood-prone areas with precision. A comprehensive flood inventory, obtained from Sentinel-1 Synthetic Aperture Radar (SAR) data using the Google Earth Engine (GEE) platform, provided empirical data for entire model training and validation. Model performance was assessed using precision, recall, F1-score, accuracy, and ROC-AUC metrics. The results highlighted Stacking Ensemble's superior predictive ability (0.965), followed closely by, XGBoost (0.934), DNN (0.929), RF (0.925), LR (0.921), and SVM (0.920) respectively, establishing the feasibility of ML applications in disaster management. The maps depicting susceptibility to flooding generated by the current research provide actionable insights for decision-makers, city planners, and authorities responsible for disaster management, guiding infrastructural and community resilience enhancements against flood risks.

ARTICLE HISTORY
Received 27 March 2024
Accepted 22 September 2024

KEYWORDS
Disaster management; Flood susceptibility; Google Earth Engine; Machine learning; Python

CONTACT Amit Kumar Srivastava AmitKumar.Srivastava@zalf.de; Subodh Chandra Pal geo.subodh@gmail.com

1. Introduction

Flooding, regarded as a perilous and recurrent natural calamity, engenders significant human and economic casualties (Saha et al. 2022) impacting around 200 million people annually (Youssef, Mahdi, et al. 2022). Over the last several decades, there has been a notable rise in the occurrence of such calamities, mostly attributed to climate change. The intensity of these hazards is predicted to worsen due to the consequences of climate change, which include more frequent and severe extreme precipitation (Antzoulatos et al. 2022). Preventive as well as emergency actions, defined as actions taken in the moments leading up to, during, or after a flood, are needed to lessen the effect of floods on the loss of property and human life. Whereas, preventative measures aim to lessen the likelihood that a certain location will flood. Under such circumstances, the execution of countermeasures requires real-time knowledge of the magnitude of the flood and the places at risk (Mitra and Das 2022; Pradhan et al. 2023). Maps indicating flood hazards, or the potential characteristics of a flood event, can be used to assess flood susceptibility and identify areas of flood vulnerability.

From 1998 to 2017, climate-related and geophysical catastrophes led to 4.4 billion people experiencing injuries, homelessness, displacement, or requiring emergency assistance. These events also claimed 1.3 million lives. While geophysical events like earthquakes and tsunamis accounted for the majority of deaths, severe meteorological events such as floods, storms, droughts, and heatwaves-were responsible for 91% of all catastrophes. Disaster-affected nations suffered direct economic losses of US$ 2908 billion between 1998 and 2017, with US$ 2245 billion, or 77% of the total, coming from climate-related catastrophes. This amounts to more than 68% (US$ 895 billion) of losses (US$ 1313 billion) that were documented between 1978 and 1997 (CRED and UNISDR 2016). Between these two 20-year periods, there was a 151% increase in the total recorded losses due to severe weather events. Current research has explored that, besides intense rainfall, recurrent storm surges resulting from tropical cyclones, extensive siltation in riverbeds, and the rivers' limited capacity to manage substantial discharges play crucial roles in exacerbating flooding in India (Mishra and Sinha 2020). Between 1995 and 2015, floods have had a global impact on over 2.3 billion individuals, representing approximately 56% of the overall population affected by hydro-meteorological disasters (CRED and UNISDR 2016).

Multiple studies have conducted assessments indicating that the Indian sub-continent is positioned at the forefront of places that are highly susceptible to river flood hazards, posing threats to both human populations and economic endeavours (Mitra and Das 2022; Saha et al. 2022). As per the findings of the National Institute of Hydrology (NIH 2000), the primary areas prone to flooding in India encompass the deltas and riverbanks of the Ganga, Brahmaputra, Kosi, Mahanadi, Godavari, Krishna, Cauvery, and their respective tributaries (Saha et al. 2021; Ghosh et al. 2022; Mitra and Das 2022; Mitra et al. 2022; Ghosh et al. 2023; Sarkar et al. 2023; Sutradhar and Mondal 2023). Significant impacts from substantial floods are observed in states like Uttar Pradesh, Bihar, West Bengal, Odisha, and Assam in India (Mishra and Sinha 2020; Das and Gupta 2021; Paul and Sarkar 2022). These states together account for 38.68% of India's total population and bear the brunt of around 70% of

the most catastrophic flood occurrences in the nation. West Bengal exhibits a significant proportion of around 42.5% of its total land area as flood susceptible, rendering it one of the foremost regions in the nation that experiences the adverse impacts of floods. The region with the greatest impact, as documented in the Annual Flood Report by the Irrigation & Waterways Department, Government of West Bengal, India (2022), measured around 37,760 sq. km as flood susceptible.

Considering these exceptional conditions, it is essential to evaluate and demarcate regions susceptible to flooding to mitigate future flood-related losses. This may be achieved by creating comprehensive flood hazard susceptibility maps. Hence, the proficient and precise modelling of flood susceptibility has the potential to mitigate the repercussions by analysing susceptibility maps. Susceptibility mapping is affected by various parameters, each leaving distinct marks. Consequently, pinpointing the optimal parameters demands a thorough process, involving an extensive literature review, insights from experts in the study area, and meticulous field observations. Numerous research studies (Mishra and Sinha 2020; Das and Gupta 2021; Ghosh et al. 2022; Mitra et al. 2022; Paul and Sarkar 2022) have identified key aspects related to flood susceptibility mapping procedures. In recent years, integrated Remote Sensing (RS) and Geographic Information System (GIS) approaches have been widely employed for accurate and efficient susceptibility, vulnerability, and risk assessment (Gharakhanlou and Pérez 2023). Significant efforts have been dedicated to evaluating flood susceptibility through the application of various decision-making methodologies, including Multi-Criteria Decision-Making (MCDM) methods (Amiri et al. 2024), Analytical Hierarchy Process (AHP) (Harshasimha and Bhatt 2023; Sarkar et al. 2023; Shah and Ai 2024), Shannon's Entropy (SE) (Bera et al. 2022), While Multi-Criteria Decision Analysis (MCDA) offers a comprehensible and direct approach to flood susceptibility assessment, its applicability to intricate events may yield less dependable outcomes (Bentivoglio et al. 2022). Additionally, owing to the subjective nature of criterion weights, determined through expert-driven knowledge-based statistical methods, these weights are susceptible to external influences.

Despite the reliability and efficiency of numerical techniques in modelling flood hazards, the expeditious and accurate simulation of floods remains a persistent challenge. In recent years, machine learning (ML) based data-driven models as they have shown to be effective and dependable in developing efficient and trustworthy solutions (Dou et al. 2020; Merghadi et al. 2020), have become more important in flood susceptibility investigations (Maharjan et al. 2024). Several studies have been carried out to assess flood susceptibility through the utilization of algorithms based on machine learning, such as Adaptive Boosting (AdaBoost) (Gharakhanlou and Pérez 2023; Dey et al. 2024), Artificial Neural Networks (Karunanayake et al. 2020; Khaniya et al. 2021), Support Vector Machine (SVM) (Tanim et al. 2022). This is because ML models analyse the historical flood data to understand the connection between flood occurrences and the elements that contribute to them (Chang et al. 2019). This approach eliminates the subjective process of determining the importance of each component in predicting flood susceptibility (Ghanim et al. 2023).

In this backdrop, our current research aimed to create six distinct machine learning (ML) models such as **Logistic Regression (LR)** (Saikh and Mondal 2023; Mansour et al. 2024), **Support Vector Machines (SVM)** (Gharakhanlou and Pérez 2023; Ren et al. 2024), **Random Forest (RF)** (Ghanim et al. 2023; Dey et al. 2024), **Extreme Gradient Boosting (XGBoost)** (Hitouri et al. 2024), **Deep neural network (DNN)** (Ghobadi and Ahmadipari 2024; Nguyen et al. 2024)**, and Stacking Ensemble** to evaluate flood susceptibility in Malda district, West Bengal. We used the Python programming language environment to train the models and extract relevant information from them. We used the stacking ensemble approach, a versatile technique for improving the prediction by combining predictions of base models such as LR, SVM, RF, XGBoost, and DNN since stacking can mitigate the weakness of individual models and produce more robust predictions, besides reducing overfitting and provide more interpretability *via* training a meta-learner on the prediction of base models. As a result, the goals of this study are to detect flood patches and generate flood inventory points for the entire Malda district using Sentinel-1 SAR data on the GEE platform. To compare the prediction accuracy of machine learning (ML) models (LR, SVM, RF, XGBoost), deep neural network (DNN) architecture, and Stacking Ensemble for the Malda district using remote sensing and geospatial techniques.

2. Dataset and study area

2.1. Study area and historical flood events in Malda district

The district of Malda is situated within the latitude range of 24°40′20″ N to 25°32′08″ N and the longitude range of 87°45′50″ E to 88°28′10″ E. Covering a geographical expanse of 3733.66 sq. km (1441.6 sq mi), Malda shares an international border with Bangladesh to the east, and is adjacent to Santhal Parganas of Jharkhand and Purnea of Bihar to the west (Figure 1). The region (Figure 1(c)) under consideration has a 165.5-km boundary that is shared with the neighbouring country of Bangladesh. The presence of a strategic position makes it a significant intersection and gateway to Siliguri from the southern region of Bengal. The river Ganga starts its ingress into the state of West Bengal near Manikchak of Malda. Due to its geographical characteristics as a low-lying basin, the area is susceptible to flooding. The Malda district mostly comprises low-lying alluvial plains that exhibit a gradual slope towards the southern direction. The district's North-Eastern region has many high areas. Certain regions of these elevated terrains have altitudes 40 m above the mean sea level. The heights are connected by deep water canals, resulting in a topographical configuration that resembles little hills. The Mahananda River, flowing from the northeast to the southeast, serves as the dividing line between the Eastern and Western regions of the district. Moreover, the river Kalindri serves as a dividing line, separating the Western area into distinct Northern and Southern sections. Each of the areas has its unique qualities. The Eastern region has a very elevated and undulating topography, while lower elevations and rich land characterize the Western region. The district may be classified into three discrete physiographic sub-regions, which are distinguished by their topography and soil characteristics. These sub-regions are referred to as Barind, Diara, and Tal.

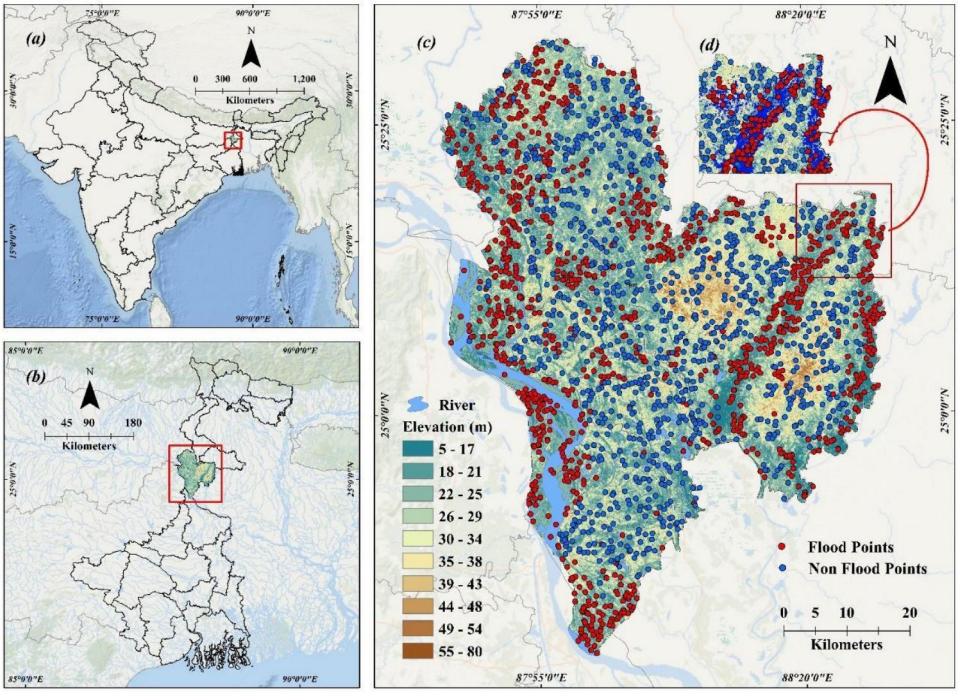

Figure 1. Location map of the study area (a) India, (b) West Bengal, (c) Malda district, and (d) spatial distribution of flood locations.

The Malda region is prone to devastating floods. Flooding has traditionally been a threat in the district during the rainy season, particularly in all the blocks and municipalities. In the following years, 1987, 1988, 1991, 1995, 1996, 1998, 1999, 2003, 2005, 2006, 2007, 2008, 2017, and 2019 all districts were hit by severe flooding catastrophes (DDMPM 2020–2021), mainly occurred during the month between July and September. Instead of being caused by rainfall in the district directly, these floods were the consequences of rainfall in the upper catchment regions as well, which caused the rivers to overflow their banks. Most district rivers and streams, including the Ganges, have their origins in the Himalayan mountains, making them vulnerable to the sudden influx of enormous water generated by the melting of snow or by the extreme rainfall in the mountains. The District Disaster Management Plan of Malda (DDMPM 2020–2021) estimated that during the 2017 flood, nearly 3,64,043 population fully affected, 236 villages were adversely affected, 12,871 houses were damaged, and 1069 hectares (ha) of agricultural land were damaged. Flooding lasted for almost 20 days, wreaking havoc on people's lives and property in the district.

2.2. Optical and radar data collection

This study compiled a geographical database of 18 flood conditioning parameters to create thematic layers within the GIS platform. The primary objective is to systematically investigate and define parameters related to flood susceptibility (FSP) using scientific methods. Various thematic data layers, sourced from reliable channels,

underwent processing within the ESRI ArcGIS 10.4.1 environment to achieve the desired outcomes. Utilizing the USGS ASTER GDEM, a digital elevation model with a spatial resolution of 30 m, served as the foundation for generating diverse thematic layers, including elevation, slope, TPI, TRI, TWI, relief amplitude, SPI, and STI. Extraction of the drainage network was performed using the ASTER GDEM dataset, and, on the ESRI ArcGIS 10.4.1 platform, thematic layers for drainage density and distance to drainage were subsequently developed.

Optical images from the USGS LANDSAT-8 OLI/TIRS, corrected for radiometric considerations, were utilized to generate three distinct land-use thematic layers: NDVI, mNDWI, and LULC. The thematic layer depicting clay content was developed using the global dataset of physical soil properties from Soil Grid. Annual rainfall and rainfall intensity thematic layers were produced using mean average annual rainfall and the modified Fournier Index (MFI) method, utilizing Information gathered by the Indian Meteorological Department (IMD) over 35 years, ranging from 1986 to 2020. The same dataset was utilized to compute district digital maps, which were then employed to create geomorphology and lithology thematic layers sourced from the BHUKOSH GSI (Geological Survey of India).

After completing the Synthetic Aperture Radar (SAR) change detection analysis to identify flood and non-flood locations, we proceeded to extract pixel values from the above-mentioned flood conditioning factors for these specific points. By overlaying the identified flood and non-flood locations onto spatial datasets corresponding to each conditioning factor, we recorded the relevant pixel values using Geographic Information System (GIS) tools and remote sensing techniques. This process enabled us to compile a comprehensive dataset reflecting the environmental and geographical conditions at both flooded and non-flooded sites, which is essential for further analysis to understand the relationship between these factors and flood occurrences, ultimately aiding in the development of predictive models and flood risk management strategies. Table 1 and Table 2 offers a comprehensive overview detailing the origins and explanations of the parameters employed in susceptibility zonation, and hyper-parameters details for applied algorithms.

2.3. Flood susceptibility parameters (FSP)

2.3.1. Elevation

Elevation is a significant criterion because low-lying flood plains and lower-elevated land are the most flood-prone zones (Meliho et al. 2022; Saikh and Mondal 2023). The district has an elevation range of 5 to 80 m. The district's western, central, and eastern parts belong to the lower elevated lands and flood plains where river flow increases downstream (Figure 2(a)). 30 m ASTER GDEM prepared the elevation map and classified by natural breaking as 5–20 m, 20–26 m, 26–32 m, 32–39 m, and 39–80 m.

