Journal of Geospatial Science and Analytics 2(2): 159-186 (2026)

Journal of Geospatial Science and Analytics

Journal homepage:
https://ejournal.foundae.com/index.php/jgsa/index
e-ISSN: 3090-8086 | p-ISSN: 3090-532X

Comparative assessment of flood risk using AHP and machine learning models in Kurigram district, Bangladesh

Pranta Bhowmik*
RWTH Aachen University,
GERMANY

Md. Aminul Haque Laskor
Bangladesh University of Engineering and Technology,
BANGLADESH

Ananta Bhowmik
Independent University,
BANGLADESH

Md. Mohaimenul Islam
Shahjalal University of Science and Technology,
BANGLADESH

Article Info

Article history:

Received: May 07, 2026
Revised: June 18, 2026
Accepted: July 20, 2026

Keywords:

Analytic Hierarchy Process
Flood Risk Index
Geographic Information System
Kurigram District
Machine Learning

DOI:
10.58524/jgsa.v2i2.139

Abstract

Flood risk assessment requires an integrated evaluation of physical flood hazards and socio-economic vulnerability, particularly in highly flood-prone regions such as northern Bangladesh. This study presents a GIS-based flood risk assessment framework integrating the Analytic Hierarchy Process (AHP) and machine learning models for spatial flood risk mapping in Kurigram District, Bangladesh. The novelty of the study lies in combining environmental flood-hazard factors and socio-economic vulnerability indicators within a hybrid GIS–AHP–ML framework for comprehensive flood risk assessment. A Flood Hazard Index (FHI) was developed using ten physical factors, while socio-economic vulnerability was assessed through a Flood Vulnerability Index (FVI). The AHP-derived weights produced acceptable consistency ratios of 0.08 for FHI and 0.014 for FVI, confirming the reliability of factor prioritization. The Flood Risk Index (FRI) was generated by integrating FHI and FVI within a GIS environment and classified into five risk categories. Results indicate that flood risk in Kurigram District is influenced more strongly by socio-economic vulnerability and limited adaptive capacity than by physical hazard intensity alone. Among the applied machine learning models, Random Forest achieved the highest predictive performance and identified 173.28 km² of Nageshwari Upazila as high-risk. Model validation using Random Forest, Support Vector Machine, and Decision Tree produced AUC values of 0.93, 0.91, and 0.86, respectively, compared to 0.82 for the conventional AHP approach. The findings demonstrate that the proposed hybrid GIS–AHP–ML framework significantly improves flood risk prediction accuracy and provides a transferable, policy-relevant tool for disaster risk reduction, land-use planning, and climate resilience building in flood-prone regions.

To cite this article: Bhowmik, P., Laskor, M. A. H., Bhowmik, A., & Islam, M. M. (2026). Comparative assessment of flood risk using AHP and machine learning models in Kurigram district, Bangladesh. *Journal of Geospatial Science and Analytics,* 2(2), 159–186.

***Corresponding Author:**
Pranta Bhowmik, RWTH Aachen University, GERMANY
Email: prantabhowmik93@gmail.com

INTRODUCTION

Floods are among the most pervasive and destructive natural hazards globally, exerting profound negative impacts on human lives, infrastructure, socio-economic development, and environmental sustainability (Liu et al., 2025). Over recent decades, the intensity and frequency of flood events have increased markedly, primarily due to climate change–induced alterations in precipitation patterns, rising occurrences of extreme rainfall, and disruptions in regional hydrological processes. These climatic drivers are further compounded by rapid population growth, unplanned urbanization, deforestation, and intensified pressure on land and water resources, collectively amplifying flood risk across diverse geographic settings (Ehteram et al., 2022; Lai et al., 2022; Polong et al., 2023; Abu El Magd et al., 2020).

Bangladesh is widely recognized as one of the most flood-prone countries in the world due to its unique geographic and hydrological setting within the low-lying deltaic floodplains of the Ganges–Brahmaputra–Meghna (GBM) river system (Mirza, 2022). The country experiences recurrent flooding driven by intense monsoon rainfall, large upstream river discharge, low elevation, and limited drainage capacity. Both seasonal riverine floods and sudden flash floods occur almost every year, leading to severe socio-economic disruptions. Floods in Bangladesh routinely cause extensive damage to agricultural land, housing, transportation infrastructure, and public utilities, while also displacing millions of people and exacerbating poverty, food insecurity, and livelihood vulnerability (Haque & Zaman, 1993; Hoq et al., 2021). The magnitude of recent events further illustrates this vulnerability; during the 2024 monsoon season, floods affected more than 5.8 million people, damaged critical infrastructure, and resulted in estimated economic losses of approximately USD 1.2 billion (Kabir Uddin et al., 2025). These recurring impacts highlight the urgent need for spatially explicit, data-driven flood risk assessments to enhance disaster risk reduction and resilience planning.

Among the flood-prone districts of Bangladesh, Kurigram—located in the northern region—exhibits exceptionally high vulnerability due to its geomorphological and hydrological characteristics. The district is traversed by several major rivers, including the Brahmaputra (Jamuna), Teesta, and Dharla, and is characterized by low relief, extensive alluvial floodplains, and a dense river network (Suman & Bhattacharya, 2015). These conditions facilitate frequent overbank flooding, prolonged waterlogging, and severe riverbank erosion, which significantly disrupt agricultural productivity and rural livelihoods (Shahriar, 2021). Recent spatial assessments suggest that approximately 33.7% of Kurigram's total area falls within very high flood susceptibility zones, indicating a critical level of exposure compared to other districts in the Rangpur Division (Haque et al., 2025). Consequently, Kurigram represents a highly relevant case study for advancing flood risk assessment using integrated geospatial and analytical approaches.

Geographic Information Systems (GIS), when combined with Multi-Criteria Decision Analysis (MCDA) techniques, have emerged as powerful tools for flood hazard and risk mapping. Among MCDA methods, the Analytic Hierarchy Process (AHP) is one of the most widely applied due to its conceptual simplicity, flexibility, and ability to incorporate expert knowledge (Saaty, 1980; Hagos et al., 2022). GIS enables the integration, visualization, and spatial analysis of diverse datasets, while AHP provides a structured framework for assigning relative weights to multiple flood-conditioning factors. Numerous studies have successfully applied GIS-AHP approaches to assess flood susceptibility using key physical variables such as elevation, slope, drainage density, rainfall intensity, distance to rivers, land use/land cover (LULC), and soil characteristics (Chen et al., 2021; Ming et al., 2022; Siam et al., 2022; Kader et al., 2024). These approaches facilitate the generation of composite flood-hazard indices and the identification of spatial patterns in flood-prone areas.

Despite their widespread use, conventional AHP-based flood assessment methods have several limitations. Most notably, many studies focus predominantly on biophysical drivers of flooding while inadequately incorporating socio-economic vulnerability, which critically influences flood impacts and recovery capacity (Haque & Zaman, 1993). Additionally, AHP's reliance on expert judgment for weight assignment introduces subjectivity and may limit model robustness and predictive reliability. To overcome these constraints, recent research has increasingly integrated machine learning (ML) techniques within GIS-based flood assessment frameworks. ML algorithms—including Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), Artificial Neural Networks (ANN), and Extreme Gradient Boosting (XGBoost)—are capable of capturing complex, non-linear relationships between environmental variables and historical flood occurrences, often outperforming traditional statistical models in predictive accuracy (Seydi et al., 2023; Duy et al., 2024; Chowdhury et al., 2025; Adnan et al., 2023; Ali et al., 2025).

The integration of ML with GIS and remote sensing data further enhances flood modelling by enabling the inclusion of dynamic environmental indicators such as the Normalized Difference Vegetation Index (NDVI), Modified Normalized Difference Water Index (MNDWI), and Topographic Wetness Index (TWI) (Antzoulatos et al., 2022). These indices improve the spatial resolution and temporal sensitivity of flood hazard mapping, thereby supporting more reliable flood prediction and risk delineation (Costache et al., 2020). Model validation techniques such as Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) metrics are commonly employed to evaluate predictive performance, with higher AUC values indicating stronger discriminatory capability between flooded and non-flooded areas (Fielding & Bell, 1997; Chen et al., 2021).

In Bangladesh, several studies have applied GIS-AHP and ML-based approaches to flood risk assessment at national and regional scales. Siam et al. (2022) integrated deep neural networks with fuzzy AHP to develop a national-scale flood risk map, identifying approximately 20.45% of Bangladesh as moderate to very high risk, although district-level socio-economic heterogeneity was not explicitly addressed. Kader et al. (2024) employed a GIS-AHP framework using eight physical factors and classified 16.03% of the country as very highly susceptible to flooding, but omitted social vulnerability indicators. At the regional level, Rana et al. (2025) incorporated both physical and selected socio-economic variables within a GIS-AHP framework for Rangpur Division, achieving an ROC-AUC value of 0.83; however, machine learning models were not used for comparative validation. Similarly, Islam et al. (2025) applied AHP-GIS methods in Sylhet Division using multiple environmental parameters, but did not integrate ML techniques. Standalone ML-based studies in northeastern Bangladesh have demonstrated high predictive accuracy using ANN, RNN, RFGB, and CatBoost models, yet these analyses largely excluded socio-economic vulnerability components (Rudra & Sarkar, 2023).