2.3.2. Slope

Slope angle is a significant factor where the lower slope leads towards the flood (Mia et al. 2022; Mehravar et al. 2023). The slope angle of the district ranges from 0° to 35°. Flat (0°–2°), gentle (2°–4°), moderate (4°–6°), moderately steep (6°–9°), and

Table 1. Source and description of the parameters used in flood susceptibility zonation.

Parameters	Descriptions	Data source	Resolution/scale	GIS data type
Elevation, slope, relief amplitude, TPI, TRI, TWI, SPI, STI, drainage density and distance to river	Derived from ASTER DEM and prepared the thematic layer using ArcGIS 10.4.1	United States of Geological Survey (USGS) Retrieved from: (https://earthexplorer.usgs.gov)	30 m	Raster
NDVI, MNDWI and LULC	Landsat-8 OLI/TIRS, all the layers were prepared after mosaicing and atmospheric correction of the image	United States of Geological Survey (USGS) Retrieved from: (https://earthexplorer.usgs.gov)	30 m	Raster
Clay content	Physical soil properties Global soil grid data	Soil Grid Retrieved from: (https://soilgrids.org/)	250 m	Raster
Annual rainfall, rainfall intensity using MFI	Gridded rainfall for the period 1986–2020 has been used for calculation	Indian Meteorological Department (IMD) Retrieved from: (http://www.imdpune.gov.in)	0.25 × 0.25	NetCDF
Geomorphology, lithology	Digital geomorphological, lithological map of the district	Geological Survey of India (GSI) Retrieved from: (http://bhukosh.gsi.gov.in/)	–	Vector

steep (9°–35°) slopes (Figure 2(b)). Slopes have been classified where flat and gentle slopes are found in the western, central, and eastern sections of the district, which are the most flood susceptible zones. The slope (degree) of the study area has been classified into 0°–2°, 2°–4°, 4°–6°, 6°–9°, and 9°–35°.

2.3.3. Topographic Position Index (TPI)

TPI is the significant parameter of FS. TPI values mainly indicate the valley, flat land, and moderately elevated and highly elevated areas (Abedi et al. 2021). Mainly negative and zero classification indicates the valley and flat regions, which leads to higher FS.

$$TPI = Z_0 - \widetilde{Z}_i \tag{1}$$

$$\widetilde{Z}_i = \frac{1}{n_R} \sum_{i \in R} Z_i \tag{2}$$

The TPI map was employed to assess FS by applying Equations (1) and (2). Topographic Position Index (TPI) illustrates the variation in altitude of a specific location (Z_0) compared to the average altitude ($\widetilde{Z}$) within a certain distance (R). Additionally, the TPI was assessed using the 'land facet corridor designer tool' within the ArcGIS platform, as outlined by Jenness et al. (2013) and Mitra and Das (2022). TPI values of the studied region ranges from −4.72 to 5.29 (Figure 2(c)). TPI values have been classified into −4.72 to −1.50, −1.49 to 0.00, 0.00–0.01, 0.02–2.50, and 2.51–5.29.

2.3.4. Topographic Ruggedness Index (TRI)

TRI is calculated by measuring the difference in elevation between adjacent cells of the ASTER GDEM dataset, providing a numerical value that reflects the heterogeneity of the landscape (Fatah et al. 2022). High TRI values indicate a rugged terrain with

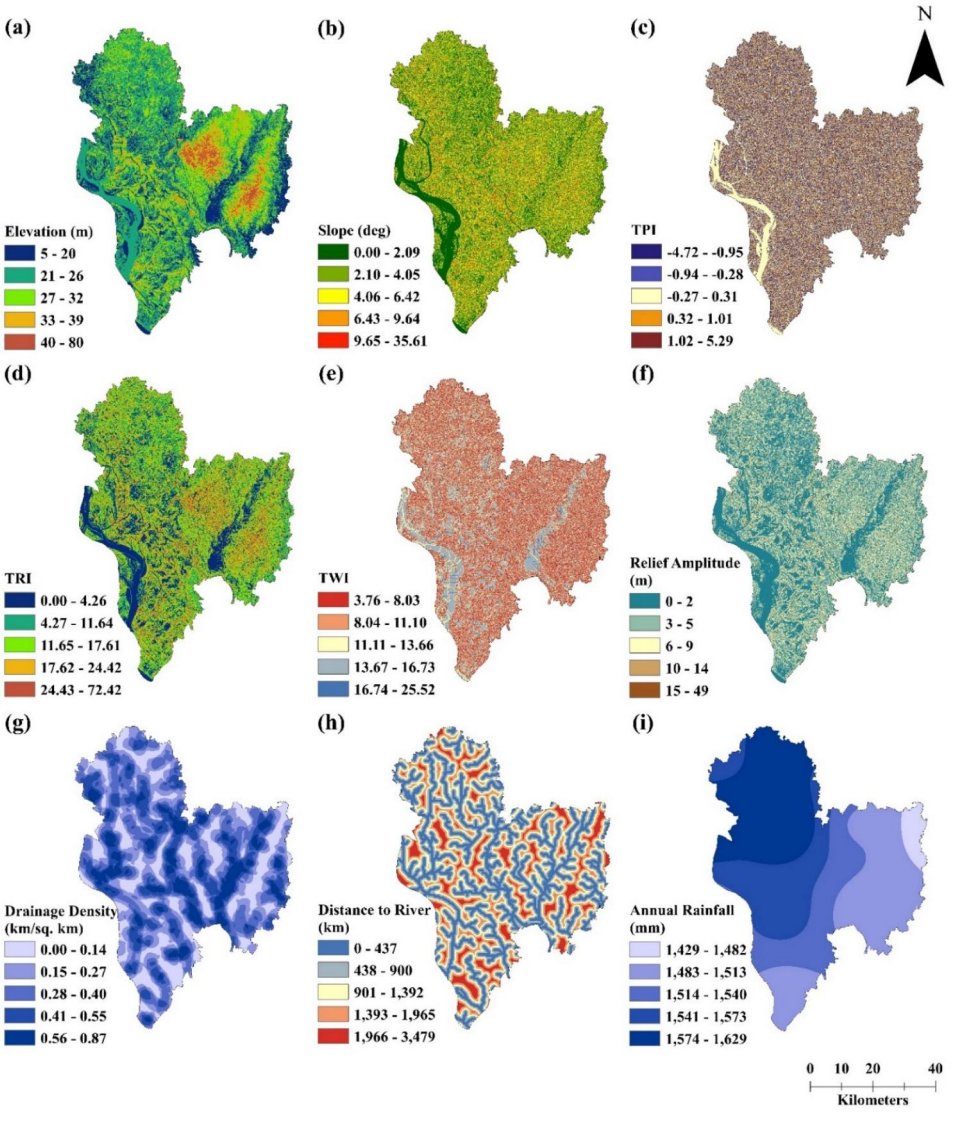

Figure 2. Flood conditioning parameters: (a) elevation, (b) slope, (c) TPI, (d) TRI, (e) TWI, (f) Relief amplitude, (g) drainage density, (h) distance to drainage, (i) rainfall, (j) MFI, (k) NDVI, (l) mNDWI, (m) SPI, (n) STI, (o) clay content, (p) LULC, (q) geomorphology, and (r) lithology.

steep slopes, while low TRI values correspond to flat or gently sloping areas. Steep terrains with high TRI values typically have rapid runoff due to limited water infiltration, potentially contributing to flash flooding downstream. Conversely, low TRI areas tend to have a slower runoff, which can lead to longer-duration floods but with generally lower depths and velocities (Equation (3)).

$$TRI = \sqrt{Abs(FSmax^2 - FSmin^2)} \tag{3}$$

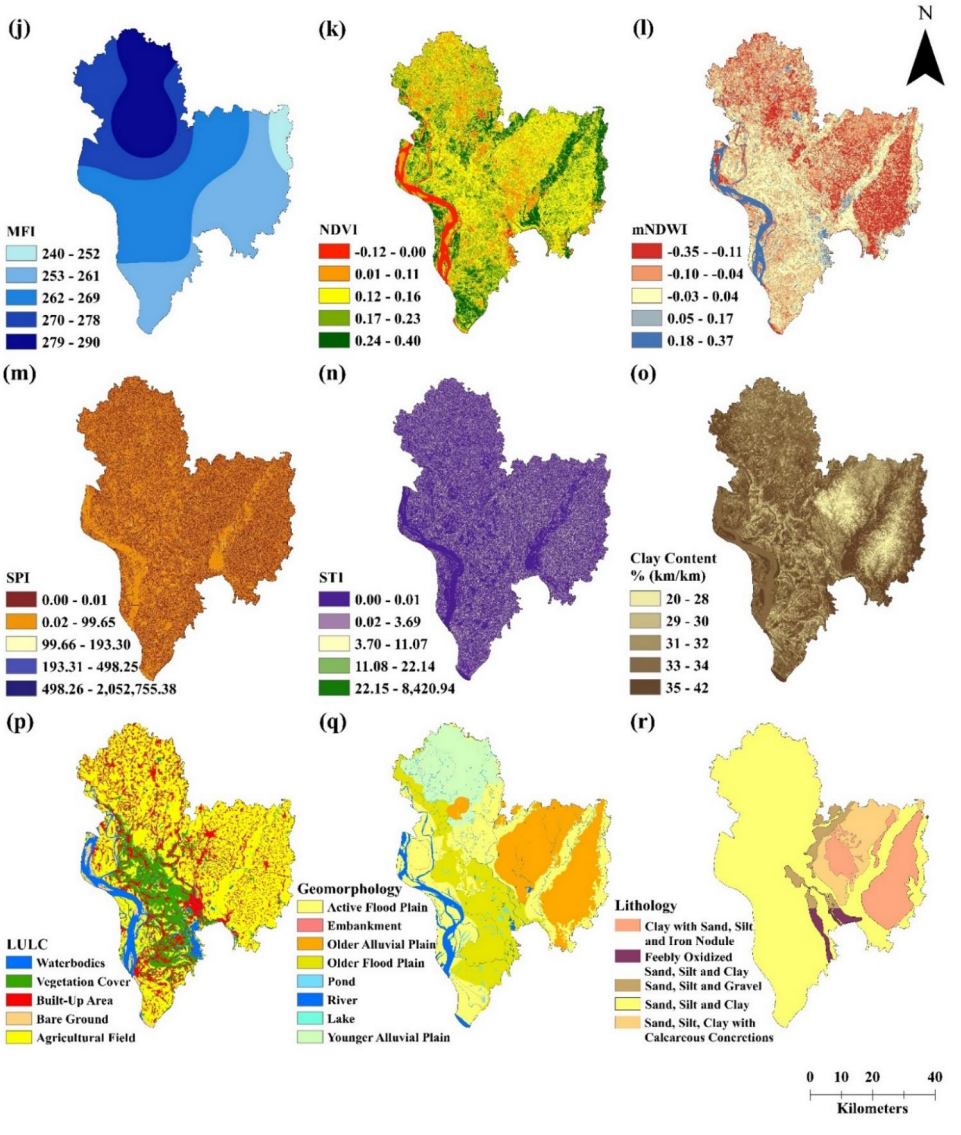

Figure 2. Continued.

Here, '*FSmax* and *FSmin* are the maximum and minimum elevation of the define region respectively'. TRI values of the studied region range from 0.00 to 72.42 (Figure 2(d)). TRI values have been classified into 0.00–4.26, 4.27–11.64, 11.65–17.61, 17.62–24.42, and 24.43–72.42.

2.3.5. Topographic Wetness Index (TWI)

TWI is a crucial landscape parameter that indicates soil erosion, soil moisture condition, the volume of flow accumulation, and runoff generation (Mousavi et al. 2022). The TWI layer is created by applying the formula outlined in Equation (4)

(Beven and Kirkby 1979; Abedi et al. 2021; Mitra et al. 2022) to the ASTER GDEM dataset,

$$TWI = Ln\left(\frac{a}{\tan\ \beta}\right) \tag{4}$$

Where, 'the variables α and $tan\ \beta$ pertain to the catchment area and slope angle, respectively, of the region under investigation, respectively. Furthermore, here $= \frac{A}{L}$, where the variable A represents the entirety of the basin's area and L denotes the length of the contour being analysed' (Beven and Kirkby 1979; Fatah et al. 2022). TWI values of the studied area varied from 3.76 to 25.52 (Figure 2(e)). TWI has been classified into 3.76–8.03, 8.04–11.10, 11.11–13.66, 13.67–16.73, and 16.74–25.52.

2.3.6. Relief amplitude

Relief amplitude serves as an indicator of potential energy in a landscape, with higher values suggesting a greater force for water to move downslope. In areas with high relief amplitude, precipitation can result in rapid surface runoff, increasing the risk of flash flooding, particularly in the absence of sufficient vegetation or soil absorbency. Low relief areas, characterized by smaller variations in elevation, are prone to slower water movement and can become zones of waterlogging and prolonged inundation (Equation (5)).

$$Relief\ \ Amplitude = (FSmax - FSmin) \tag{5}$$

Where '*FSmax* and *FSmin* are the maximum and minimum elevation of that region respectively'. Relief amplitude values of the studied region range from 0 to 49 m (Figure 2(f)). Relief amplitude values have been classified into 0–2 m, 3–5 m, 6–9 m, 10–14 m, and 15–49 m.

2.3.7. Drainage density

Higher drainage density leads to a greater amount of surface runoff. Regions with higher stream networks have a great chance to be flooding frequently (Meliho et al. 2022; Wang et al. 2023) (Equation (6)).

$$Dd = \sum_{i=1}^{i=n} \frac{D_i}{A}(km^{-1}) \tag{6}$$

Where,'$\sum D_i$ is the stream's overall length within the grid (km) is represented by, while A denotes the grid area (km^2)'. The drainage density of the area ranges from 0.00 to 0.87 km./sq. km. (Figure 2(g)). Drainage density values are 0.00–0.14 km./sq. km, 0.15–0.27 km./sq. km, 0.28–0.40 km./sq. km, 0.41–0.55 km./sq. km, and 0.56–0.87 km./sq. km.

2.3.8. Distance to river

Areas near rivers are more susceptible to flooding than those near the river network. Inland storage of water like large ponds, lakes, reservoirs, and dams is also likely to

flood locally (Meliho et al. 2022; Mia et al. 2022). Distance to the river ranges from 0.00 to 3478.97 km (Figure 2(h)). Distance to river values has been classified into: 0.00–436.58 km, 436.59–900.44 km, 900.45–1391.59 km, 1391.60–1964.59 km, and 1964.60–3478.97 km.

2.3.9. Annual rainfall

Annual Rainfall is a critical parameter in flood susceptibility as it provides a baseline for the amount of water introduced to a catchment area over a year (Mia et al. 2022; Dutta et al. 2023). High annual rainfall levels can saturate the ground, making it less capable of absorbing additional rainfall, thus increasing flood potential (Meliho et al. 2022). Monitoring annual rainfall can help in predicting potential flood events, especially when combined with other hydrological factors. In the district, Annual Rainfall values range from 1429.11 to 1628.97 mm (Figure 2(i)). Annual Rainfall values have been classified into 1429.11–1481.62 mm, 1481.63–1512.97 mm, 1512.98–1540.41 mm, 1540.42–1573.32 mm, and 1573.33–1628.97 mm.

2.3.10. Modified Fournier Index (MFI):

Rainfall intensity map with the MFI method has been calculated by 35 years (1981–2020) IMD data of the total district. After the creation of the spatiotemporal map, the higher MFI values indicate higher FS (Eslaminezhad et al. 2022; Mitra and Das 2022) (Equation (7)).

$$MFI = \sum_{i=1}^{12} \frac{P_i^2}{P'} \tag{7}$$

Where, 'P_i represents the average amount of precipitation in a month, while P' represents the average amount of precipitation in a year'. In the district, MFI values range from 239.54 to 290.37 mm/year (Figure 2(j)). MFI values have been classified into 239.54–252.50 mm/year, 252.51–260.67 mm/year, 260.68–269.24 mm/year, 269.25–278.01 mm/year, and 278.02–290.37 mm/year.