Therefore, the present study proposes an integrated GIS-AHP and machine-learning framework to develop a comprehensive Flood Risk Index (FRI) for the Kurigram District, Bangladesh. The framework combines a Flood Hazard Index (FHI), derived from ten environmental variables—rainfall, distance to rivers, drainage density, elevation, slope, soil texture, TWI, LULC, MNDWI, and NDVI—with a Flood Vulnerability Index (FVI) constructed using five socio-economic indicators, namely population density, female population density, distance to roads, distance to healthcare facilities, and literacy rate. By integrating hazard intensity with socio-economic vulnerability, this study provides a multidimensional assessment of flood risk that identifies highly vulnerable upazilas within Kurigram District. The findings are expected to provide robust scientific support for disaster risk reduction planning, resource prioritization, and resilience-building initiatives in one of Bangladesh's most flood-affected regions, while also offering a transferable methodological framework for flood-prone areas worldwide.

METHODS

Study Area

Kurigram is a district in the Rangpur Division of northern Bangladesh, situated along the country's border with India. It covers an area of 2245.04 sq km, located between 25°23' and 26°14' north latitude and between 89°27' and 89°54' east longitude. The West Bengal state of India bounds it on the north and west, Gaibandha and Jamalpur districts on the south, Assam and Meghalaya states of India on the east, as well as Rangpur and Lalmonirhat districts on the west (BNP, 2018).The district consists of 9 upazilas, 72 unions, and 1,872 villages (BBS, 2022). Several rivers flow through the centre of Kurigram. The major rivers are the Brahmaputra (now called the Jamuna), Dharla, and Teesta. Minor rivers include the Dudhkumar, Phulkuar, Gangadhar, Jinjiram, and others (ASB, 2003).

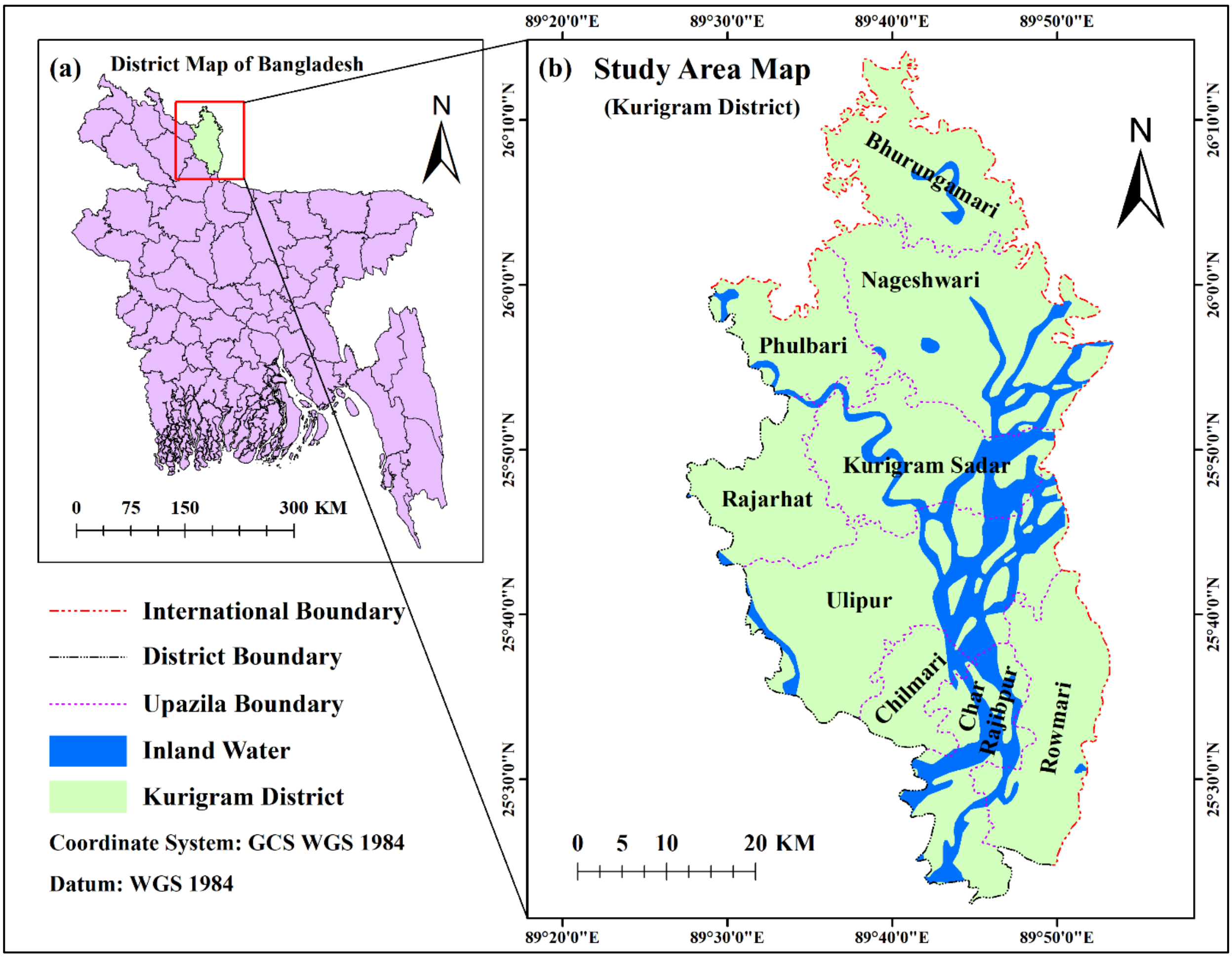

Figure 1. Geographic Location of the Study Area: (a) Bangladesh and (b) Kurigram District

The average maximum temperature is around 32–33 degrees Celsius, and the average minimum temperature is around 5–10 degrees Celsius. Like other parts of Bangladesh, the district experiences heavy rainfall during the rainy season, with an average annual rainfall of about 3,000 millimetres (BNP, 2018). The population density was 1,037 people per km². Due to the lower literacy rate (65.13%) than the national average (74.80%), agriculture is the main source of income (BBS, 2022). Kurigram experiences annual flooding primarily due to its location along major transboundary rivers such as the Brahmaputra, Teesta, Dharla, and Dudhkumar, combined with heavy monsoon rainfall, low-lying topography, and limited flood management infrastructure, a situation further exacerbated by climate change (Alam et al., 2022; Hossain et al., 2021; Rahman et al., 2023). In 2024, the Brahmaputra River at Chilmari rose 78 cm above the danger level, submerging new areas and affecting over 150,000 residents across 42 unions in eight upazilas (The Daily Star, 2024). The floods disrupted daily life, with 37 government primary schools and numerous secondary schools and madrassas suspending classes. At the same time, agricultural lands were also inundated, highlighting the district's ongoing vulnerability and the need for effective flood management (Dhaka Tribune, 2024). The geographic location and administrative boundary of the study area are presented in Figure 1.

Research Framework

The flood risk assessment of Kurigram District, Bangladesh, was conducted using a GIS-based multi-criteria approach that integrated flood hazard factors and flood vulnerability factors. This methodology allows for spatially explicit mapping of flood risk across the district, supporting disaster management and planning. The methodological framework adopted for flood risk assessment in this study is illustrated in Figure 2.