2.3.11. Normalized Difference Vegetation Index (NDVI):

NDVI values typically range from −1.0 to +1.0. Areas with NDVI values (0.1 or less) indicate non-vegetated regions, such as low-lying floodplains, bare land, and sand. Moderate NDVI values (0.2–0.3) indicate shrub, grassland, and rangeland, while higher values (0.6–0.8) suggest moderate to dense forest (Zhao et al. 2021; Mia et al. 2022). Less vegetation cover regions are more susceptible to flood than those with higher vegetation cover (Equation (8)).

$$NDVI = \frac{NIR - Red}{NIR + Red'} \tag{8}$$

Where, '*Red* signifies red band and *NIR* signifies near-infrared band'. NDVI values of the district range from −0.12 to 0.40 (Figure 2(k)). NDVI values have been classified into: −0.12to 0.00, 0.01–0.11, 0.12–0.16, 0.17–0.23, and 0.24–0.40.

2.3.12. Modified Normalized Difference Water Index (mNDWI):

Normally, mNDWI values lie between −1.0 and +1.0. According to different values, vegetation has smaller values that differentiate vegetation from waterbodies, built-up features have values between 0.00 and 0.20, and waterbodies value is greater than 0.5 (Equation (9)).

$$mNDWI = \frac{Green - MIR}{Green + MIR'} \tag{9}$$

Where, '*Green* signifies green band and *MIR* signifies middle infra-red band'. mNDWI values of the study area range from −0.35 to 0.37 (Figure 2(l)). mNDVI values are: −0.35 to −0.11, −0.10 to −0.04, −0.03 to 0.04, 0.05–0.17, and 0.18–0.37.

2.3.13. Stream Power Index (SPI)

SPI is a critical factor that indicates the stream erosive power level and sediment transport to a specific area with respect to the catchment basin. High stream power leads to channel transformation and creates flooding conditions (Abedi et al. 2021; Penki et al. 2022) (Equation (10)).

$$SPI = Ai \times \tan\beta \tag{10}$$

Where, 'Ai Represents the specific area and $\tan\beta$ represents the gradient'. SPI values of the district range from 0.00 to 20,52,755.36 (Figure 2(m)). SPI values have been classified into 0.00–0.01, 0.02–99.65, 99.66–193.30, 193.31–498.25, and 498.26–20,52,755.36.

2.3.14. Sediment Transport Index (STI)

STI is an important parameter of FS and shows the general runoff figure of a catchment basin. Higher runoff indicates a higher amount of sediment transportation and lower FS (Fatah et al. 2022; Mia et al. 2022) (Equation (11)).

$$STI = \left[\frac{\left(\frac{F_a}{\delta x}\right)^2}{\left(\frac{Sig\ \ n(S_a)}{\delta y}\right)^2} \right] \tag{11}$$

Where, 'F_a represents the flow accumulation and S_a represents the slope raster, derived from ASTER GDEM, and δx and δy represent the constant'. STI values of the studied region varied from 0.00 to 8420.94 (Figure 2(n)). STI values have been classified into 0.00–0.01, 0.02–3.69, 3.70–11.07, 11.08–22.14, and 22.15–8420.94.

2.3.15. Clay content

Clay content in soils influences flood susceptibility due to its unique particle size and cohesive properties. Compared to sand or silt, clay particles are much smaller, leading to tighter soil pore spaces and decreased permeability. As a result, clay-rich soils tend to retain water on the surface longer than coarser soils, exacerbating surface runoff

during heavy rain. In areas with high clay content, water infiltration rates are lower, causing the water table to rise more quickly during rain events. Clay content values of the studied region varied from 20 to 42% (km/km) (Figure 2(o)). Clay content values have been classified into 20–28% (km/km), 29–30% (km/km), 31–32% (km/km), 33–34% (km/km), and 35–42% (km/km).

2.3.16. Land Use and Land Cover (LULC)

Land Use and Land Cover (LULC) changes are a key factor in flood events (Youssef et al. 2023). These changes influence how water moves across the landscape (Saha et al. 2023). LULC of the district signifies the flood condition with diverse landscapes (Abedi et al. 2021; Mia et al. 2022). LULC of the Malda district has been classified into five different distinctive classes, i.e. waterbody (7.78%), vegetation cover (13.68%), agricultural area (59.12%), built-up area (17.54%) and bare ground (1.87%) (Figure 2(p)). The above classification shows that the agricultural area mostly occupies the Malda district.

2.3.17. Geomorphology

Geomorphology plays a critical role in understanding and predicting flood behaviour, as it encompasses the study of landforms and the processes that have shaped them (Mitra and Das 2022). When considering flood events, the geomorphological features and processes of a region can influence the magnitude, frequency, duration, and extent of flooding (Saikh and Mondal 2023). The geomorphological map of the entire district has been classified into eight distinctive zones based on BHUKOSH GSI (Figure 2(q)). Classified zones are active flood plain (31.70%), embankment (0.21%), older alluvial plain (25.58%), older flood plain (21.13%), younger alluvial plain (14.66%), pond (1.02%), river (5.68%) and lake (0.03%).

2.3.18. Lithology

In this research, lithology is regarded as an additional component influencing flood conditions (Youssef et al. 2023). Measuring the hydrological processes, stagnation, and percolation of water, rock permeability, and lithology is a crucial factor in flooding (Abedi et al. 2021; Fatah et al. 2022). Even drainage network features of each given location are influenced by this phenomenon. The lithological map of the entire district has been classified into five distinctive zones based on BHUKOSH GSI (Figure 2(r)). Classified zones are clay with sand, silt, and iron nodule (13.44%), feebly oxidized sand, silt and clay (1.63%), sand, silt, and gravel (3.67%), sand, silt, and clay (72.24%) and sand, silt, clay with calcareous concretions (9.02%).

2.4. Multicollinearity test

The Multicollinearity testing method is used to assess the relationship between chosen flood affecting elements. Prior to doing regression analysis in susceptibility models, multicollinearity tests are often performed using tolerances, variance inflation factors (VIF), conditional index, and Pearson's correlation coefficients (Debnath et al. 2023). This

research used tolerances and VIF (Variance Inflation Factor) to evaluate the level of multicollinearity among the 18 chosen flood-affecting components (Ghosh et al. 2023).

Tolerance and VIF (Variance Inflation Factor) are closely related concepts used to assess multicollinearity in multiple regression models. Tolerance is the reciprocal of VIF, and both can be calculated as follows. Tolerance measures the proportion of variance in a predictor variable that can be explained by the other predictor variables in the model. It is calculated for each predictor variable as follows (Equation (12)):

Tolerance of the ith predictor variable

$$(T_i) = 1 - R_i^2 \tag{12}$$

Where, 'T_i is the tolerance for the ith predictor variable. R_i^2 is the coefficient of determination when the ith variable is regressed against all the other predictor variables'. High tolerance values (close to 1) indicate that a predictor variable is not affected by multicollinearity and has little variance explained by other predictors. Conversely, low tolerance values suggest that a variable is highly influenced by multicollinearity.

VIF quantifies how much the variance of the estimated regression coefficients is increased due to multicollinearity. It is the inverse of tolerance and is calculated as follows (Equation (13)):

VIF of the ith predictor variable

$$(VIF_i) = \frac{1}{T_i} \tag{13}$$

Where, 'VIF_i is the VIF for the ith predictor variable, T_i is the tolerance for the ith predictor variable'. A high VIF (typically greater than 10) suggests strong multicollinearity, while a low VIF indicates that a predictor variable is not highly correlated with others in the model.

3. Methodology

3.1. Flood inventory map

Microwave remote sensing techniques, such as Synthetic Aperture Radar (SAR), offer additional advantages over optical remote sensing by enabling the observation and monitoring of the Earth's surface irrespective of prevailing weather conditions (Ramayanti et al. 2022). The backscatter variability in the VV channel is generally more responsive to water surface roughness compared to the VH channel (Zhao et al. 2022). Consequently, the anticipated backscatter reduction due to flooding may not always occur, as turbulent floodwaters can produce higher backscatter in the VV channel than various natural and artificial bare surfaces. Conversely, the VH channel tends to show a more consistent decrease in backscatter across different land cover types during flooding, as it produces very low backscatter returns over water, is less sensitive to surface roughness, and exhibits minimal double-bounce scattering (Pelich et al. 2022). The VH polarization is particularly efficient in identifying waterlogged

areas, particularly in small regions of concern (Parida et al. 2022). In both VH and VV polarizations, water bodies appear black due to the smoothness of the surface of the water bodies, which causes radar signals to be reflected away from the sensor, resulting in low backscatter (Saini et al. 2020). This low return signal creates dark patches in the image, effectively highlighting water bodies against the more reflecting surrounding environment. However, for our study site, VH polarization was chosen due to its more consistent backscatter reduction across different land cover types during flooding.

The flood inventory map, which provides a comprehensive record of historical and contemporary flood incidents, plays a critical role in identifying flood-prone areas (Nohani et al. 2019; Zhao et al. 2021; Tavus et al. 2022). In our study, we derived a flood inventory map through a systematic approach. We conducted an in-depth analysis of the historical flood events that occurred in the years 2017, 2019, and 2021. These flood events varied in magnitude and spatial extent. We extracted the flood patches by performing a comprehensive change detection analysis on the VH polarization band of the SAR GRD (Ground Range Detected) data hosted on the Google Earth Engine cloud platform, comparing pre- and post-flood conditions within our specified region of interest.

The process begins with applying the refined Lee algorithm to both the pre-and post-flood VH polarized images to reduce the salt and pepper noise caused by radar scattering (Yommy et al. 2015). Then, we compute the ratio by dividing the dB values of the post-flood image by those of the pre-flood image, which allows us to highlight significant changes in backscatter, indicative of flooding. A threshold of 1.25 is applied to this ratio image based on empirical analysis to create an initial estimate of flooded pixels, indicating a significant change that is indicative of flooding. However, this initial estimate can contain false positives due to factors like permanent water bodies, steep slopes (where water is unlikely to accumulate), and isolated pixels (which are likely errors). The Global Surface Water (GSW) dataset is used to identify areas with permanent or semi-permanent water, which are masked out as they do not represent fresh instances of flooding. The SRTM digital elevation model (DEM) is used to calculate slope, and areas with slopes greater than 5 degrees are masked out, as water is less likely to accumulate on steep slopes. In GEE, a connected-component analysis is performed on the remaining flooded areas. Small clusters of pixels below a threshold of 20 are considered isolated and removed, as they are likely errors or not representative of true flood extents. This threshold is based on empirical analysis and is tailored to the problem and the study area's characteristics. Subsequently, the refined flood extents from each of the three flood years under consideration were integrated into a composite binary raster, providing a comprehensive overview of areas inundated during any of the analysed flood events. Finally, a morphological closing operation is applied to fill in small holes and gaps within the flood patches, effectively smoothing the boundaries. The result was a more continuous representation of the flood extent. This refined flood layer and multi-year analysis enabled a more robust assessment of flood vulnerability and a better understanding of the spatial and temporal dynamics of flooding in the study region.

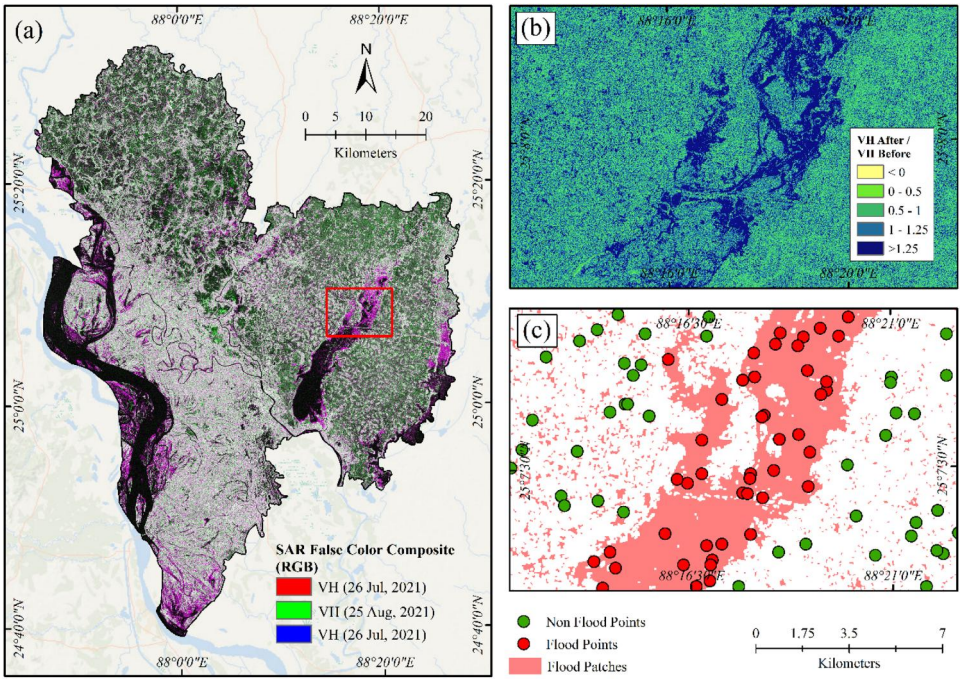

Figure 3. Flood inventory map (a) false colour composite of SAR VH bands, with pre-flood VH in red and blue channels and post-flood VH in the green channel, highlighting changes in the backscatter response due to flooding, (b) ratio image of pre- and post-flood VH polarization bands, identifying areas affected by flooding, and (c) combined flood patches from three different flood years, overlaid with randomly generated flood and non-flood points.

We generated random points within the study area and adopted an approach where points lying within the extracted flood patches were assigned a value of 1, indicating them as flood points. Subsequently, in terms of non-flood points, we created a buffer zone with a radius of 500 m around the flood patches and generated random points again. But this time, the points that fell outside the buffer zone and did not intersect with the flood patches were selected and assigned a value of 0, indicating them as non-flood points. The entire process of generating flood inventory points was efficiently performed on the GEE cloud platform. We acquired a total of 2260 data points to generate a robust and well-balanced dataset for our flood susceptibility zonation (Figure 3(b)). Subsequently, to facilitate model training and evaluation, we partitioned this dataset using a popular approach known as the train-test split. We allocated 70% of the data (1582 points) for training our flood susceptibility models, enabling them to learn patterns and relationships from the majority of the dataset. The remaining 30% of the data (678 points) was left for testing and evaluating the models' performance (Figure 3(c)). This methodical split of data ensured that we could examine the model's ability to generalize and generate accurate predictions on fresh, unseen data, a vital step in flood susceptibility mapping (Antzoulatos et al. 2022; Tavus et al. 2022; Risling et al. 2024; Wahba et al. 2024).

3.2. Data pre-processing

Normalization of the explanatory variables is an important data pre-processing step to avoid any kind of bias in the classification. In normalization, the data have typically been rescaled between 0 and 1 so that no single feature significantly dominates the classification results. This makes the interpretation of the normalized values straightforward. We chose Min-Max scaler because it preserves the true relationship among the original data points, thereby maintaining the proportionate distances between different values. Additionally, normalization can also bring robustness to the models and make them less prone to overfitting. For example, elevation in the study area typically ranges from 6 to 80 m, whereas indices like NDVI and NDWI range between −1 and +1. Some machine learning algorithms compute the weight of each explanatory variable based on its sheer magnitude. Normalization is performed as per Equation (14):

$$X_{normalized} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{14}$$

However, not all machine learning models require feature scaling. For instance, algorithms like logistic regression, SVM, MLP, DNN, and kNN perform better with feature scaling, whereas tree-based models such as decision trees, random forests, and gradient boosting remain unaffected.

Handling categorical variables is another essential data pre-processing step, as most machine learning algorithms cannot handle categorical features unless we convert them into numerical features. There are several techniques to handle categorical features, such as one-hot encoding, label encoding, ordinal encoding, frequency encoding, etc. One-hot encoding is a commonly used method to convert nominal features into binary vectors. It creates a new binary column for each category within a categorical feature and assigns a value of 1 or 0 to the columns, where 1 represents the presence and 0 represents the absence of a specific category.

One-hot encoding has been efficiently implemented in the Python environment. We applied one-hot encoding to the categorical features such as geomorphology, lithology, and LULC. These features were initially represented as text labels and needed to be converted into a numerical input format suitable for machine learning algorithms. However, one-hot encoding, while effective, can introduce the 'curse of dimensionality' in the dataset, which in turn can lead to increased memory and computational requirements.The methodological flow diagram shows (Figure 4) the complete blueprint of the entire research.