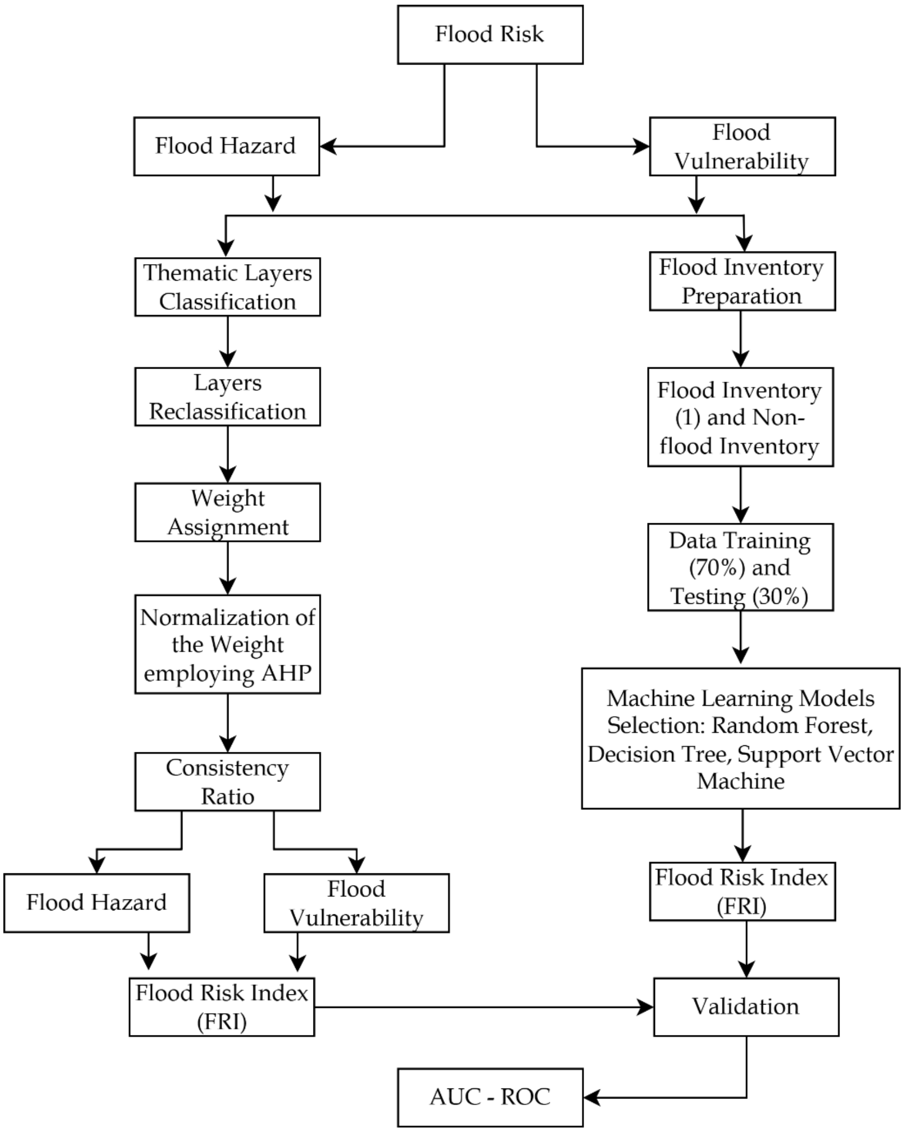

Figure 2. Integrated Research Framework for Flood Risk Assessment in Kurigram District

Data Collection

In this study, various types of data from different sources were utilized to generate spatial layers of different factors using geospatial tools. The study comprises two main factors: the Flood Hazard factor and the Flood Vulnerability factor, to assess the Flood Risk Map of the area. Table 1 presents the type of data acquired, sources, date of acquisition, and the output layers used in this study.

Table 1. Data Types and Sources Used in the Study

Data	Date	Source	Output
Digital Elevation Model (DEM)	23/09/2014	USGS Earth Explorer (SRTM, 1-arc second)	Elevation, Slope, Drainage Density, TWI Map
Land Use Land Cover (LULC)	08/09/2024	USGS Earth Explorer (Landsat 8-9 OLI/TIRS, 30 m resolution)	LULC, MNDWI, NDVI Map
Soil	29/06/2006	Bangladesh Agricultural Research Council (BARC)	Soil Texture Map
Rainfall	2020 - 2024	CHRS Data Portal	Rainfall Map
Distance to River	15/08/2018	Local Government Engineering Department (LGDE), Bangladesh	Distance to River Map
Population Density, Female Density, Literacy Rate	2022	BBS 2022	Population Density, Female Density, Literacy Rate Map
Distance to Road	08/02/2025	OpenStreetMap	Distance to Road Map
Distance to Hospital	12/01/2025	OpenStreetMap	Distance to Hospital Map

Data Preparation

To obtain data on flood hazard factors such as elevation, slope, drainage density, and TWI, the DEM file was analyzed using hydrological and spatial analyst tools downloaded from USGS Earth Explorer. The DEM was first corrected with a fill operation to represent the elevation for this study. Slope was calculated using the slope tool in ArcMap 10. 8. Drainage density was created by applying flow direction and flow accumulation tools, and streamlines were generated with the raster calculator. These streamlines were compared with the base map and reshaped before calculating the drainage density for the study area. The TWI map was produced in ArcGIS 10. 8 using the equation proposed by Beven and Kirkby (1979), where catchment area (As) and slope (β) were derived from DEM data and used in the formula TWI = ln (As /tan β) to depict the potential water accumulation across the region.

The land-use land-cover map for the study area was generated from Landsat 8-9 OLI/TIRS satellite imagery at 30 m resolution in 2024. Supervised classification in ArcGIS 10.8 grouped land cover into five classes. The Normalized Difference Vegetation Index (NDVI) and Modified Normalized Difference Water Index (MNDWI) were calculated from Landsat imagery and categorized into five classes. NDVI was computed as NDVI = (Band 5 − Band 4) / (Band 5 + Band 4), while MNDWI was calculated as MNDWI = (Band 3 − Band 6) / (Band 3 + Band 6). Rainfall data, representing the peak rainy season (June- August), were used to assess rainfall, since most floods occur during this period in Kurigram District, Bangladesh. Monthly average rainfall was spatially interpolated using the Inverse Distance Weighting (IDW) method to create a rainfall map.

The soil texture map was obtained from the Bangladesh Agricultural Research Council (BARC). The distance to the river was generated using the Euclidean distance tool, and additional data were collected from the Local Government Engineering Department (LGED), Bangladesh. Five flood vulnerability factors- such as population density, female population density, literacy rate, distance to road, and distance to hospital- were tabulated for the nine upazilas of Kurigram district from the Bangladesh Bureau of Statistics and OpenStreetMap. Raster layers for each socio-economic factor were generated and reclassified into five classes according to their relative influence on flood vulnerability. Although the datasets used in this study were collected from different years due to data availability constraints, they represent the most reliable regional-scale information available for flood risk assessment in the study area. Finally, the weighted flood hazard and flood vulnerability layers were integrated through raster overlay analysis to generate the final Flood Risk Index (FRI) map (Hu et al., 2017; Luu & Von Meding, 2018).

Weight Assignment Using Analytic Hierarchy Process (AHP)

Analytical Hierarchy Process (AHP) (Saaty, 1977) is one of the successfully implemented multi-criterion decision methods (MCDM) for the assessment of probability, vulnerability, and risk of natural hazards (Sinha et al., 2008). The AHP technique utilizes relative priorities assigned to each criterion in achieving a specific objective. The main steps in the AHP modelling process are the construction of the matrix for pairwise comparisons, the normalization of the table derived from the matrix, the computation of the eigenvector, the calculation of the weighting coefficient, and the determination of the consistency ratio. Firstly, class-wise ranking schemes and corresponding weights were assigned for both the Flood Hazard Index (FHI) (Table 2) and the Flood Vulnerability Index (FVI) (Table 5). Subsequently, pairwise comparison matrices were constructed for the FHI (Table 3) and FVI (Table 6) factors based on their relative importance. The indicator scores obtained from the pairwise matrices were then normalized to generate normalized comparison tables (Tables 4 and 7). Finally, the average value of each row in the normalized matrices was computed to derive the final weight for each factor.

Flood Hazard Index (FHI)

Table 2. AHP-Based Ranking and Weights of Flood Hazard Factors for FHI Mapping

Flood Hazard Factors	Classification Range	Weight	Influence (%)	Rank
Rainfall (mm)	121.28 – 127.34	0.266	26.6	1
	127.34 – 133.41			2
	133.41 – 139.48			3
	139.48 – 145.55			4
	145.55 – 151.62			5
Distance to River (km)	0 – 1.45	0.187	18.7	5
	1.45 – 2.90			4
	2.90 – 4.36			3
	4.36 – 5.81			2
	5.81 – 7.26			1
Drainage Density (km/sq.km)	0 – 31.24	0.136	13.6	1
	31.24 – 62.49			2
	62.49 – 93.73			3
	93.73 – 124.98			4
	124.98 – 156.23			5
Elevation (m)	0 – 21	0.099	9.9	5
	21 – 26			4
	26 – 30			3
	30 – 33			2
	33 – 101			1
Slope (degree)	0 – 1.00	0.074	7.4	5
	1.00 – 2.21			3
	2.21 – 4.02			5
	4.02 – 7.03			2
	7.03 – 51.26			1
Soil Texture	Clayey Soils	0.057	5.7	5
	Loamy Soils			2
	Sandy Soils			1
	Silt Loam Soils			3
	Silty Clay Loam Soils			4
TWI	-8.40 – -.395	0.046	4.6	1
	-3.95 – -2.19			2
	-2.19 – 0.24			3
	0.24 – 3.68			4
	3.68 – 13.01			5

Flood Hazard Factors	Classification Range	Weight	Influence (%)	Rank
LULC	Waterbody	0.041	4.1	4
	Settlement			5
	Vegetation			1
	Agricultural Land			3
	Barren Land			2
MNDWI	-0.43 – -0.15	0.047	4.7	1
	-0.15 – -0.12			2
	-0.12 – -0.04			3
	-0.04 – 0.07			4
	0.07 – 0.18			5
NDVI	-0.10 – 0.02	0.047	4.7	5
	0.02 – 0.13			4
	0.13 – 0.21			3
	0.21 – 0.27			2
	0.27 – 0.4			1

Table 3. AHP-Based Pairwise Comparison Matrix of Flood Hazard Factors

Factors	RF	DTR	DD	EL	SL	ST	TWI	LULC	MNDWI	NDVI
RF	1	2	3	4	5	6	7	7	3	3
DTR	0.50	1	2	3	4	5	6	6	2	2
DD	0.33	0.50	1	2	3	4	5	5	2	2
EL	0.25	0.33	0.50	1	2	3	4	4	2	2
SL	0.20	0.25	0.33	0.50	1	2	3	3	2	2
ST	0.17	0.20	0.25	0.33	0.50	1	2	2	2	2
TWI	0.14	0.17	0.20	0.25	0.33	0.50	1	2	2	2
LULC	0.14	0.17	0.20	0.25	0.33	0.50	0.50	1	2	2
MNDWI	0.33	0.50	0.50	0.50	0.50	0.50	0.50	0.50	1	1
NDVI	0.33	0.50	0.50	0.50	0.50	0.50	0.50	0.50	1	1
Total	3.39	5.62	8.48	12.33	17.16	23	29.5	31	19	19

Table 4. Normalized AHP Matrix and Derived Weights for Flood Hazard Factors

Factors	RF	DTR	DD	EL	SL	ST	TWI	LULC	MNDWI	NDVI	W
RF	0.295	0.356	0.354	0.324	0.293	0.261	0.237	0.226	0.159	0.159	0.266
DTR	0.147	0.178	0.236	0.243	0.233	0.217	0.203	0.194	0.105	0.105	0.187
DD	0.097	0.089	0.117	0.162	0.175	0.174	0.169	0.162	0.105	0.105	0.136
EL	0.074	0.059	0.059	0.081	0.116	0.13	0.136	0.129	0.105	0.105	0.099
SL	0.059	0.045	0.039	0.041	0.058	0.087	0.102	0.097	0.105	0.105	0.074
ST	0.05	0.035	0.029	0.027	0.029	0.043	0.068	0.064	0.105	0.105	0.057
TWI	0.042	0.03	0.024	0.02	0.019	0.022	0.034	0.064	0.105	0.105	0.046
LULC	0.042	0.03	0.024	0.02	0.019	0.022	0.017	0.032	0.105	0.105	0.041
MNDWI	0.097	0.089	0.059	0.041	0.029	0.022	0.017	0.016	0.053	0.053	0.047
NDVI	0.097	0.089	0.059	0.041	0.029	0.022	0.017	0.016	0.053	0.053	0.047
Total	1	1	1	1	1	1	1	1	1	1	1

Flood Vulnerability Index (FVI)

Table 5. AHP-Based Ranking and Weights of Flood Vulnerability Factors for FVI Mapping

Flood Vulnerability Factors	Classification Range	Weight	Influence (%)	Rank
Population Density (Pop./Sq.km)	585 – 706	0.417	41.7	1
	706 – 873			2
	873 – 1079			3
	1079 – 1222			4
	1222 – 1307			5
Female Density (Female Pop./Sq.km)	293 – 357	0.262	26.2	1
	357 – 443			2
	443 – 569			3
	569 – 612			4
	612 – 662			5
Distance to Road (km)	0 – 4.02	0.161	16.1	5
	4.02 – 8.05			4
	8.05 – 12.08			3
	12.08 – 16.10			2
	16.10 – 20.13			1
Distance to Hospital (km)	0 – 12.85	0.098	9.8	5
	12.85 – 25.71			4
	25.71 – 38.57			3
	38.57 – 51.42			2
	51.42 – 64.28			1
Literacy Rate (%)	60.52 – 61.36	0.062	6.2	5
	61.36 – 63.04			4
	63.04 – 64.68			3
	64.68 – 66.37			2
	66.37 – 70.27			1

Table 6. AHP-Based Pairwise Comparison Matrix of Flood Vulnerability Factors

Factors	PD	FD	DRD	DTH	LR
PD	1	2	3	4	5
FD	0.5	1	2	3	4
DRD	0.33	0.5	1	2	3
DTH	0.25	0.33	0.5	1	2
LR	0.2	0.25	0.33	0.5	1
Total	2.28	4.08	6.83	10.5	15

Table 7. Normalized AHP Matrix and Derived Weights for Flood Vulnerability Factors

Factors	PD	FD	DRD	DTH	LR	W
PD	0.438	0.49	0.439	0.381	0.333	0.417
FD	0.219	0.245	0.293	0.286	0.267	0.262
DRD	0.145	0.123	0.147	0.19	0.2	0.161
DTH	0.11	0.081	0.073	0.095	0.133	0.098
LR	0.088	0.061	0.048	0.048	0.067	0.062
Total	1	1	1	1	1	1

Where RF = Rainfall, DTR = Distance to River, DD = Drainage Density, EL = Elevation, SL = Slope, ST = Soil Texture, TWI = Topographic Wetness Index, LULC = Land Use Land Cover, MNDWI = Modified Normalized Difference Water Index, NDVI = Normalized Difference Vegetation Index, PD = Population Density, FD = Female Density, DRD = Distance to Road, DTH = Distance to Hospital, LR = Literacy Rate, W = Weight.

Consistency Ratio (CR) Check

The consistency ratio (CR) was calculated using the following equation to check the inconsistency in assigning weights. The CR formula: CR = CI/RI, with CR = Consistency ratio, CI = Consistency index, and RI = Random index. CI was calculated by the equation: $CI = \frac{\lambda_{max} - n}{n-1}$, where λ_{max} is the largest eigenvalue and n is the number of indicators considered for the study. The AHP model will be acceptable if the CR is less than 0.1. Otherwise, the assigned relative importance needs to be adjusted, and the CR value needs to be recomputed until acceptable. A CR value of zero means the model has the perfect level of consistency.

Table 8. Consistency Statistics of Flood Hazard Index and Flood Vulnerability Index

Index	Flood Hazard	Flood Vulnerability
Number of factors (n)	10	5
λmax	11.073	5.063
CI	0.119	0.016
RI	1.49	1.12
CR	0.08	0.014
Consistency	Acceptable	Acceptable

FHI and FVI Mapping

Weighted hazard and vulnerability layers were combined to generate a vulnerability map by using the following equation (Meraj et al., 2015):

Flood Hazard Index (FHI) Map:

$$\text{Rainfall} \times 0.266 + \text{Distance to River} \times 0.187 + \text{Drainage Density} \times 0.136 + \text{Elevation} \times 0.099 + \text{Slope} \times 0.074 + \text{Soil Texture} \times 0.057 + \text{TWI} \times 0.046 + \text{LULC} \times 0.041 + \text{MNDWI} \times 0.047 + \text{NDVI} \times 0.047.$$

Flood Vulnerability Index (FVI) Map:

$$\text{Population Density} \times 0.417 + \text{Female Density} \times 0.262 + \text{Distance to Road} \times 0.161 + \text{Distance to Hospital} \times 0.098 + \text{Literacy Rate} \times 0.062.$$

Generation of Flood Risk Index (FRI) Map

Flood risk was calculated using the formula (Masood & Takeuchi, 2012; Wisner et al., 2003):

Flood Risk Index (FRI) Map:

$$\text{Flood Hazard Index (FHI)} \times \text{Flood Vulnerability Index (FVI)}.$$

The flood risk index has been calculated by multiplying the hazard index by the vulnerability index. So, flood risk depends on both factors. This means that flood risk will be high for any region or society when the physical hazard is large, and the society of that region is highly vulnerable to that hazard. If the physical hazard is large but society can handle it, the risk will be lower. So, flood risk depends on the size of the flood hazard and the region or society's vulnerability to flooding. The physical hazard and social vulnerability were multiplied using the Raster Calculator tool in ArcGIS to produce the Flood Risk Index (FRI) map. The output FRI map was then classified into five classes, namely very low risk, low risk, moderate risk, high risk, and very high risk, using the natural break classification algorithm in ArcGIS.

Machine Learning Algorithms for FRI Mapping Study Area

Flood Inventory and Training Data

Flood inventory mapping is a fundamental requirement for supervised machine learning–based flood risk assessment, as it provides the target variable necessary for model training and validation. In this study, the flood inventory dataset was prepared using Sentinel-1 Synthetic Aperture Radar (SAR) imagery, historical flood information, and ancillary satellite observations corresponding to major flood events in 2020, 2022, and 2024. Flood-affected areas were identified through visual interpretation and comparison between flood-affected and non-flooded conditions. To maintain consistency with Landsat-derived environmental variables used in this study, the final flood inventory dataset was prepared at a spatial resolution of 30 m. The flood extents from the three flood events were integrated into a single composite flood inventory layer representing the observed flood-prone areas within the study region. Areas affected during any flood event were classified as flood presence (1), while non-flood locations (0) were randomly selected from areas with no recorded flooding history. To reduce sampling bias, non-flood samples were spatially distributed across different parts of the district and different land-cover conditions. Potential uncertainties associated with SAR-based flood interpretation, including smooth water surfaces, shadow effects, and mixed pixels, were minimized through visual verification and comparison with historical flood information and ancillary satellite observations. The final flood inventory dataset was randomly divided into training and testing subsets, where 70% of the samples were used for model training, and the remaining 30% were reserved for independent validation. This inventory dataset served as the dependent variable for all supervised machine learning models applied in the study.