3.3. Feature selection

Feature selection is a critical step, as one-hot encoding significantly increases the dimensionality of the dataset. Feature selection removes irrelevant features from the dataset while keeping the most relevant ones, which helps to increase the overall performance of the models. Mutual information (entropy) gain, which measures the

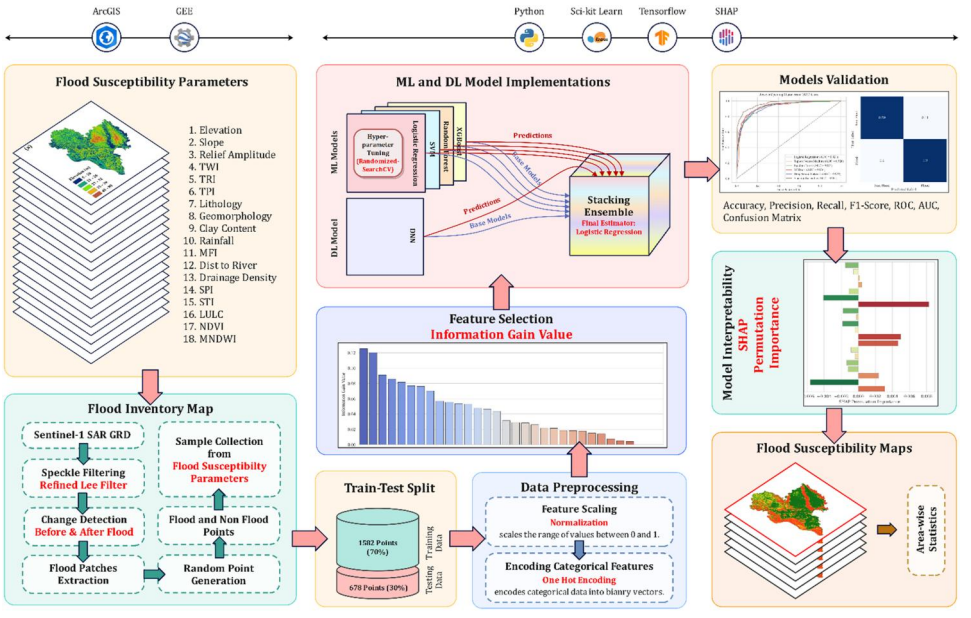

Figure 4. Methodological flow chart.

dependence and information shared between the features and the target variable, is a useful method for feature selection.

Information gain is closely related to entropy, which quantifies the amount of information contained in a random variable or feature. A skewed distribution, where one event dominates, generally has low entropy as there is less uncertainty. In contrast, a distribution where events have equal probability has higher entropy, indicating greater uncertainty and disorder in the dataset. In a binary classification scenario, entropy can be calculated as per Equation (15):

$$H(S) = -p_0(p_0) - p_1(p_1) \tag{15}$$

Where, $H(S)$ is the entropy of the dataset, p_0 and p_1 are the proportions of instances in class 0 and class 1 respectively. The entropy ranges between 0 and 1, where 0 indicates a complete pure dataset and 1 indicates a complete impure dataset.

To quantify the gain of each feature in the dataset, the IG value is computed as per Equation (16):

$$IG(S,A) = H(S) - \sum_{v \in Values(A)} \frac{|s_v|}{|s|} \cdot H(S_v) \tag{16}$$

Where, $IG(S,A)$ is the information gain by splitting the dataset S on feature A, $H(S)$ is the entropy of the dataset S, $Values(A)$ is the set of all possible values of feature A, s_v is the subset of data in which feature A takes the values v, $|s|$ is the total number of instances in the dataset S, $|s_v|$ is the total number of instances in the subset s_v, and $H(S_v)$ is the entropy of the subset S_v.

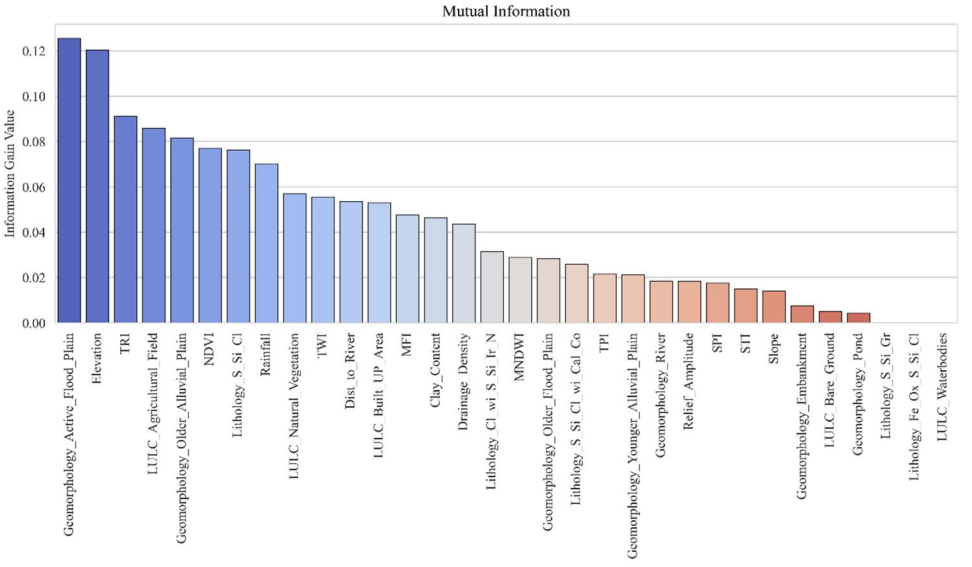

Figure 5. Information gain value.

We utilized Scikit-learn's 'mutual_info_classif' class to quantify the information gain value for each feature in the dataset. Among the various explanatory variables considered for flood susceptibility, it was found that active floodplain and elevation exhibited the highest information gain values following other variables such as TRI, the agricultural field, older alluvial plain, etc. (Figure 5).

Scikit-learn's 'SelectKBest' class has been incorporated to effectively reduce the initial one-hot encoded set of 32 features to a more manageable 20 features. This process, in turn, reduces the dimensionality of the data while preserving the most valuable features for further analysis.

3.4. Supervised learning

We utilized Python, a powerful and popular programming language for data analysis and machine learning tasks, to create flood susceptibility maps through a supervised learning procedure. More specifically, we used scikit-learn, a Python-based machine learning library (https://scikit-learn.org/stable/) that is easy to use and effective, to create our flood susceptibility models. This library ensures great performance and user-friendliness by making the training and evaluation of our models easier. Our analysis encompassed the utilization of four machine learning algorithms: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), along with a Deep Neural Network (DNN). These algorithms were selected for their outstanding classification capabilities and predictive accuracy. LR offers a simple and interpretable baseline for understanding relationships between flood conditioning factors and flood susceptibility. SVM excels in handling high-dimensional data and finding optimal class separation, enhancing the model's precision in identifying susceptible areas. RF is effective in managing non-linear relationships and interactions between variables, making it robust for predicting flood-prone regions. XGBoost is known for its high performance

and efficiency in capturing intricate patterns within the data, improving predictive accuracy. DNNs provide the flexibility to learn complex, non-linear representations, which is crucial for accurately mapping flood susceptibility. The stacking ensemble model integrates these diverse approaches to leverage their unique strengths, resulting in improved overall performance and robustness. The combination of these models provides a comprehensive approach to flood prediction, leveraging both linear and non-linear relationships, simple and complex models, and individual and ensemble learning strategies. This diverse set of models ensures robustness and reliability in our flood prediction analysis.

3.4.1. Logistic Regression (LR)

Logistic regression (LR) is a statistical technique primarily known for binary classification tasks, where the objective is to predict whether an instance belongs to one of two groups. It assumes that independent variables and the binary outcome can be represented by the logistic (sigmoid) function (Ghosh et al. 2022; Saikh and Mondal 2023). This function maps an input feature's linear combination onto a range of probabilities between 0 and 1, with the anticipated probability indicating the possibility of an instance being part of the positive class. The mathematical expression of the logistic regression model is as follows (Equation (17)):

$$P(Y=1) = \frac{1}{1+e^{-(\beta_0+\beta_1X_1+\beta_2X_2+\ldots+B_nX_n)}} \tag{17}$$

Here, $P(Y=1)$ is the probability of being positive class, β_0 is the intercept, β_i are the coefficients and X_i are the input features.

3.4.2. Support Vector Machine (SVM)

Support vector machine (SVM) algorithm is a widely used machine learning technique that efficiently solves classification and regression problems. Its core concept is to establish an optimal hyperplane that maximizes the distance between classes in the feature space (Dou et al. 2019; Saravanan and Abijith 2022). The combination of SVM and a radial basis function (RBF) kernel is a potent solution that excels at handling nonlinear decision boundaries in intricate datasets. The RBF kernel transforms the initial feature space into a higher-dimensional space, enabling SVM to effectively capture complex relationships. The mathematical expression for the decision function of SVM with an RBF kernel is as follows (Equation (18)):

$$f(X) = \sum_{i=1}^{n} a_iK(X,X_i) + \beta \tag{18}$$

Here, $f(X)$ stands for the decision function, while a_i are the Lagrange multipliers and X_i represents the support vectors. $K(X,X_i)$ is the RBF kernel, and β is the bias term. The RBF kernel can be expressed as (Equation (19)):

$$K\left(X,X^{'}\right) = expexp\left(-\frac{\|X-X^{'}\|^2}{2\sigma^2}\right) \tag{19}$$

Where, $\|X-X^{'}\|^2$ represents the squared Euclidean distance between input feature vectors X and $X^{'}$, and σ is the kernel width parameter.

3.4.3. Random Forest (RF)

Random forest is a widely used tree-based ensemble learning method that is capable of handling both classification and regression tasks with great proficiency. This algorithm builds multiple decision trees during training and then computes the mode of the classes in the case of classification (Li et al. 2022; Ghanim et al. 2023; Gharakhanlou and Pérez 2023). It introduces randomness by selecting a random subset of features for each tree and bootstrapping the data. The decision function of the random forest model can be expressed mathematically as an aggregation over the decision trees (Equation (20)):

$$F(X) = \left\{f_1(X) + f_2(X) + \cdots + f_N(X)\right\} \tag{20}$$

Here, the prediction for a feature vector is denoted by $F(X)$, *mode* indicates the class that appears most frequently among the predictions made by each tree, and N denotes the total number of trees in the random forest ensemble. Each $f_i(X)$ refers to the prediction of the i-th decision tree, which is trained on a random subset of features and a bootstrapped sample of the dataset.

3.4.4. Extreme Gradient Boosting (XGBoost)

XGBoost stands as a state-of-the-art algorithm famous for its exceptional performance in classification and regression problems. It is a gradient-boosting framework that sequentially generates an ensemble of weak learners, often decision trees, and steadily refines the prediction accuracy (Youssef, Pourghasemi, et al. 2022). The formulation of XGBoost combines the capabilities of gradient boosting and regularization techniques to generate robust and accurate models (Equation (21)).

$$F(X) = \sum_{m=1}^{M} f_m(X) \tag{21}$$

Where, $F(X)$ is the final prediction, M is the total number of weak learners, and $f_m(X)$ is the prediction of the m-th tree. Each tree is trained to optimize the loss of its predecessors, thereby adapting, and enhancing the model's overall accuracy.

XGBoost's effectiveness can be attributed to its incorporation of regularization elements within the objective function, which helps strike a balance between predictive performance and model complexity. The objective function is formulated as (Equation (22)):

$$Obj = \sum_{i=1}^{n} L\left(y_i, \hat{y}_i\right) + \sum_{m=1}^{M} \Omega(f_m) \tag{22}$$

Here, $L\left(y_i, \hat{y}_i\right)$ signifies the loss function, evaluating the difference between the true label y_i and the predicted label $\hat{y}_i$, while $\Omega(f_m)$ represents the regularization term for the m-th tree.

3.4.5. Deep Neural Network (DNN)

A Deep Neural Network is a sophisticated architecture of artificial neural networks characterized by its multi-layered structures which refers to the presence of numerous hidden layers between the input and output layers. It uses techniques like forward and backward propagation for parameter learning and activation functions for introducing non-linearities. Although there are difficulties in computation and a requirement for a large amount of labelled data, DNNs have played a crucial role in advancing the field of artificial intelligence. The model architecture was developed using the TensorFlow library, having a setup suited for robust performance. The initial layer has 128 neurons with the rectified linear unit (ReLU) activation function, adding L2 regularization (0.01) to mitigate overfitting. A Dropout layer with a 0.5 dropout rate follows, boosting generalization by randomly deactivating neurons during training. Batch Normalization stabilizes the activations before continuing to a second Dense layer with 64 neurons, ReLU activation, and additional L2 regularization (0.01). A Dropout layer (0.3) and Batch Normalization further improve the network. The final layer, a Dense layer with one neuron and a sigmoid activation function is suited for binary classification problems. This architecture, defined by varied activation functions, regularization techniques, and dropout layers, amalgamates to build a resilient DNN model capable of sophisticated classification and robust generalization.

3.4.6. Stacking Ensemble

Hybrid modelling, under the framework of a stacking ensemble, involves using different base models to construct a more robust and precise prediction model. Stacking, or stacked generalization, is a meta-learning method in which many distinct base models are trained to predict the target variable (Zhu et al. 2024). The predictions made by these models are then used as inputs for a higher-level model, commonly called a meta-model. The training data for the ensemble model is created by including the predictions provided by the preceding models, specifically Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Deep Neural Network (DNN). The stacking ensemble comprises four core classifiers: LR, SVM, RF, and XGBoost, and a Logistic Regression model as the final estimator, combining the accumulated knowledge from the base models to improve forecast accuracy. To thoroughly evaluate the effectiveness of the model, a 10-fold cross-validation approach is employed, which guarantees a full assessment over several subsets of the dataset. This strategy seeks to use the synergistic advantages of multiple models, mitigating individual limitations and enhancing overall prediction efficacy.

We trained all our supervised machine learning models using the comprehensive pixel samples dataset to classify flood and non-flood locations (Ghobadi and Ahmadipari 2024). Once the models were trained, we applied them to the entire image containing the flood causative factors (with each factor represented as a separate band). We set the model output in the form of class probabilities, specifically the probability of each pixel belonging to the positive class (i.e. flood-prone areas) (Diaconu et al. 2024). We interpreted these probabilities as flood susceptibility,

creating flood susceptibility maps that highlight areas with higher probabilities of flooding (Bhatta and Adhikari 2024). In short, our methodology involves delineating flood patches from SAR images, extracting data from flood conditioning factors for flood and non-flood locations, training machine learning models on these data, and generating flood susceptibility maps based on the class probabilities output by the models (Amitrano et al. 2024; Ghosh et al. 2024). This approach allows us to effectively translate SAR-derived flood information into a comprehensive assessment of flood susceptibility.

3.5. Hyperparameter optimization

Hyperparameter optimization is a crucial stage in the development of machine learning models, aimed at optimizing the configuration of algorithmic parameters to enhance prediction performance. In this study, we have adopted Randomized Search Cross-Validation (RandomizedSearchCV) to iteratively search for the optimal hyperparameters. It is a robust and fast optimization technique that involves randomly sampling from a defined range of hyperparameters and evaluating model performance using cross-validation. This approach allows us to efficiently explore a broad hyperparameter space without the exhaustive computational cost associated with Grid Search. We have conducted a randomized search with a substantial 1000 iterations and five-fold cross-validation. This five-fold cross-validation strategy produces a robust estimate of model performance across various hyperparameter configurations.

For the LR model, the choice of the SAGA solver, L1 penalty, a maximum of 200 iterations, and a balanced class weight indicates a balanced consideration of computational performance, regularization, and handling of imbalanced class distributions. The regularization strength (C) of 100 suggests a relatively high regularization, favouring a more parsimonious model. In the refinement of the SVM, the specified hyperparameters include the choice of the radial basis function (RBF) kernel, suggesting a capacity to capture complicated, non-linear relationships within the data. The 'gamma' parameter determines the influence of individual training samples, with higher values resulting in a more localized influence. Additionally, the regularization parameter 'C' is calibrated to a value of 10, regulating the trade-off between creating a smooth decision border and accurately classifying training data.