Random Forest (RF) Algorithm

The Random Forest (RF) algorithm was applied in this study to classify flood-prone and non-flood-prone areas using the prepared flood inventory dataset and fifteen flood-conditioning factors. The model was implemented in the R programming environment using the "randomForest" package (Breiman et al., 2001). The flood inventory dataset was randomly divided into training (70%) and testing (30%) subsets to ensure independent model validation. The RF model was developed using multiple decision trees generated through bootstrap aggregation (bagging), which improves prediction stability and reduces overfitting. A total of 500 trees (ntree = 500) were used to achieve stable classification performance, while the number of predictor variables randomly selected at each node split (mtry) was automatically optimized during model training. The model incorporated all flood-conditioning variables, including rainfall, distance to rivers, drainage density, elevation, slope, soil texture, TWI, LULC, MNDWI, NDVI, population density, female population density, distance to roads, distance to healthcare facilities, and literacy rate. The RF algorithm was selected because of its strong capability to model complex nonlinear relationships between environmental variables and flood occurrence, while also handling noisy and high-dimensional datasets effectively (Yariyan et al., 2020; Zanotti et al., 2019). Variable importance analysis was additionally used to evaluate the relative contribution of flood-conditioning factors to flood susceptibility prediction.

Decision Tree (DT) Algorithm

The Decision Tree (DT) algorithm was utilized to classify flood and non-flood areas based on the prepared flood inventory dataset and selected flood-conditioning variables. The model was implemented in the R environment using the "rpart" package. Similar to the RF model, the inventory dataset was randomly divided into 70% training data and 30% testing data for model development and independent validation. The DT model recursively partitions the predictor variables into homogeneous groups using binary splitting rules to maximize class purity and improve classification accuracy (Tsangaratos et al., 2016). To reduce model overfitting and improve generalization capability, pruning techniques were applied during model optimization. The model automatically determined the optimal tree depth and node splitting criteria during training. All environmental and socio-economic flood-conditioning variables were incorporated into the model to evaluate their combined influence on flood occurrence. The DT algorithm was selected because of its interpretability, computational efficiency, and ability to identify hierarchical relationships among flood-conditioning factors in spatial flood modelling studies (Can et al., 2023).

Support Vector Machine (SVM) Algorithm

The Support Vector Machine (SVM) algorithm was applied to classify flood susceptibility zones using the prepared flood inventory dataset and flood-conditioning factors. The model was implemented in the R programming environment using the "e1071" package. Before model training, all predictor variables were normalized to minimize scale-related bias and improve classification performance. The flood inventory dataset was randomly divided into 70% training data and 30% testing data to maintain consistency with the RF and DT models. In this study, the radial basis function (RBF) kernel was adopted because of its effectiveness in modelling nonlinear relationships between flood occurrence and environmental variables (Pham et al., 2019). The kernel parameters, including the cost parameter (C) and gamma value, were automatically optimized during model training to achieve improved predictive accuracy and generalization performance. The SVM model incorporated all physical hazards and socio-economic vulnerability variables used in the flood risk assessment framework. The SVM algorithm was selected because of its strong classification capability in high-dimensional datasets and its robustness in handling complex spatial relationships commonly observed in flood susceptibility and flood risk modelling studies (Huang & Zhao, 2018).

Validation of the Map Using AUC-ROC

Evaluating the performance of predictive models is a critical step to assess their accuracy, reliability, and practical utility. Validation results provide insights into the comparative performance of different models, enabling informed selection of the most effective approach. A widely used method for this purpose is the receiver operating characteristic (ROC) curve, commonly applied in flood prediction studies (Khosravi et al., 2018; Termeh et al., 2018; Pham et al, 2021). The ROC curve plots the false positive rate (FPR, or 1 minus specificity) on the x-axis and the true positive rate (TPR, or sensitivity) on the y-axis. Specificity measures the proportion of non-flooded areas correctly identified as non-flooded, while sensitivity indicates the proportion of flooded areas correctly detected. Adjusting the classifier threshold produces multiple TPR–FPR pairs, each corresponding to a point on the ROC plot. Connecting these points forms a curve extending from (0,0) to (1,1). The area under the curve (AUC) provides a quantitative summary of overall model performance, with higher AUC values reflecting greater predictive accuracy, robustness, and reliability in flood risk assessment and mapping.

RESULTS

Spatial Distribution and Summary Statistics of Flood Hazard Index

The spatial distributions of the major flood hazard factors used in the Flood Hazard Index (FHI) assessment are presented in Figure 3, while their class-wise distributions and proportional area coverage are summarized in Table 9. Rainfall distribution indicates that low-to-moderate rainfall classes dominate large portions of the district, with the 121.28–127.34 mm rainfall class covering the largest proportion (26.79%) of the study area. In contrast, very high rainfall zones remain comparatively localized. Areas situated close to major river systems, including the Brahmaputra, Teesta, and Dharla rivers, exhibit comparatively higher flood susceptibility due to increased overflow potential during monsoon periods, while the 5.81–7.26 km river distance class occupies 36.99% of the district. Drainage density analysis reveals that low drainage concentration zones dominate the study area, accounting for 34.22%, which may contribute to localized water accumulation and prolonged inundation conditions. Similarly, moderate elevation floodplains (26–30 m) occupy 36.69% of the district, reflecting the low-lying topographic characteristics of the floodplain environment. Slope analysis further indicates that moderate-to-steep slope classes dominate most parts of the district, accounting for 38.92% of the total area, which may influence runoff concentration and flood flow dynamics across the region.

The spatial characteristics of the remaining flood hazard factors used in the Flood Hazard Index (FHI) assessment are presented in Figure 4. At the same time, their class-wise distributions and area coverage are summarized in Table 10. Soil texture analysis indicates that sandy and clayey soils occupy substantial portions of the district, accounting for 23.29% and 26.79% of the study area, respectively, reflecting spatial variations in infiltration and runoff generation potential. TWI analysis shows that low wetness zones (−8.40 to −3.95) dominate much of the district, covering 39.81% of the total area, whereas areas with very high moisture accumulation remain comparatively limited and are mainly concentrated near riverine and low-lying environments. LULC analysis reveals that barren land constitutes the dominant landscape category (42.40%), followed by agricultural land (22.52%), reflecting extensive exposed and cultivated surfaces across the floodplain environment. MNDWI results indicate that moderate surface water conditions are widespread across the district, while zones with very high surface water presence are localized around major river channels and waterlogged regions. NDVI analysis further demonstrates that low-to-moderate vegetation density classes dominate large portions of the study area, indicating varying influences on infiltration and surface runoff regulation.

Table 9. Summary Statistics of Flood Hazard Factors in Kurigram District

Flood Hazard Factors	Class Range	Area Coverage (%)	Flood Implication
Rainfall (mm)	121.28–127.34	26.79	Low rainfall zone
	127.34–133.41	17.30	Moderate rainfall zone
	133.41–139.48	23.29	Moderately high rainfall zone
	139.48–145.55	17.74	High rainfall zone
	145.55–151.62	14.88	Very high rainfall zone
Distance to River (km)	0–1.45	4.24	Very high flood exposure zone
	1.45–2.90	11.02	High flood exposure zone
	2.90–4.36	18.15	Moderate flood exposure zone
	4.36–5.81	29.60	Relatively lower flood exposure
	5.81–7.26	36.99	Low flood exposure zone
Drainage Density (km/km²)	0–31.24	34.22	Low drainage concentration
	31.24–62.49	19.89	Moderate drainage density
	62.49–93.73	23.80	Relatively dense drainage
	93.73–124.98	15.02	High drainage density
	124.98–156.23	7.07	Very high drainage density
Elevation (m)	0–21	6.72	Very low-lying floodplain
	21–26	18.35	Low elevation zone
	26–30	36.69	Moderate elevation floodplain
	30–33	27.61	Relatively elevated area
	33–101	10.63	High elevation zone
Slope (°)	0–1.00	0.82	Nearly flat terrain
	1.00–2.21	5.75	Gentle slope
	2.21–4.02	18.32	Moderate slope
	4.02–7.03	38.92	Moderate-to-steep slope
	7.03–51.26	36.19	Steep slope terrain