When it comes to tree-based algorithms, the RF model is parameterized with great consideration: it comprises 50 trees (n_estimators), each requiring a minimum of 8 samples to split (min_samples_split) and 8 samples at a leaf node (min_samples_leaf). The model takes the whole dataset for each tree (max_samples: 1), evaluates 40% of features for each split (max_features: 0.4), and allows trees to develop without a stated maximum depth. The criterion for splitting nodes is based on entropy, and bootstrap sampling is enabled throughout tree construction. The XGBoost model is meticulously designed with precise hyperparameters. A subsample fraction of 0.7 defines the random sampling during each boosting round, and a total of 150 boosting rounds (n_estimators) are conducted. The minimal child weight is set to 1, indicating the minimum sum of instance weight needed in a child. The maximum depth of each tree is defined as 8. A low learning rate of 0.05 is applied to regulate the step size

Table 2. Utilized hyperparameters optimized through randomized search CV for ML models and proposed architecture for DL model.

Classifier	Hyperparameters		Classifier	Hyperparameters		
LR	Solver	Saga	DNN	Input layer	Dimension	20
	Penalty	l1		Hidden layer 1	Units	128
	Max_iter	200			Activation	Relu
	Class_weight	Balanced			Regularizer	L2
	C	100			Dropout	0.5
SVM	Kernel	Rbf			Batch normalization	True
	Gamma	1		Hidden layer 2	Units	64
	C	10			Activation	Relu
RF	n_estimators	50			Regularizer	L2
	Min_samples_split	8			Dropout	0.3
	Min_samples_leaf	8			Batch normalization	True
	Max_samples	1		Output layer	Units	1
	Max_features	0.4			Activation	Sigmoid
	Max_depth	None	Stacking	Base models	LR	Probability
	Criterion	Entropy	Ensemble		SVM	Probability
	Bootstrap	True			RF	Probability
XGBoost	Subsample	0.7			XGBoost	Probability
	n_estimators	150			DNN	Probability
	Min_child_weight	1		Final estimator	Logistic regression	
	Max_depth	8		CV	10	
	Learning_rate	0.05		Stack_method	Predict_proba	
	Gamma	0				
	Colsample_bytree	0.3				

shrinking and prevent overfitting. The gamma parameter is set to 0, necessitating a minimal loss reduction for further partition on a leaf node. Lastly, the fraction of features to be randomly selected for each tree (colsample_bytree) is determined as 0.3.

The specified DNN architecture comprises three layers with 128, 64, and 1 neuron(s) respectively, the model employs ReLU activation functions in the hidden layers and a sigmoid activation function in the output layer, tailored for binary classification tasks. The utilization of L2 regularization and Dropout layers significantly reduces the overfitting. Additionally, Batch Normalization is employed after each hidden layer, contributing to stable and efficient training.

The specialized hyperparameter selections work together to create a group of models that are accurately tuned, leading to trustworthy and context-specific predictions for areas that are prone to floods. By varying the parameters across the models, we can take advantage of each algorithm's strengths and achieve superior performance in the context of flood susceptibility mapping.

4. Evaluation of model performance

4.1. Multicollinearity results

To assess the presence of multicollinearity and choose variables for regression analysis, the Variance Inflation Factor (VIF) is used. However, VIF alone does not determine the factors that influence the dependent variable. To assess the extent of multicollinearity in the flood susceptibility modelling and to evaluate the significance of the chosen food-controlling factors, we calculated both the Variance Inflation Factor (VIF) and the tolerance value for multicollinearity. A collinearity problem is

Table 3. Multicollinearity test (tolerance and VIF values) of the food conditioning factors.

Collinearity statistics		
Factors	Tolerance	VIF
Elevation	0.431	2.322
Slope	0.723	1.382
TPI	0.661	1.513
TRI	0.52	1.922
TWI	0.457	2.19
Relief amplitude	0.702	1.424
Drainage density	0.769	1.301
Distance to drainage	0.746	1.341
Rainfall	0.685	1.459
MFI	0.768	1.301
NDVI	0.716	1.397
mNDWI	0.685	1.459
SPI	0.731	1.368
STI	0.874	1.144
Clay content	0.955	1.047
LULC	0.854	1.17
Geomorphology	0.953	1.05
Lithology	0.926	1.08

often identified when the Variance Inflation Factor (VIF) exceeds 10 and the tolerance value is below 0.1. The VIF and tolerance values for all flood conditioning factors in this research, as shown in Table 3, fall within the range of 10 to 0.01. This indicates that none of the chosen food conditioning variables exhibit any issues of multicollinearity. Based on the findings, there is no problem of multicollinearity.

4.2. Interpretability of machine learning by Shapley method

Shapley values, derived from cooperative game theory, offer a robust method for interpreting the contribution of each player (in our context, each machine learning model) towards the predictive performance of a coalition (ensemble). This is particularly useful in complex ensemble methods where understanding individual contributions is key to improving overall model performance and transparency.

The Shapley value can be calculated as (Equation (23)):

$$\phi_j(val) = \sum_{S \subseteq \{x_1, \ldots, x_p\}\ \{x_j\}} \frac{|S|!(p - |S| - 1)!}{p!} \left[val\left(S \cup \{x_j\}\right) - val(S) \right] \tag{23}$$

Where, S is a subset of the features used in the model, x is the vector of feature values of instance to be explained, p the number of features, and $val(S)$ is the prediction for feature values in set S marginalized over features that are not included in set S.

4.3. Statistical measures criteria

In this research, the performance of the machine learning (ML) models was evaluated using statistical metrics such as precision, recall, F1-score, accuracy, and area under

the curve (AUC) (Equations (24)–(29)). The use of these performance assessment methodologies is prevalent in academic literature (Eslaminezhad et al. 2022; Sellami et al. 2022; Ghanim et al. 2023; Saber et al. 2023).

4.3.1. Precision

Precision, a statistical metric employed to assess the precision of positive predictions made by a model, is determined by dividing the number of true positive predictions by the total number of positive predictions. This calculation aids in gauging the model's ability to minimize false positives, as demonstrated in Equation (24) (Eslaminezhad et al. 2022).

4.3.2. Recall

Recall, alternatively known as sensitivity or true positive rate, serves as a statistical measure for assessing a model's proficiency in correctly recognizing all pertinent instances among the overall positive instances. Its ability to minimize false negatives is emphasized in Equation (25) (Sellami et al. 2022), showcasing the model's effectiveness.

4.3.3. F1-score

The F1-score serves as a comprehensive metric that harmoniously combines accuracy and recall, striking a balance through their harmonic mean. It helps evaluate a model's overall performance in binary classification tasks by considering both false positives and false negatives shown in Equation (26) (Ghanim et al. 2023; Saber et al. 2023).

4.3.4. Accuracy

Accuracy serves as a statistical metric employed to assess the accuracy of predictions generated by a model. It is computed by dividing the count of accurate predictions by the overall number of predictions. However, its applicability may be limited in scenarios involving imbalanced datasets, where the emphasis on false positives or false negatives is highlighted, as demonstrated in Equation (27) (Eslaminezhad et al. 2022; Saber et al. 2023).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{24}$$

$$Recall = \frac{TP}{TP + FN} \tag{25}$$

$$F1 - Score = \frac{2(Precision \times Recall)}{Precision + Recall} \tag{26}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{27}$$

Where, 'TP = True Positive, and TN = True Negative are correctly predicted pixel numbers; FP = False Positive, and FN = False Negative are falsely predicted pixel numbers; P = Total number of floods pixels; N = Total number of non-floods pixels'.

4.4. Validation of evaluation of model performance

The validation of flood susceptibility maps holds considerable importance in the process of identifying and characterizing areas prone to flooding (Sellami et al. 2022). This study aims to evaluate the accuracy of flood susceptibility maps generated by various models using receiver operating characteristics (ROC) analysis (Ha et al. 2022). ROC analysis is a highly influential and widely utilized methodology in spatial modelling (Ghosh et al. 2022).

The ROC curve, a visual representation, illustrates a model's sensitivity in predicting the proportion of pixels accurately relative to the total predicted by the model. Constructing this curve involves plotting two statistical measures, 'sensitivity' and '1-specificity', on the y-axis and x-axis, respectively (Saber et al. 2023). Higher sensitivity indicates a significant proportion of correct predictions (true positives), while higher specificity indicates a low probability of false positives. The research graph depicts the correlation between the false positive rate (1-specificity) on the X-axis and the true positive rate (sensitivity) on the Y-axis (Ghanim et al. 2023) Equations (28) and (29).

$$X = 1 - specificity = 1 - \left[\frac{TN}{(TN + FP)}\right] \tag{28}$$

$$Y = sensitivity = \left[\frac{TN}{(TP + FN)}\right] \tag{29}$$

The evaluation of the predicted accuracy of flood susceptibility models involves the assessment of the area under the curve (AUC) of the forecasting degree (Youssef, Pradhan, et al. 2022). The categorization of the qualitative relationship between the area under the ROC curve (AUC) and the prediction performance of a model has been classified into five separate groups (Li and Hong 2023). The classes are classified as follows: 0.5–0.6 (poor), 0.6–0.7 (average), 0.7–0.8 (good), 0.8–0.9 (very good), and 0.9–1 (excellent).

5. Results

5.1. Flood susceptibility zonation (FSZ)

Flood susceptibility maps of the Malda area are developed using six advanced machine learning (ML) models: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Deep Neural Network (DNN) and Stacking Ensemble. These models are implemented inside a RS and GIS context. The final map is categorized into five distinct classifications, namely very low, low, moderate, high, and very high susceptibility to floods (Table 4). As seen in Figure 6, the maps illustrate that the regions exhibiting significant sensitivity are mostly localized within the region and along the course of the River. Furthermore, it is evident that the categories with a significantly high susceptibility include a substantial quantity of training data points that were derived from the affected areas during the occurrence of the flood. This empirical evidence serves to validate the robust performance of the

Table 4. Area of flood susceptibility zonation of Malda district, WB.

	Logistic Regression (LR)		Support Vector Machines (SVM)		Random Forest (RF)		Extreme Gradient Boosting (XGBoost)		Deep Neural Network (DNN)		Stacking Ensemble	
Flood susceptibility levels	Area in sq. km	Area in (%)	Area in sq. km	Area in (%)	Area in sq. km	Area in (%)	Area in sq. km	Area in (%)	Area in sq. km	Area in (%)	Area in sq. km	Area in (%)
Very low susceptibility	*1537*.61	41.94	1751.92	47.79	1270.56	34.66	1535.85	41.90	1097.99	29.95	1703.27	46.46
Low susceptibility	438.22	11.95	238.24	6.50	553.27	15.09	421.01	11.48	707.38	19.30	300.36	8.19
Moderate susceptibility	451.43	12.31	214.43	5.85	558.33	15.23	361.35	9.86	521.63	14.23	239.24	6.53
High susceptibility	406.42	11.09	434.95	11.86	562.02	15.33	404.78	11.04	416.72	11.37	265.28	7.24
Very high susceptibility	832.26	22.70	1026.40	28.00	721.78	19.69	942.97	25.72	922.23	25.16	1157.81	31.58
Total area	3665.95	100.00	3665.94	100.00	3665.95	100.00	3665.95	100.00	3665.95	100.00	3665.95	100.00

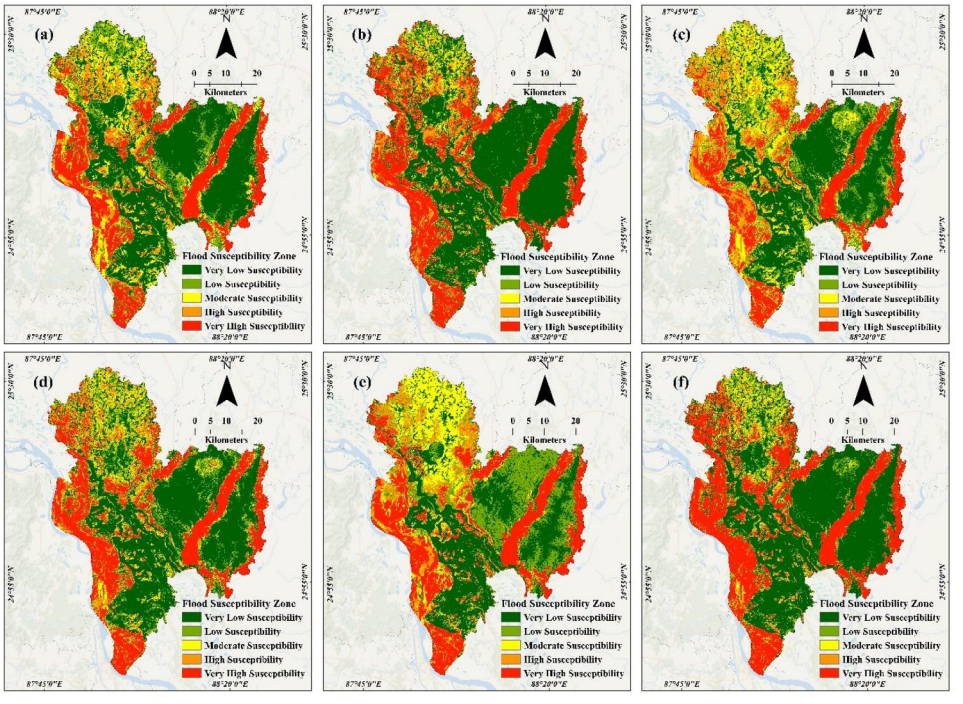

Figure 6. Flood susceptibility zone (FSZ) (a) Logistic Regression (LR), (b) Support Vector Machines (SVM), (c) Random Forest (RF), (d) Extreme Gradient Boosting (XGBoost), (e) Deep Neural Network (DNN), and (f) Stacking Ensemble.

model (Ramayanti et al. 2022). The classifications are denoted as 'very high', 'high', 'moderate', 'low', and 'very low' in sequential order, signifying the accurate training of the models (Figure 7). This was further validated by a comparison between the predicted data and real-world data. According to the Logistic Regression (LR) model (Figure 6(a)), the study region is classified as having 832.26 sq. km (22.70% of total area) with very high susceptibility, 406.42 sq. km (11.09% of total area) with high susceptibility, 451.43 sq. km (12.31% of total area) with a moderate susceptibility, 438.22 sq. km (11.95% of total area) with low susceptibility, and 1537.61 sq. km (41.94% of total area) with a very low susceptibility (Figure 7). According to the findings of the Support Vector Machine (SVM) model (Figure 6(b)), it has been determined that 1026.40 sq. km (28.00% of total area) of the study area exhibits a significantly very high susceptibility to flooding, 434.95 sq. km (11.86% of total area) of the study area demonstrates a high susceptibility, 214.43 sq. km (5.85% of total area) displays a moderate susceptibility, 238.24 sq. km (6.50% of total area) showcases a low susceptibility, and the majority, 1751.92 sq. km (47.79% of total area), exhibits a very low susceptibility to flooding (Figure 7). According to the findings of the Random Forest (RF) model (Figure 6(c)), it has been determined that within the study region, 721.78 sq. km (19.69% of total area) exhibits a very high susceptibility to flooding, 562.02 sq. km (15.33% of total area) of the research area demonstrates a high susceptibility, 558.33 sq. km (15.23% of total area) displays a moderate susceptibility, 553.27 sq. km (15.09% of total area) showcases a low susceptibility, and *1270*.56 (34.66% of total)

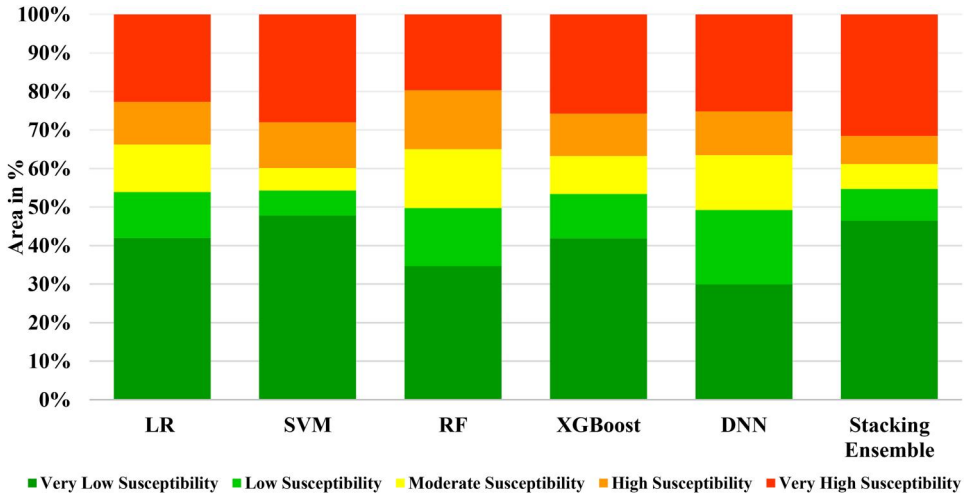

Figure 7. FSZ area coverage by different models.

area exhibits a very low sensitivity to flooding (Figure 7). The final model used in this research is the Extreme Gradient Boosting (XGBoost) model (Figure 6(d)). The results indicate that 942.97 sq. km (25.72% of total area) of the study region exhibits a very high susceptibility to flooding, 404.78 sq. km (11.04% of total area) of the area demonstrates a high susceptibility, 361.35 sq. km (9.86% of total area) displays a moderate susceptibility, 421.01 sq. km (11.48% of total area) exhibits a low susceptibility, and the remaining 1535.85 sq. km (41.90% of total area) showcases the lowest vulnerability to flooding (Figure 7). For Deep Neural Network (DNN) (Figure 6(e)), the distribution across zones was as follows: very high zone 922.23 sq. km (25.16% of total area), high zone 416.72 sq. km (11.37% of total area), moderate zone 521.63 sq. km (14.23% of total area), low zone 707.38 sq. km (19.30% of total area), very low zone *1097*.99 sq. km (29.95% of total area) (Figure 7). For Stacking Ensemble (Figure 6(f)), the distribution across zones was as follows: very high zone *1157*.81 sq. km (31.58% of total area), high zone 265.28 sq. km (7.24% of total area), moderate zone 239.24 sq. km (6.53% of total area), low zone 300.36 sq. km (8.19% of total area), very low zone 1703.27 sq. km (46.46% of total area) (Figure 7).