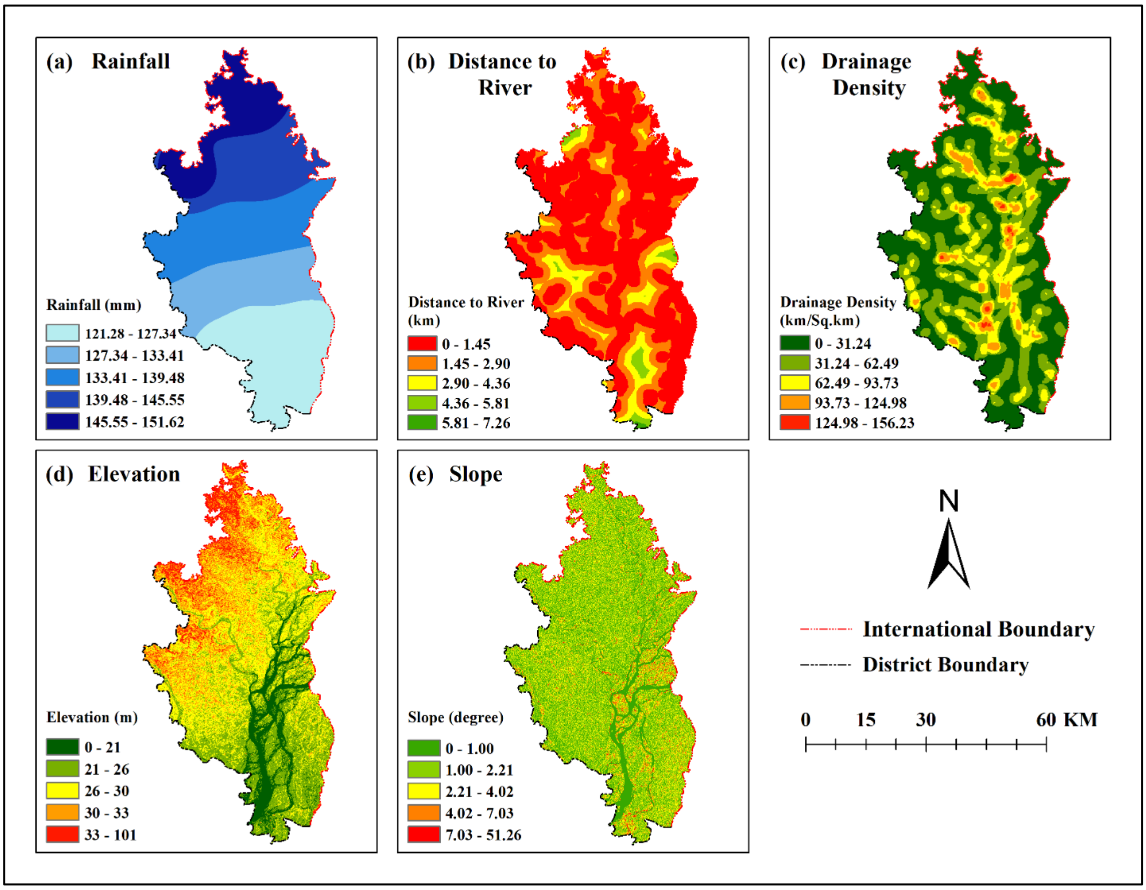

Figure 3. Flood Hazard Index (FHI) Thematic Maps showing (a) Rainfall, (b) Distance to River, (c) Drainage Density, (d) Elevation, and (e) Slope

Table 10. Summary Statistics of Flood Hazard Factors in Kurigram District

Flood Hazard Factors	Class Range	Area Coverage (%)	Flood Implication
Soil Texture	Clayey Soil	26.79	Low infiltration potential
	Loamy Soil	17.30	Moderate infiltration zone
	Sandy Soil	23.29	High infiltration capacity
	Silt Loam Soil	17.74	Moderate runoff potential
	Silty Clay Loam	14.88	High moisture retention
Topographic Wetness Index (TWI)	−8.40 to −3.95	39.81	Low moisture zone
	−3.95 to −2.19	31.61	Moderate moisture zone
	−2.19 to 0.24	17.05	Intermediate wetness zone
	0.24 to 3.68	9.80	High moisture zone
	3.68 to 13.01	1.73	Very high wetness zone
Land Use Land Cover (LULC)	Waterbody	11.29	Surface water areas
	Settlement	7.87	Built-up and urbanized
	Vegetation	15.92	Vegetation-covered areas
	Agricultural Land	22.52	Cultivated areas
	Barren Land	42.40	Sparsely vegetated land
Modified Normalized Difference Water Index (MNDWI)	−0.43 to −0.15	20.51	Low surface water presence
	−0.15 to −0.12	28.32	Moderate water presence
	−0.12 to −0.04	22.58	Moderate surface moisture
	−0.04 to 0.07	13.91	High surface moisture zone
	0.07 to 0.18	14.68	Very high surface water
Normalized Difference Vegetation Index (NDVI)	−0.10 to 0.02	22.16	Sparse vegetation cover
	0.02 to 0.13	23.02	Low vegetation density
	0.13 to 0.21	19.61	Moderate vegetation
	0.21 to 0.27	21.56	Moderately high vegetation
	0.27 to 0.45	13.65	Dense vegetation coverage

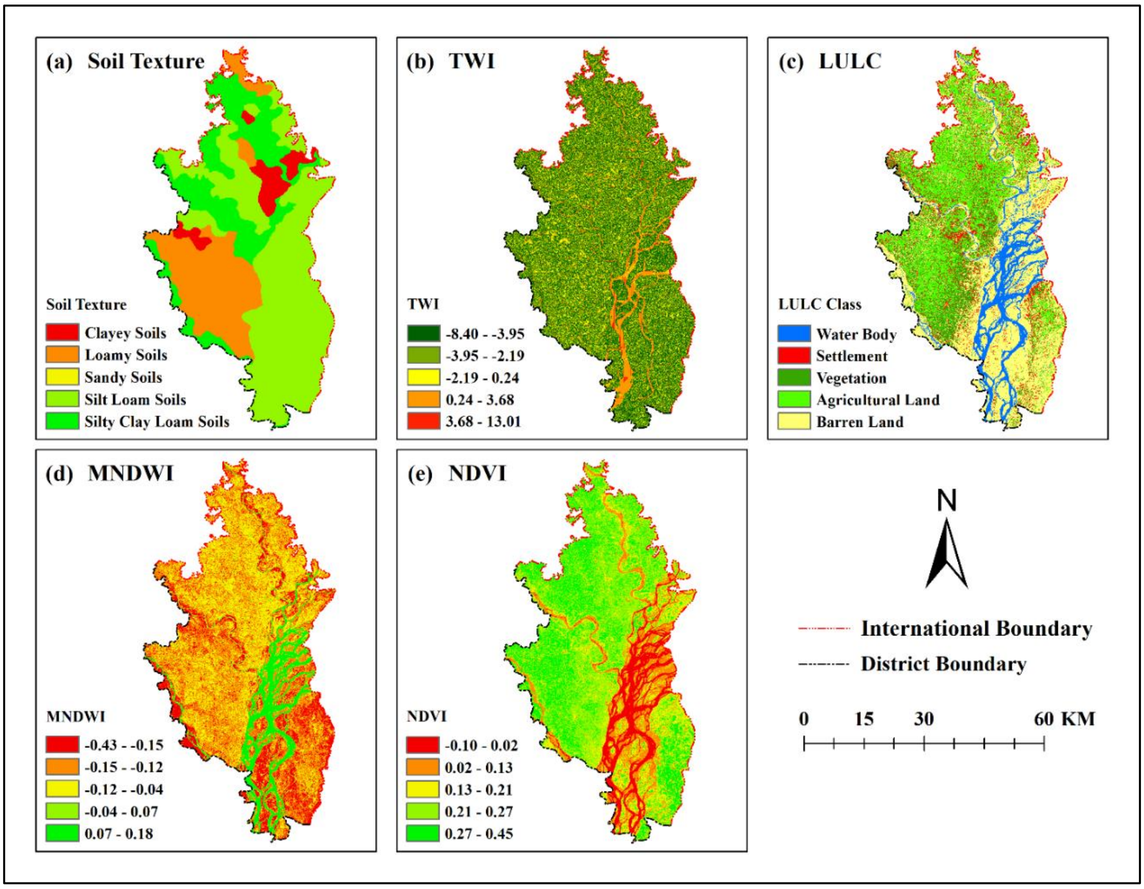

Figure 4. Flood Hazard Index (FHI) Thematic Maps showing (a) Soil Texture, (b) TWI, (c) LULC, (d) MNDWI, and (e) NDVI

Spatial Distribution and Summary Statistics of Flood Vulnerability Index

Flood vulnerability in Kurigram District is shaped by several socio-economic indicators, including population density, female population density, accessibility to transportation and healthcare facilities, and literacy conditions. The spatial distributions of these vulnerability factors are illustrated in Figure 5, while their class-wise distributions and proportional area coverage are summarized in Table 11. Population density analysis indicates that moderate-to-high population concentration zones dominate substantial portions of the district, with the 873–1079 persons/km² class covering 28.83% of the study area, increasing potential human exposure during flood events. Similarly, female population density is relatively high across several areas, where the 443–569 persons/km² class alone accounts for 35.95% of the district, reflecting spatial variations in socially sensitive populations. Accessibility-related indicators further reveal that large portions of the district are situated far from road networks and healthcare facilities, with the 16.10–20.13 km road-distance class and the 51.42–64.28 km hospital-distance class covering 31.84% and 37.67% of the study area, respectively. These conditions may substantially limit emergency response efficiency and post-disaster recovery capacity during severe flooding events. Literacy rate distribution demonstrates moderate variations in educational attainment and adaptive capacity across the district, with the 64.68–66.37% literacy class occupying the largest proportion (28.71%) of the study area. Areas characterized by lower literacy levels may experience comparatively reduced access to flood awareness information, preparedness measures, and adaptive responses. Together, these socio-economic conditions contribute significantly to the spatial variability of flood vulnerability throughout Kurigram District.