The results emphasize a significant flood susceptibility (FS) in the Tal and Diara regions located in the western part of the district. The younger, lower floodplains along the Ganga, Fulhar, Kalindi, and Mahananda rivers demonstrate an increased vulnerability to floods. Several factors, such as lower altitude, gradual slope, intense rainfall, high drainage density, and elevated water levels, contribute to this susceptibility. During the rainy season, there is a heightened risk of riverbed siltation and back thrust action at confluences. In contrast, the Barind region in the eastern half generally faces a comparatively lower threat of flooding, except for the Tangon and Punarbhava riverine floodplains. The eastern part exhibits a moderate to low level of flood susceptibility due to its higher elevation, moderate slope, and undulating terrain. A detailed block-level analysis in Malda District, comprising 15 administrative blocks, underscores 11 blocks with the highest susceptibility (Figure 8). Notably, areas

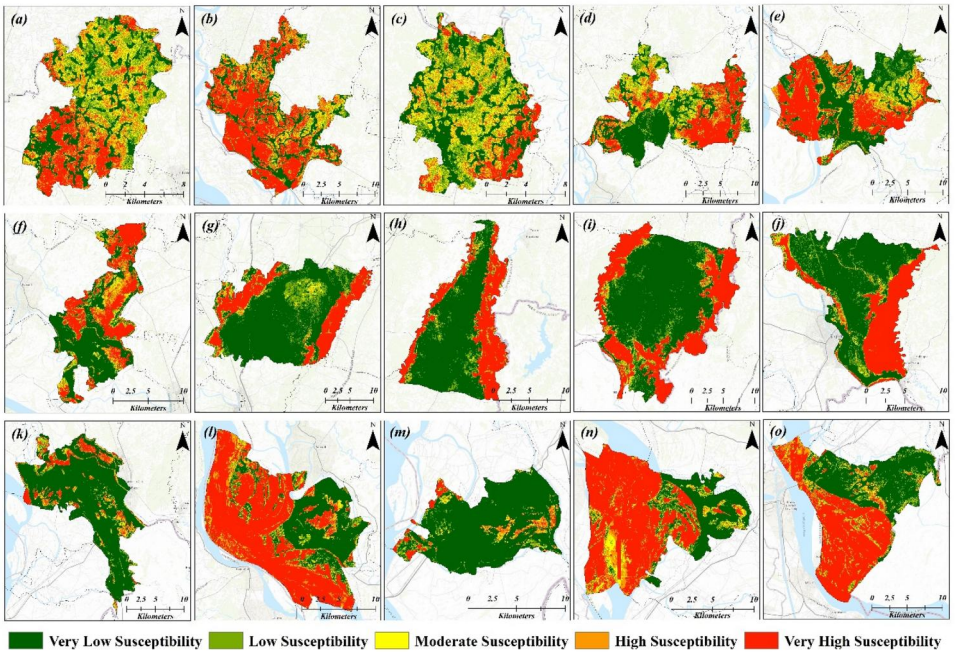

Figure 8. CD Block-wise flood susceptibility zone distribution (a) Harischandrapur-I, (b) Harischandrapur-II, (c) Chanchal-I, (d) Chanchal-II, (e) Ratua-I, (f) Ratua-II, (g) Gazole, (h) Bamangola, (i) Habibpur, (j) Maldah (Old), (k) English Bazar, (l) Manikchak, (m) Kaliachak-I, (n) Kaliachak-II, (o) Kaliachak-III.

at the highest risk include Harischandrapur-I & II, Kaliachak-II & III, Manikchak, and Ratua-I & II blocks, situated amidst the Ganga, Fulahar, Kalindi, and Mahananda rivers (Figure 8). Conversely, regions with a moderate susceptibility level include the Bamangola, Gajole, Habibpur, and Old Malda blocks, located between the Mahananda and Tangon rivers, as well as the Punarbhava and Tangon interfluves (Figure 8). The conclusive findings of this study firmly establish Malda as a flood-prone district in West Bengal, where specific environmental conditions contribute to the annual recurrence of floods.

5.2. Influence of explanatory variables

SHAP values, also known as SHapley Additive exPlanations, are an effective method for evaluating the results of machine learning models by attributing a unique contribution to each feature in creating predictions. These values are based on cooperative game theory and ensure an equitable distribution of the model's prediction across each feature (Lyu and Yin 2023; Pradhan et al. 2023). In the context of our investigation, SHAP values were applied to acquire insights into the impact of explanatory variables across all the models tested in the current study. Specifically, a SHAP bar plot was constructed to graphically represent the distribution and amplitude of the SHAP values for each feature throughout the dataset (Liu et al. 2023).

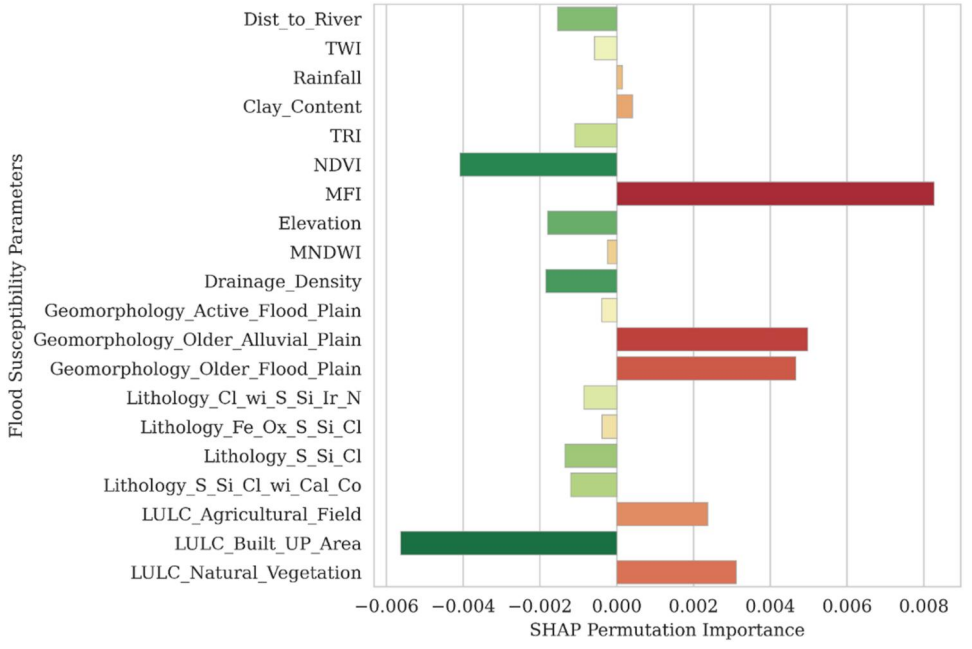

Figure 9. Permutation importance of explanatory variables across models using SHAP values.

A significant spread in permutation importance of the Modified Fourier Index (MFI, a value of + 0.00826) and Rainfall (value of + 0.000148), concomitant with other affecting factors, has been noticed (Figure 9). This remarkable heterogeneity in the range of SHAP values reflects the diversified and large impact of MFI and Rainfall on the predicted outcomes of these models. On the other hand, a dispersion in the features of Geomorphologically Active Flood Plain and Elevation, in conjunction with other influencing factors, has been observed. This large variety in the distribution of SHAP values demonstrates the diversified and influential nature of these two factors on the predicted outcomes of these ensemble models.

5.3. Model evaluation

Six distinct machine learning (ML) algorithms were applied to predict flood susceptibility in the district. These algorithms include Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGboost), Deep Neural Network (DNN) and Stacking Ensemble. Table 5 illustrates the validation metrics for each of the four machine-learning models utilized in our study. These metrics provide a quantitative assessment of the models' performance and their suitability for flood susceptibility prediction.

For Logistic Regression (LR) statistical metricises are, Precision: 0.825 (for output 0) and 0.835 (for output 1), Recall: 0.817 (for output 0) and 0.842 (for output 1), F1-score: 0.821 (for output 0) and 0.839 (for output 1), Accuracy: 0.83. Being a linear model, LR exhibits decent performance with a balanced F1-score for both binary classes. It reaches an accuracy of 83%. For Support Vector Machine (SVM) statistical

Table 5. Validation metrics.

Classifier	Output	Precision	Recall	F1-score	Accuracy	AUC
Logistic Regression (LR)	0	0.825	0.817	0.821	0.83	0.921
	1	0.835	0.842	0.839		
Support Vector Machine (SVM)	0	0.869	0.820	0.844	0.855	0.92
	1	0.845	0.887	0.865		
Random Forest (RF)	0	0.849	0.836	0.842	0.851	0.925
	1	0.853	0.865	0.859		
Extreme Gradient Boosting (XGBoost)	0	0.847	0.842	0.845	0.853	0.934
	1	0.857	0.862	0.860		
Deep Neural Network (DNN)	0	0.866	0.820	0.843	0.854	0.929
	1	0.844	0.885	0.864		
Stacking Ensemble	0	0.885	0.885	0.885	0.891	0.965
	1	0.896	0.896	0.896		

metricises are, Precision: 0.869 (for output 0) and 0.845 (for output 1), Recall: 0.82 (for output 0) and 0.887 (for output 1), F1-score: 0.844 (for output 0) and 0.865 (for output 1), Accuracy: 0.855. SVM demonstrates improved overall performance with higher precision, recall, F1-score, and accuracy compared to LR. This suggests its competence to handle difficult decision boundaries and properly classify data. SVM exhibits an overall accuracy of about 85.5%. Ensemble models like RF, which combine multiple decision trees, are typically known for their robustness and accuracy. For Random Forest (RF) statistical metricises are, Precision: 0.849 (for output 0) and 0.853 (for output 1), Recall: 0.836 (for output 0) and 0.865 (for output 1), F1-score: 0.842 (for output 0) and 0.859 (for output 1), Accuracy: 0.851. RF performs well, obtaining balanced performance metrics for both classes. Its accuracy of 85.1% implies that it successfully exploits the benefits of ensemble learning, leading to good classification results. XGBoost, a powerful ensemble learning algorithm, leverages the power of advanced gradient boosting techniques. It is quite popular for its strong predictive accuracy and scalability. For Extreme Gradient Boosting (XGBoost), Precision: 0.847 (for output 0) and 0.857 (for output 1), Recall: 0.842 (for output 0) and 0.862 (for output 1), F1-score: 0.845 (for output 0) and 0.86 (for output 1), Accuracy: 0.853. XGBoost displays performance on par with RF, displaying good precision, recall, and F1-scores. The model's accuracy of 85.3% implies that it efficiently employs gradient boosting to give competitive classification results. For Deep Neural Network (DNN), Precision: 0.866 (for output 0) and 0.844 (for output 1), Recall: 0.820 (for output 0) and 0.885 (for output 1), F1-score: 0.843 (for output 0) and 0.864 (for output 1), Accuracy: 0.854. For Stacking Ensemble, Precision: 0.885 (for output 0) and 0.896 (for output 1), Recall: 0.885 (for output 0) and 0.896 (for output 1), F1-score: 0.885 (for output 0) and 0.896 (for output 1), Accuracy: 0.891.

The Receiver Operating Characteristic (ROC) curve is a significant tool for analysing the classification performance of machine learning models. It provides insights into a model's efficacy in discriminating between positive and negative instances. The Area Under the ROC Curve (AUC) is a statistic that quantifies the overall performance of a model, with a higher AUC value suggesting more discriminatory power. The AUC scores for our machine learning models were as follows: Stacking Ensemble achieved the highest AUC (0.965), followed closely by, XGBoost (0.934), DNN (0.929), RF (0.925), LR (0.921), and SVM (0.920). These AUC values represent the

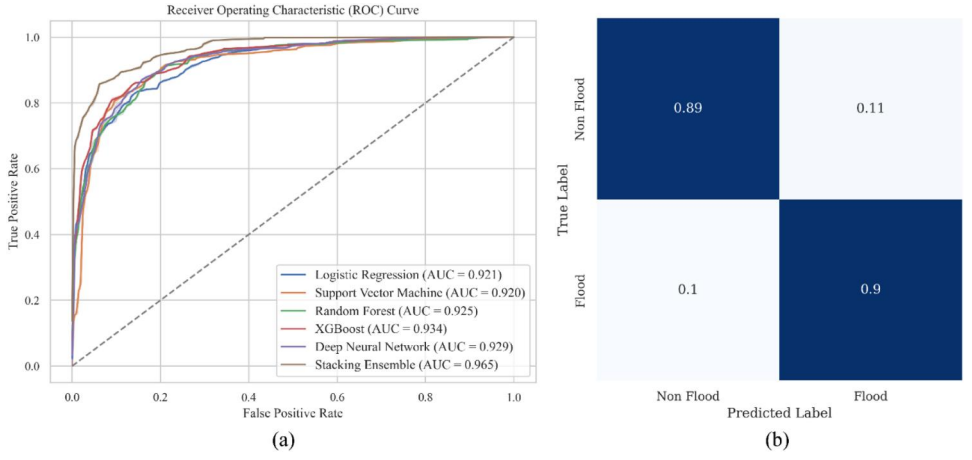

Figure 10. (a) Performance of Logistic Regression (LR), Support Vector Machines (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Deep Neural Network (DNN), and Stacking Ensemble models based on AUC-ROC curves, (b) Confusion matrix of Stacking Ensemble model.

models' ability to differentiate between flood-prone and non-flood-prone locations (Figure 10).

The Stacking Ensemble model shows the highest performance across almost all metrics, notably achieving the highest accuracy (0.891) and AUC (0.965). This suggests that the ensemble approach, which combines multiple models, provides a significant improvement over individual model. This improvement is seen in its ability to generalize better and make more accurate predictions. The performance metrics indicate the trade-offs between different models; for example, while the SVM shows high precision for output 0, the Stacking Ensemble balances high precision, recall, and F1-scores across both output classes, leading to its overall superior performance.