Table 11. Summary Statistics of Flood Vulnerability Factors in Kurigram District

Flood Vulnerability Factors	Class Range	Area Coverage (%)	Flood Implication
Population Density (persons/km²)	585–706	13.07	Low population exposure
	706–873	20.16	Moderate population density
	873–1079	28.83	Moderate-to-high density
	1079–1222	25.85	High population concentration
	1222–1307	12.09	Very high population exposure
Female Density (persons/km²)	293–357	13.07	Low female population density
	357–443	20.16	Moderate female population
	443–569	35.95	Dominant female population
	569–612	18.72	High female population density
	612–662	12.10	Very high female population
Distance to Road (km)	0–4.02	10.05	High accessibility zone
	4.02–8.05	12.82	Moderate accessibility
	8.05–12.08	19.74	Intermediate accessibility
	12.08–16.10	25.55	Relatively low accessibility
	16.10–20.13	31.84	Poorly accessible areas
Distance to Hospital (km)	0–12.85	9.22	High healthcare accessibility
	12.85–25.71	15.51	Moderate healthcare access
	25.71–38.57	16.44	Intermediate healthcare access
	38.57–51.42	21.16	Limited healthcare accessibility
	51.42–64.28	37.67	Poor healthcare accessibility
Literacy Rate (%)	60.52–61.36	19.89	Low adaptive capacity
	61.36–63.04	11.36	Relatively low literacy level
	63.04–64.68	20.23	Moderate literacy condition
	64.68–66.37	28.71	Dominant literacy category
	66.37–70.27	19.81	High literacy category

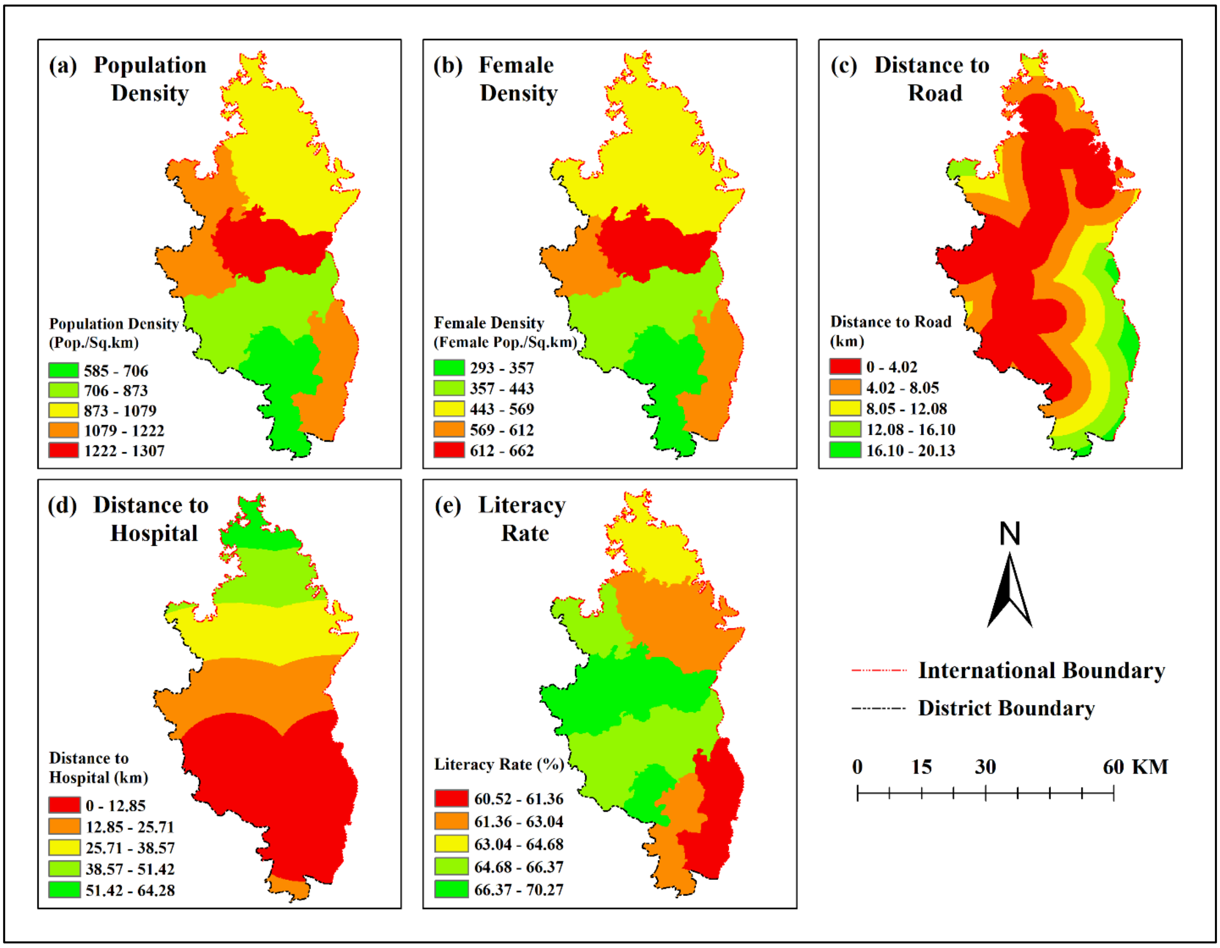

Figure 5. Flood Vulnerability Index (FVI) Thematic Maps showing (a) Population Density, (b) Female Density, (c) Distance to Road, (d) Distance to Hospital, and (e) Literacy Rate

FHI and FVI Mapping

The Flood Hazard Index (FHI) map (Figure 6a) illustrates the spatial distribution of flood hazards across Kurigram District, revealing a balanced yet spatially heterogeneous pattern. The index was derived based on ten physical hazard factors and validated through the Analytic Hierarchy Process (AHP), which yielded an acceptable consistency ratio (CR = 0.08), confirming the reliability of the weighting scheme (Saaty, 1988). Moderate (25.18%; 565.29 km²) and high (25.55%; 573.60 km²) hazard zones predominate, collectively covering over half of the district, reflecting the influence of low-lying floodplains, active fluvial systems, and seasonal monsoonal rainfall. Very low and low hazard areas occupy 13.00% (291.85 km²) and 22.38% (502.43 km²) of the district, respectively, while very high hazard zones remain spatially limited (13.89%; 311.83 km²), indicating localized but intense flood-prone environments. Among the upazilas, Nageshwari exhibits the highest proportions of high (31.41%) and very high (37.42%) hazard classes, followed by Bhurungamari (24.13% very high), highlighting strong exposure to riverine processes. Kurigram Sadar presents a mixed hazard pattern due to both river proximity and urban development, whereas Ulipur and Raomari are dominated by very low to low hazard classes (>60%). Charrajibpur and Chilmari predominantly experience low-hazard conditions, reflecting relatively stable geomorphology. These results demonstrate that flood hazard in Kurigram District is strongly controlled by riverine morphology and spatial proximity to major watercourses, producing concentrated high-hazard belts amid broader zones of moderate to low hazard.