Conclusively, our analysis shows that areas along the Ganga, Fulhar, Kalindi, and Mahananda rivers, particularly the Tal and Diara regions, exhibit the highest flood susceptibility due to factors like low altitude, gradual slopes, and high rainfall. In contrast, the Barind region to the east, with its higher elevation and moderate slopes, faces lower flood risks, except for the floodplains of the Tangon and Punarbhava rivers. Different models provided varying susceptibility rates, with the Stacking Ensemble model achieving the highest performance, identifying 31.58% of the area as highly susceptible and showing the highest accuracy (89.1%) and AUC (0.965). SHAP values highlighted the significant impact of features like the Modified Fourier Index and rainfall on flood predictions. The findings emphasizes Malda's vulnerability to flooding, particularly in specific blocks such as Harischandrapur-I & II and Kaliachak-II & III, and emphasize the importance of these models in guiding effective flood risk management and disaster preparedness strategies in the district. Moreover, high-risk zones, identified in the current study, such as the Tal and Diara regions, frequently experience floods due to low altitude, gradual slopes, intense rainfall, high drainage density, and elevated water levels, which are exacerbated during the rainy season by riverbed siltation and back thrust action at confluences. The flood susceptibility maps produced not only correspond to past flood events but also provide an

accurate representation of flood-prone areas, enhancing the reliability of the models for guiding flood risk management and disaster preparedness strategies. Utilizing these outcomes, authorities can better manage flood-vulnerable regions through targeted interventions. Tal and Diara regions can benefit greatly from improved infrastructure, such as enhanced drainage systems, flood barriers, flood-proofing homes, creating community-based disaster management plans and early warning systems. Additionally, optimized land-use planning and tailored emergency response plans will minimize flood damage and focus on the most vulnerable areas. The data produced in the current work also informs community education programs, ensuring residents are better prepared for flood events. Authorities can prioritize investments in flood defences and resilient construction practices, minimizing damage costs and insurance premiums. Farmers can make informed decisions about crop planting and harvesting, reducing losses and enhancing food security.

6. Discussion

The present study conducted a comprehensive flood susceptibility mapping analysis in the Malda district of West Bengal, utilizing six distinct machine learning (ML) algorithms: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Deep Neural Network (DNN) and Stacking Ensemble. The research integrated a wide range of flood conditioning parameters to create a holistic understanding of the flood susceptibility dynamics within the region. The flood conditioning parameters considered in this study included elevation, slope, TPI, TRI, TWI, relief amplitude, drainage density, distance to river, annual rainfall, MFI, NDVI, mNDWI, SPI, STI, LULC, geomorphology, and lithology. The use of multiple machine learning (ML) algorithms in this research enabled a robust and reliable assessment of flood susceptibility in the Malda district. Each algorithm contributed unique advantages to the analysis. Logistic Regression (LR) was employed to establish a foundational understanding of the relationship between flood susceptibility and the selected conditioning parameters (Li et al. 2022; Saikh and Mondal 2023). It provided insights into the statistical significance and direction of influence of individual parameters, facilitating the identification of key drivers of flood susceptibility. Support Vector Machine (SVM) was employed to capture complex nonlinear relationships within the dataset and define a hyperplane that maximized the separation between flood-prone and non-flood-prone areas (Gharakhanlou and Pérez 2023). Its ability to work well with high-dimensional datasets made it well-suited for integrating the numerous conditioning parameters considered in this study. Random Forest (RF) an ensemble learning technique, was utilized to address the challenge of overfitting in the model (Satarzadeh et al. 2021; Li et al. 2022). By aggregating the predictions of multiple decision trees, RF improved the model's predictive performance while considering the importance of individual parameters. This allowed for a more accurate assessment of flood susceptibility. Extreme Gradient Boosting (XGBoost) another ensemble method, was employed to further enhance predictive accuracy by minimizing both bias and variance in the model (Ghanim et al. 2023). Its ability to handle imbalanced datasets was particularly

advantageous for flood susceptibility mapping, where flood events are typically rare compared to non-flood events. Deep Neural Networks (DNN) are sophisticated ML algorithms that mimic the workings of the human brain, capable of capturing complex patterns through their multiple layers and neurons. In this study, a DNN architecture was designed to process large volumes of data and extract intricate patterns related to flood susceptibility. This approach allowed for a deep learning-based interpretation of the data, leveraging the ability of DNNs to model complex nonlinear relationships. Stacking Ensemble is an advanced machine-learning technique that combines multiple prediction models *via* a meta-learner. In this context, the predictions from LR, SVM, RF, XGBoost, and DNN were aggregated to create a final ensemble model. This method aims to capitalize on the strengths of each model and mitigate its weaknesses, providing a robust and accurate flood susceptibility prediction. By integrating these diverse machine learning algorithms, the study aimed to harness the complementary strengths of each method, thereby enhancing the reliability and accuracy of flood susceptibility maps for the Malda district. This comprehensive approach not only contributes to the scientific understanding of flood risks but also serves as a valuable tool for planners and policymakers in implementing effective flood risk management strategies.

The incorporation of a diverse set of flood conditioning parameters provided a comprehensive view of the factors influencing flood susceptibility in the Malda district. Key factors such as elevation, drainage density, and distance to rivers were expected to play crucial roles due to their direct influence on flood dynamics. The inclusion of topographic indices (TPI, TRI, TWI) helped capture terrain characteristics that contribute to flood susceptibility, while meteorological parameters (annual rainfall, MFI) and hydrological parameters (SPI, STI) provided insights into precipitation and hydrological conditions. Vegetation indices (NDVI and mNDWI) were used to assess the impact of land cover and water bodies on flood susceptibility, and the consideration of LULC, geomorphology, and lithology allowed for the evaluation of land-use patterns and geological factors. The findings of this research offer valuable insights into flood susceptibility patterns in the Malda district, which can inform disaster management, land-use planning, and infrastructure development. Additionally, the methodology and machine learning (ML) algorithms employed in this study can be adapted for flood susceptibility mapping in other regions, contributing to improved flood risk assessment and resilience-building efforts.

Despite the robust findings, the study's reliance on Synthetic Aperture Radar (SAR) data for certain parameters introduces several limitations. SAR data often suffer from limited temporal resolution, hindering the capture of rapid changes in flood dynamics during extreme weather events. Additionally, the spatial resolution of SAR data can vary, potentially missing finer details of flood-prone areas, particularly in complex terrains. The accuracy of SAR signals can also be compromised by vegetation and surface roughness, leading to potential inaccuracies in regions with dense vegetation or uneven terrain. In particular, the SAR thresholding technique employed in this study has intrinsic limitations when it comes to identifying floods in urban and densely vegetated areas. Urban areas and densely vegetated regions have complex surface characteristics that can hinder the accurate detection of floodwater using

SAR-based methods. The method's insensitivity to these conditions sometimes masks out these areas, leading to false negatives in the training data (Zhao et al. 2021). Furthermore, obtaining high-quality, up-to-date SAR data can be challenging, which limits the model's applicability in regions where such data are scarce or outdated. These limitations highlight the need for supplementary data sources and advanced processing techniques to enhance the accuracy and reliability of flood susceptibility assessments.

7. Conclusion

Floods are widely acknowledged as highly destructive natural disasters that pose significant threats to various aspects of human life. The development of a flood susceptibility map is an essential element in the execution of a complete flood management plan, since it aims to reduce human fatalities and mitigate flood-induced losses. A significant obstacle encountered in several nations throughout the process of generating susceptibility maps is the insufficiency of appropriate and current field data that can be used for the training phase of machine learning techniques. Additionally, the generation of flood susceptibility maps is often influenced by two primary factors: effectiveness and accuracy. This research used a combination of RS & GIS, GEE, and Python capabilities to achieve the necessary levels of speed and accuracy in generating susceptibility maps. Climate change and global warming are phenomena that cannot be avoided, and natural disasters are prevalent worldwide. One example is the flood, which can result in significant damage. This prompts us to consider identifying more efficient solutions to mitigate the magnitude of this issue. Therefore, the model that has been produced may be used in other places by gathering the flood conditioning components that are relevant to the specific region being examined. The research may also be used in the field of urban planning to identify locations with a high susceptibility to flooding, therefore informing decisions about the approval of new structures in order to mitigate future risks. This research also has the potential to provide benefits to a range of stakeholders, including policymakers, government ministries, authorities, local administrative bodies, environmentalists, planners, and engineers, who work together to minimize vulnerability to flooding. While the study on flood susceptibility mapping in the Malda district using advanced machine learning models shows promising results, several limitations must be acknowledged. Firstly, the study relies on complex machine learning models which, although powerful, pose a risk of overfitting due to the high dimensionality of the data and the complexity of the models themselves. This overfitting may limit the generalizability of the findings, making it difficult to apply these results to other regions with different geographical or climatic conditions. Additionally, the accuracy and effectiveness of the susceptibility maps are highly dependent on the quality and availability of relevant field data. Although the Malda district benefits from relatively comprehensive datasets, such data may be outdated or unavailable in other regions, thereby limiting the reliability and applicability of the models elsewhere.

Furthermore, the models used may incorporate simplified hydrological assumptions, potentially overlooking complex processes like river flow dynamics,

groundwater interactions, and urban drainage systems that are crucial for a comprehensive flood risk assessment. Adding features form process based hydrological dynamics may further improve the accuracy of the models used in the study. Our approach assumes that all years are independent, which may be too strong of an assumption if consecutive years have correlated environmental and climatic condition, which could have impact on model accuracy. The study also does not take into account socioeconomic factors such as population density, land use planning, or infrastructure resilience, which are essential for a holistic understanding of flood risk and its impact on communities. The training data's focus on flood-affected areas might introduce bias, leading to models that are overly sensitive to these areas and potentially underestimating susceptibility in regions with less historical flood data. Lastly, the analysis is based on historical data and does not consider potential future changes in flood patterns due to climate change, thus reducing the long-term applicability of the susceptibility maps for future flood risk management and planning. These limitations underscore the need for further research and methodological improvements to enhance the robustness and applicability of flood susceptibility mapping.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This publication was supported by the Deanship of Scientific Research at the King Faisal University, Saudi Arabia (grant: 5823). We also received the funding from the German Federal Ministry of Education and Research (BMBF) in the framework of the funding measure 'Soil as a Sustainable Resource for the Bioeconomy—BonaRes', project BonaRes (Module A): BonaRes Center for Soil Research, subproject 'Sustainable Subsoil Management—Soil3'(Grant 031B0151A). In addition, the work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2070—390732324. In addition, we also received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2070—390732324 and COINS (Grant 01LL2204C).

Data availability statement

Data are available on request from the authors.

References

Abedi R, Costache R, Shafizadeh-Moghadam H, Pham QB. 2021. Flash-flood susceptibility mapping based on XGBoost, random forest and boosted regression trees. Geocarto Int. 37(19):5479–5496. doi: 10.1080/10106049.2021.1920636.

Amiri A, Soltani K, Ebtehaj I, Bonakdari H. 2024. A novel machine learning tool for current and future flood susceptibility mapping by integrating remote sensing and geographic information systems. J Hydrol. 632:130936. doi: 10.1016/j.jhydrol.2024.130936.

Amitrano D, Di Martino G, Di Simone A, Imperatore P. 2024. Flood detection with SAR: a review of techniques and datasets. Remote Sens. 16(4):656. doi: 10.3390/rs16040656.

Antzoulatos G, Kouloglou I, Bakratsas M, Moumtzidou A, Gialampoukidis I, Karakostas A, Lombardo F, Fiorin R, Norbiato D, Ferri M, et al. 2022. Flood hazard and risk mapping by applying an explainable machine learning framework using satellite imagery and GIS data. Sustainability. 14(6):3251. doi: 10.3390/su14063251.

Bentivoglio R, Isufi E, Jonkman SN, Taormina R. 2022. Deep learning methods for flood mapping: a review of existing applications and future research directions. Hydrol Earth Syst Sci. 26(16):4345–4378. doi: 10.5194/hess-26-4345-2022.

Bera S, Das A, Mazumder T. 2022. Evaluation of machine learning, information theory and multi-criteria decision analysis methods for flood susceptibility mapping under varying spatial scale of analyses. Remote Sens Appl Soc Environ. 25:100686. doi: 10.1016/j.rsase.2021.100686.

Beven KJ, Kirkby MJ. 1979. A physically based, variable contributing area model of basin716-hydrology. Hydrol Sci J. 24(1):43–69. doi: 10.1080/71702626667909491834.718.

Bhatta S, Adhikari BR. 2024. Comprehensive risk evaluation in Rapti Valley, Nepal: a multi-hazard approach. Prog Disaster Sci. 23:100346. doi: 10.1016/j.pdisas.2024.100346.

Chang K, Merghadi A, Yunus AP, Pham BT, Dou J. 2019. Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques. Sci Rep. 9(1):12296. doi: 10.1038/s41598-019-48773-2.

CRED and UNISDR. 2016. The human cost of weather related disasters 1995-2015. https://www.unisdr.org/2015/docs/climatechange/COP21_WeatherDisastersReport_2015_FINAL.pdf.

Das S, Gupta A. 2021. Multi-criteria decision based geospatial mapping of flood susceptibility and temporal hydro-geomorphic changes in the Subarnarekha basin, India. Geosci Front. 12(5):101206. doi: 10.1016/j.gsf.2021.101206.

Debnath J, Sahariah D, Nath N, Saikia A, Lahon D, Islam MN, Hashimoto S, Meraj G, Kumar P, Singh SK, et al. 2023. Modelling on assessment of flood risk susceptibility at the Jia Bharali River basin in Eastern Himalayas by integrating multicollinearity tests and geospatial techniques. Model Earth Syst Environ. 10(2):2393–2419. doi: 10.1007/s40808-023-01912-1.

Dey H, Shao W, Moradkhani H, Keim BD, Peter BG. 2024. Urban flood susceptibility mapping using frequency ratio and multiple decision tree-based machine learning models. Nat Hazards. 120(11):10365–10393. doi: 10.1007/s11069-024-06609-x.

Diaconu DC, Costache R, Islam ARMT, Pandey M, Pal SC, Mishra AP, Pande CB. 2024. Developing flood mapping procedure through optimized machine learning techniques. Case study: Prahova river basin, Romania. J Hydrol Region Stud. 54:101892. doi: 10.1016/j.ejrh.2024.101892.

Dou J, Yunus AP, Bui DT, Merghadi A, Sahana M, Zhu Z, Chen C, Han Z, Pham BT. 2019. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. Landslides. 17(3):641–658. doi: 10.1007/s10346-019-01286-5.

Dou J, Yunus AP, Merghadi A, Shirzadi A, Nguyen H, Hussain Y, Avtar R, Chen Y, Pham BT, Yamagishi H. 2020. Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning. Sci Total Environ. 720:137320. doi: 10.1016/j.scitotenv.2020.137320.

Dutta M, Saha S, Saikh NI, Sarkar D, Mondal P. 2023. Application of bivariate approaches for flood susceptibility mapping: a district level study in Eastern India. HydroResearch. 6:108–121. doi: 10.1016/j.hydres.2023.02.004.

Eslaminezhad SA, Eftekhari M, Azma A, Kiyanfar R, Akbari M. 2022. Assessment of flood susceptibility prediction based on optimized tree-based machine learning models. J Water Clim Change. 13(6):2353–2385. doi: 10.2166/wcc.2022.435.

Fatah KK, Mustafa YT, Hassan I. 2022. Flood susceptibility mapping using an analytic hierarchy process model based on remote sensing and GIS approaches in Akre district, Kurdistan region, Iraq. IGJ. 55(2C):121–149. doi: 10.46717/igj.55.2C.10ms-2022-08-23.

Ghanim AaJ, Shaf A, Ali T, Zafar M, Al-Areeq AM, Alyami SH, Irfan M, Rahman S. 2023. An improved flood susceptibility assessment in Jeddah, Saudi Arabia, using advanced machine learning techniques. Water. 15(14):2511. doi: 10.3390/w15142511.

Gharakhanlou NM, Pérez L. 2023. Flood susceptible prediction through the use of geospatial variables and machine learning methods. J Hydrol. 617:129121. doi: 10.1016/j.jhydrol.2023.129121.

Ghobadi M, Ahmadipari M. 2024. Enhancing flood susceptibility modeling: a hybrid deep neural network with statistical learning algorithms for predicting flood prone areas. Water Resour Manage. 38(8):2687–2710. doi: 10.1007/s11269-024-03770-7.

Ghosh A, Chatterjee U, Pal SC, Islam ARMT, Alam E, Islam MK. 2023. Flood hazard mapping using GIS-based statistical model in vulnerable riparian regions of sub-tropical environment. Geocarto Int. 38(1). doi: 10.1080/10106049.2023.2285355.

Ghosh A, Dey P, Ghosh T. 2022. Integration of RS-GIS with frequency ratio, fuzzy logic, logistic regression and decision tree models for flood susceptibility prediction in lower gangetic plain: a study on Malda District of West Bengal, India. J Indian Soc Remote Sens. 50(9):1725–1745. doi: 10.1007/s12524-022-01560-5.

Ghosh B, Garg S, Motagh M, Martinis S. 2024. Automatic flood detection from Sentinel-1 data using a nested UNet model and a NASA benchmark dataset. PFG. 92(1):1–18. doi: 10.1007/s41064-024-00275-1.

Ghosh S, Saha S, Bera B. 2022. Flood susceptibility zonation using advanced ensemble machine learning models within Himalayan foreland basin. Nat Hazards Res. 2(4):363–374. doi: 10.1016/j.nhres.2022.06.003.

Ha H, Bui QD, Khuc TD, Tran DT, Pham BT, Mai SH, Nguyen LP, Luu C. 2022. A machine learning approach in spatial predicting of landslides and flash flood susceptible zones for a road network. Model Earth Syst Environ. 8(4):4341–4357. doi: 10.1007/s40808-022-01384-9.

Harshasimha AC, Bhatt CM. 2023. Flood vulnerability mapping using MaxEnt machine learning and analytical hierarchy process (AHP) of Kamrup Metropolitan District, Assam. Environ Sci Proc. 25(1):73. doi: 10.3390/ecws-7-14301.

Hitouri S, Mohajane M, Lahsaini M, Ali SA, Setargie TA, Tripathi G, D'Antonio P, Singh SK, Varasano A. 2024. Flood susceptibility mapping using SAR data and machine learning algorithms in a small watershed in northwestern Morocco. Remote Sens. 16(5):858. doi: 10.3390/rs16050858.

Irrigation & Waterways Department, Government of West Bengal, India. 2022. Annual flood report 2022. https://www.wbiwd.gov.in/uploads/anual_flood_report/Annual_Flood_Report_2022.pdf.

Jenness J, Brost B, Beier P. 2013. Land facet corridor designer. USDA Forest Service Rocky Mountain Research Station.

Karunanayake C, Gunathilake MB, Rathnayake U. 2020. Inflow forecast of Iranamadu Reservoir, Sri Lanka, under projected climate scenarios using artificial neural networks. Appl Comput Intell Soft Comput. 2020:1–11. doi: 10.1155/2020/8821627.

Khaniya B, Gunathilake MB, Rathnayake U. 2021. Ecosystem-based adaptation for the impact of climate change and variation in the water management sector of Sri Lanka. Math Prob Eng. 2021:1–10. doi: 10.1155/2021/8821329.

Li J, Zhang H, Zhao J, Guo X, Rihan W, Deng G. 2022. Embedded feature selection and machine learning methods for flash flood susceptibility-mapping in the mainstream Songhua River Basin, China. Remote Sens. 14(21):5523. doi: 10.3390/rs14215523.

Li Y, Hong H. 2023. Modelling flood susceptibility based on deep learning coupling with ensemble learning models. J Environ Manage. 325(Pt A):116450. doi: 10.1016/j.jenvman.2022.116450.

Liu J, Liu K, Wang M. 2023. A residual neural network integrated with a hydrological model for global flood susceptibility mapping based on remote sensing datasets. Remote Sens. 15(9):2447. doi: 10.3390/rs15092447.

Lyu H, Yin Z. 2023. Flood susceptibility prediction using tree-based machine learning models in the GBA. Sustain Cities Soc. 97:104744. doi: 10.1016/j.scs.2023.104744.

Maharjan M, Timilsina S, Ayer S, Singh B, Manandhar B, Sedhain A. 2024. Flood susceptibility assessment using machine learning approach in the Mohana-Khutiya River of Nepal. Nat Hazards Res. 4(1):32–45. doi: 10.1016/j.nhres.2024.01.001.

Mansour A, Mrad D, Djebbar Y. 2024. Advanced modeling for flash flood susceptibility mapping using remote sensing and GIS techniques: a case study in Northeast Algeria. Environ Earth Sci. 83(2) 60. doi: 10.1007/s12665-023-11324-0.

Mehravar S, Razavi-Termeh SV, Moghimi A, Ranjgar B, Foroughnia F, Amani M. 2023. Flood susceptibility mapping using multi-temporal SAR imagery and novel integration of nature-inspired algorithms into support vector regression. J Hydrol. 617:129100. doi: 10.1016/j.jhydrol.2023.129100.

Meliho M, Khattabi A, Zejli D, Orlando CA. 2022. Spatial prediction of flood-susceptible zones in the Ourika watershed of Morocco using machine learning algorithms. ACI. doi: 10.1108/ACI-09-2021-0264.

Merghadi A, Yunus AP, Dou J, Whiteley J, ThaiPham B, Bui DT, Avtar R, Abderrahmane B. 2020. Machine learning methods for landslide susceptibility studies: a comparative overview of algorithm performance. Earth Sci Rev. 207:103225. doi: 10.1016/j.earscirev.2020.103225.

Mia MU, Rahman M, Elbeltagi A, Abdullah-Al-Mahbub M, Sharma G, Islam HMT, Pal SC, Costache R, Islam ARMT, Islam MM, et al. 2022. Sustainable flood risk assessment using deep learning-based algorithms with a blockchain technology. Geocarto Int. 38(1):1–29. doi: 10.1080/10106049.2022.2112982.

Mishra K, Sinha R. 2020. Flood risk assessment in the Kosi megafan using multi-criteria decision analysis: a hydro-geomorphic approach. Geomorphology. 350:106861. doi: 10.1016/j.geomorph.2019.106861.

Mitra R, Das J. 2022. A comparative assessment of flood susceptibility modelling of GIS-based TOPSIS, VIKOR, and EDAS techniques in the Sub-Himalayan foothills region of Eastern India. Environ Sci Pollut Res. 30(6):16036–16067. doi: 10.1007/s11356-022-23168-5.

Mitra R, Saha P, Das J. 2022. Assessment of the performance of GIS-based analytical hierarchical process (AHP) approach for flood modelling in Uttar Dinajpur district of West Bengal, India. Geomatics Nat Hazards Risk. 13(1):2183–2226. doi: 10.1080/19475705.2022.2112094.

Mousavi SM, Ataie-Ashtiani B, Hosseini SM. 2022. Comparison of statistical and MCDM approaches for flood susceptibility mapping in northern Iran. J Hydrol. 612:128072. doi: 10.1016/j.jhydrol.2022.128072.

Nguyen HD, Nguyen Q, Bui Q. 2024. Solving the spatial extrapolation problem in flood susceptibility using hybrid machine learning, remote sensing, and GIS. Environ Sci Pollut Res. 31(12):18701–18722. doi: 10.1007/s11356-024-32163-x.

[NIH] National Institute of Hydrology. 2000. Dam break study of Gogahoda project, Orissa. Technical report, CS-AR-15/96-97. Roorkee (India): National Institute of Hydrology, Jal Vigyan Bhawan.

Nohani N, Moharrami N, Sharafi N, Khosravi N, Pradhan N, Pham N, Lee N, Melesse N. 2019. Landslide susceptibility mapping using different GIS-Based bivariate models. Water. 11(7):1402. doi: 10.3390/w11071402.

Parida BR, Pandey AC, Kumar S, Tripathi G. 2022. Comparative flood area analysis based on change detection and binarization methods using Sentinel-1 synthetic aperture radar data. In: Radar remote sensing. Elsevier; p. 93–108.

Paul P, Sarkar R. 2022. Flood susceptible surface detection using geospatial multi-criteria framework for management practices. Nat Hazards. 114(3):3015–3041. doi: 10.1007/s11069-022-05503-8.

Pelich R, Chini M, Hostache R, Matgen P, Pulvirenti L, Pierdicca N. 2022. Mapping floods in urban areas from dual-polarization InSAR coherence data. IEEE Geosci Remote Sens Lett. 19:1–5. doi: 10.1109/LGRS.2021.3110132.

Penki R, Basina SS, Tanniru SR. 2022. Application of geographical information system-based analytical hierarchy process modeling for flood susceptibility mapping of Krishna District in Andhra Pradesh. Environ Sci Pollut Res. 30(44):99062–99075. doi: 10.1007/s11356-022-22924-x.

Pradhan B, Lee S, Dikshit A, Kim H. 2023. Spatial flood susceptibility mapping using an explainable artificial intelligence (XAI) model. Geosci Front. 14(6):101625. doi: 10.1016/j.gsf.2023.101625.

Ramayanti S, Nur AS, Syifa M, Panahi M, Achmad AR, Park S, Lee C. 2022. Performance comparison of two deep learning models for flood susceptibility map in Beira area, Mozambique. Egypt J Remote Sens Space Sci. 25(4):1025–1036. doi: 10.1016/j.ejrs.2022.11.003.

Ren H, Pang B, Bai P, Zhao G, Liu S, Liu Y, Li M. 2024. Flood susceptibility assessment with random sampling strategy in ensemble learning (RF and XGBoost). Remote Sens. 16(2):320. doi: 10.3390/rs16020320.

Risling A, Lindersson S, Brandimarte L. 2024. A comparison of global flood models using Sentinel-1 and a change detection approach. Nat Hazards. 120(12):11133–11152. doi: 10.1007/s11069-024-06629-7.

Saber M, Boulmaiz T, Guermoui M, Abdrabo KI, Kantoush SA, Sumi T, Boutaghane H, Hori T, Binh DV, Nguyen BQ, et al. 2023. Enhancing flood risk assessment through integration of ensemble learning approaches and physical-based hydrological modeling. Geomatics Nat Hazards Risk. 14(1). doi: 10.1080/19475705.2023.2203798.

Saha A, Pal SC, Arabameri A, Blaschke T, Panahi S, Chowdhuri I, Chakrabortty R, Costache R, Arora A. 2021. Flood susceptibility assessment using novel ensemble of hyperpipes and support vector regression algorithms. Water. 13(2):241. doi: 10.3390/w13020241.

Saha S, Gayen A, Bayen B. 2022. Deep learning algorithms to develop flood susceptibility map in Data-Scarce and Ungauged River Basin in India. Stoch Environ Res Risk Assess. 36(10):3295–3310. doi: 10.1007/s00477-022-02195-1.

Saha UD, Alam MJ, Bhattacharya S, Majumder A. 2023. Nature of flood and channel sedimentation in the Torsa River: a hydro-geomorphic study. In: Springer geography. p. 37–61. doi: 10.1007/978-3-031-21086-0_3.

Saikh NI, Mondal P. 2023. Gis-based machine learning algorithm for flood susceptibility analysis in the Pagla river basin, Eastern India. Nat Hazards Res. 3(3):420–436. doi: 10.1016/j.nhres.2023.05.004.

Saini O, Bhardwaj A, Chatterjee RS. 2020. Detection of water body using very high-resolution UAV SAR and Sentinel-2 images. In: Proceedings of UASG 2019: Unmanned Aerial System in Geomatics 1. Springer International Publishing, p. 53–65.

Saravanan S, Abijith D. 2022. Flood susceptibility mapping of northeast coastal districts of Tamil Nadu India using multi-source geospatial data and machine learning techniques. Geocarto Int. 37(27):15252–15281. doi: 10.1080/10106049.2022.2096702.

Sarkar D, Saha S, Sarkar T, Mondal P. 2023. Delineation of flood susceptibility zone using analytical hierarchy process and frequency ratio methods: a case study of Dakshin Dinajpur District, India. J Indian Soc Remote Sens. 51(12):2447–2465. doi: 10.1007/s12524-023-01777-y.

Satarzadeh E, Sarraf A, Hajikandi H, Sadeghian MS. 2021. Flood hazard mapping in western Iran: assessment of deep learning vis-à-vis machine learning models. Nat Hazards. 111(2):1355–1373. doi: 10.1007/s11069-021-05098-6.

Sellami EM, Maanan M, Rhinane H. 2022. Performance of machine learning algorithms for mapping and forecasting of flash flood susceptibility in Tetouan, Morocco. Int Arch Photogramm Remote Sens Spatial Inf Sci. XLVI-4/W3-2021:305–313. doi: 10.5194/isprs-archives-XLVI-4-W3-2021-305-2022.

Shah SA, Ai S. 2024. Flood susceptibility mapping contributes to disaster risk reduction: a case study in Sindh, Pakistan. Int J Disaster Risk Reduct. 108:104503. doi: 10.1016/j.ijdrr.2024.104503.

Sutradhar S, Mondal P. 2023. Prioritization of watersheds based on morphometric assessment in relation to flood management: a case study of Ajay river basin, Eastern India. Watershed Ecol Environ. 5:1–11. doi: 10.1016/j.wsee.2022.11.011.

Tanim AH, McRae CB, Tavakol-Davani H, Goharian E. 2022. Flood detection in urban areas using satellite imagery and machine learning. Water. 14(7):1140. doi: 10.3390/w14071140.

Tavus B, Can R, Kocaman S. 2022. A CNN-based flood mapping approach using sentinel-1 data. ISPRS Ann Photogramm Remote Sens Spatial Inf Sci. V-3-2022:549–556. doi: 10.5194/isprs-annals-V-3-2022-549-2022.

Tavus B, Kocaman S, Gokceoglu C. 2022. Flood damage assessment with Sentinel-1 and Sentinel-2 data after Sardoba dam break with GLCM features and random forest method. Sci Total Environ. 816:151585. doi: 10.1016/j.scitotenv.2021.151585.

Wahba M, Sharaan M, Elsadek WM, Kanae S, Hassan HS. 2024. Examination of the efficacy of machine learning approaches in the generation of flood susceptibility maps. Environ Earth Sci. 83(14). doi: 10.1007/s12665-024-11696-x.

Wang F, Ahmad I, Zelenakova M, Fenta A, Dar MA, Teka AH, Belew AZ, Damtie M, Berhan M, Shafi SN. 2023. Exploratory regression modeling for flood susceptibility mapping in the GIS environment. Sci Rep. 13(1). doi: 10.1038/s41598-023-27447-0.

Yommy AS, Liu R, Wu S. 2015. SAR image despeckling using refined Lee filter. In: 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics. Vol. 2. IEEE, p. 260–265.

Youssef AM, Mahdi A, Pourghasemi HR. 2022. Optimal flood susceptibility model based on performance comparisons of LR, EGB, and RF algorithms. Nat Hazards. 115(2):1071–1096. doi: 10.1007/s11069-022-05584-5.

Youssef AM, Mahdi A, Al-Katheri MM, Pouyan S, Pourghasemi HR. 2023. Multi-hazards (landslides, floods, and gully erosion) modeling and mapping using machine learning algorithms. J Afr Earth Sci. 197:104788. doi: 10.1016/j.jafrearsci.2022.104788.

Youssef AM, Pourghasemi HR, El-Haddad BA. 2022. Advanced machine learning algorithms for flood susceptibility modeling—performance comparison: Red Sea, Egypt. Environ Sci Pollut Res Int. 29(44):66768–66792. doi: 10.1007/s11356-022-20213-1.

Youssef AM, Pradhan B, Dikshit A, Mahdi A. 2022. Comparative study of convolutional neural network (CNN) and support vector machine (SVM) for flood susceptibility mapping: a case study at Ras Gharib, Red Sea, Egypt. Geocarto Int. 37(26):11088–11115. doi: 10.1080/10106049.2022.2046866.

Zhao F, Wang T, Zhang L, Feng H, Yan S, Fan H, Xu D, Wang Y. 2022. Polarimetric persistent scatterer interferometry for ground deformation monitoring with VV-VH Sentinel-1 data. Remote Sens. 14(2):309. doi: 10.3390/rs14020309.

Zhao J, Pelich R, Hostache R, Matgen P, Cao S, Wagner W, Chini M. 2021. Deriving exclusion maps from C-band SAR time-series in support of floodwater mapping. Remote Sens Environ. 265:112668. doi: 10.1016/j.rse.2021.112668.

Zhu X, Guo H, Huang JJ. 2024. Urban flood susceptibility mapping using remote sensing, social sensing and an ensemble machine learning model. Sustain Cities Soc. 108:105508. doi: 10.1016/j.scs.2024.105508.