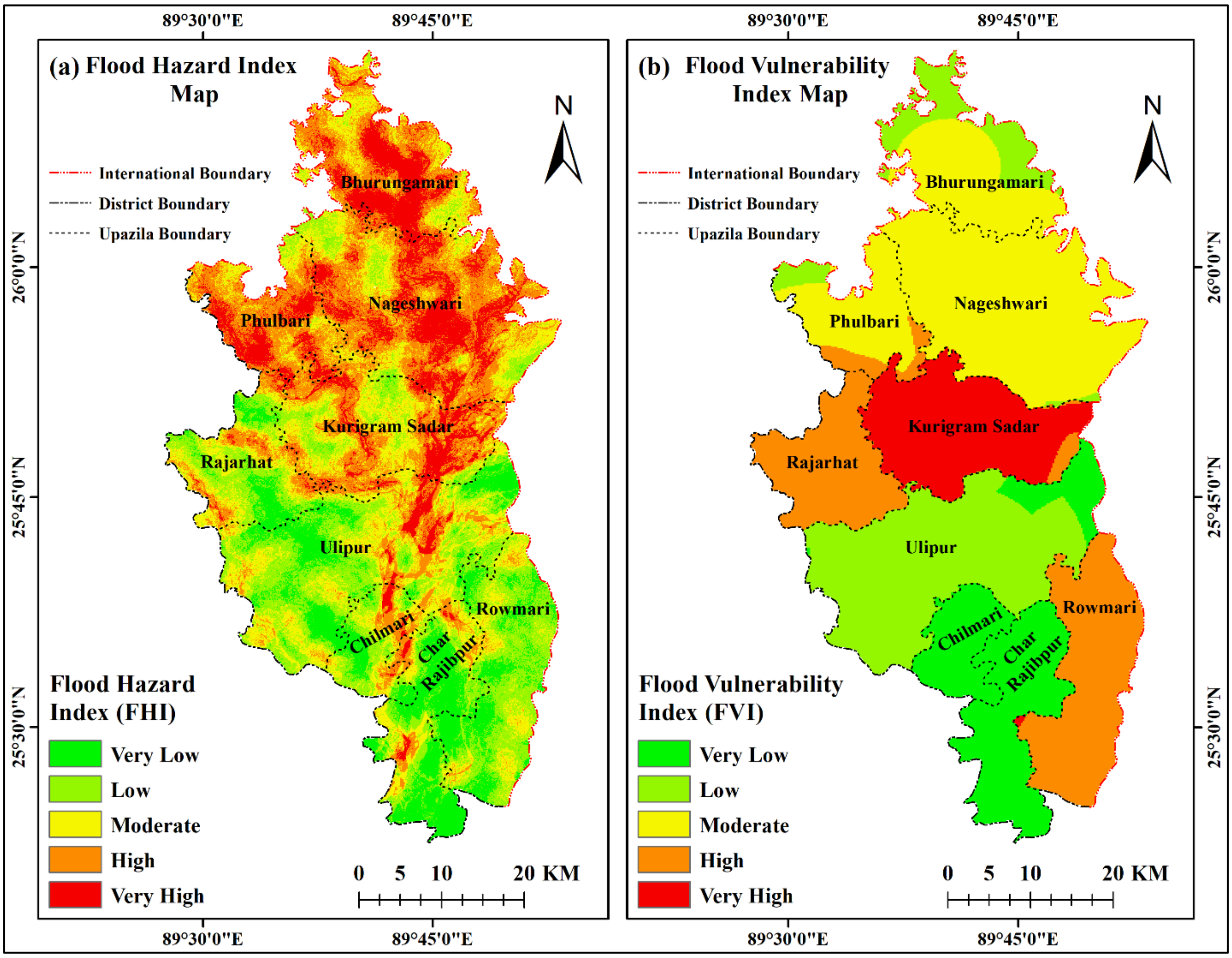

Figure 6. (a) Flood Hazard Index (FHI) Map, (b) Flood Vulnerability Index (FVI) Map

The Flood Vulnerability Index (FVI) map (Figure 6b) demonstrates that flood vulnerability in Kurigram is primarily driven by socio-economic sensitivity and adaptive capacity rather than physical hazard alone. The index was validated through AHP with a consistency ratio of 0.014. The FVI incorporates five socio-economic factors with robust weighting. Moderate vulnerability is the most widespread, covering 30.06% (674.85 km²) of the district, followed by low (22.40%; 502.88 km²) and high (20.36%; 457.08 km²) vulnerability zones. Very low and very high vulnerability areas account for 15.41% (345.95 km²) and 11.77% (264.24 km²), respectively, reflecting a generally even distribution with localized extremes. Kurigram Sadar emerges as the most critical hotspot, with 98.72% of its area classified as very high vulnerability, attributable to dense population, high infrastructure exposure, and limited coping capacity. Elevated vulnerability is also observed in Raomari (53.47%) and Rajarhat (37.70%), whereas Charrajibpur and Chilmari exhibit predominantly very low vulnerability, indicating stronger resilience. Moderate vulnerability dominates in Nageshwari (61.43%), while Ulipur remains largely low vulnerable despite moderate hazard exposure. These patterns emphasize the critical role of socio-economic resilience in shaping flood vulnerability.

Analysis of Flood Risk Index (FRI) Map

The AHP-based flood risk map (Figures 7a and 8) indicates a dominance of very low and low risk zones in Charrajibpur and Chilmari, where very low risk accounts for 63.15% (196.88 km²) and 17.01% (53.14 km²) of the respective upazila areas. Ulipur is largely classified as low risk under AHP, with 64.55% (307.05 km²) of its area falling into this category. Moderate flood risk zones are mainly observed in Raomari (28.38%, 187.50 km²) and Ulipur (22.71%, 150.15 km²).

High flood risk areas are concentrated in Nageshwari, occupying 38.62% (218.40 km²), while very high flood risk is overwhelmingly dominant in Kurigram Sadar, accounting for 57.48% (132.21 km²) of the upazila area.

The RF-based model (Figures 7b and 8) produces a comparatively conservative delineation of low-risk zones and a clearer identification of higher-risk classes. Very low flood risk remains prominent in Charrajibpur (38.98%, 187.00 km²), whereas Ulipur shows a reduced low-risk extent (49.35%, 24.13 km²) compared to AHP. Moderate risk zones are widespread in Nageshwari (26.27%, 106.37 km²) and Raomari (26.68%, 112.83 km²). High flood risk is most pronounced in Nageshwari, comprising 41.62% (173.28 km²), while very high flood risk is strongly concentrated in Kurigram Sadar, where it reaches 61.70% (38.01 km²).

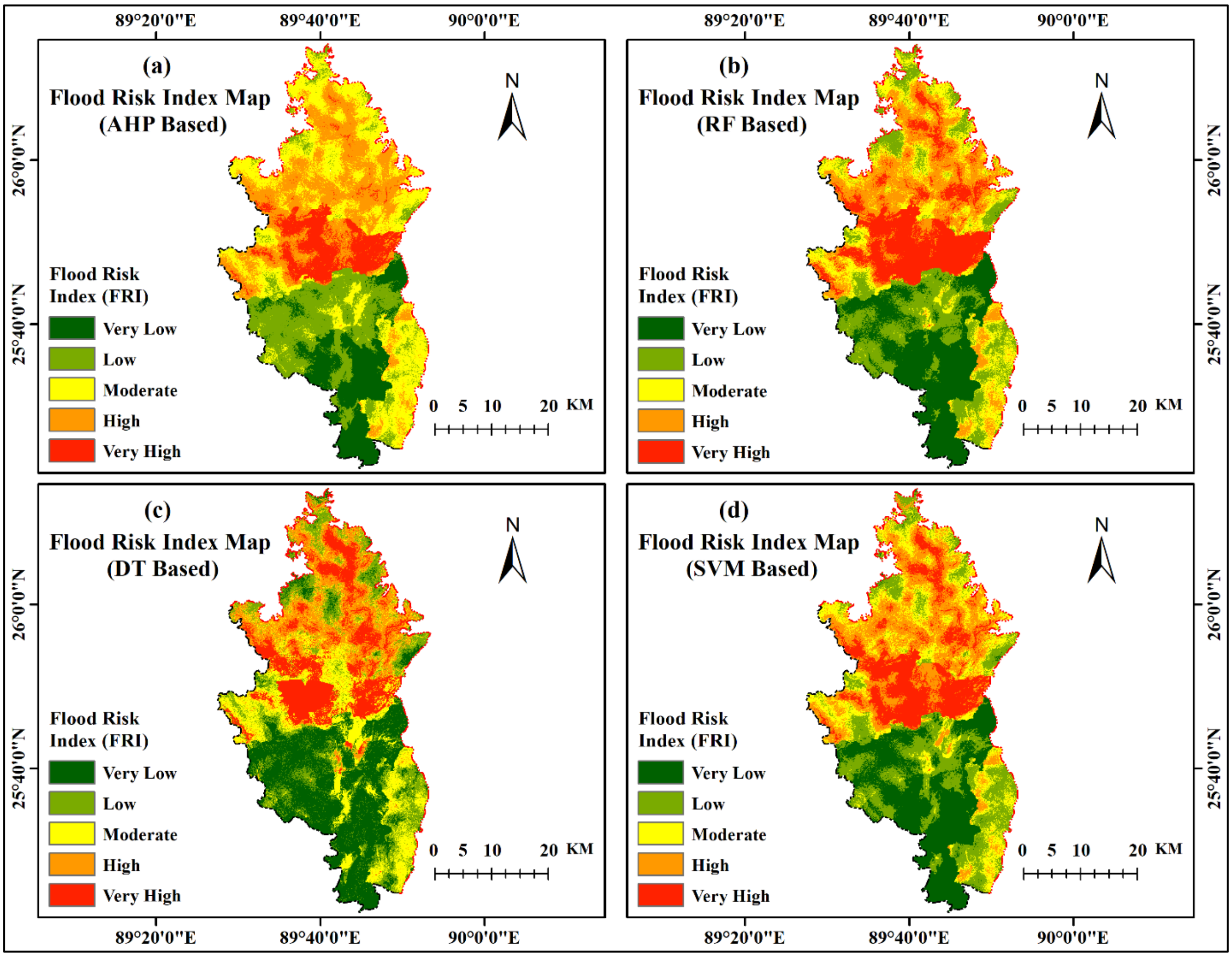

Figure 7. Flood Risk Index Map based on (a) AHP, (b) RF, (c) DT, (d) SVM

The DT-based flood risk map (Figures 7c and 8) exhibits more aggressive spatial segmentation and generally larger extents of high-risk areas. Very low flood risk is notable in Ulipur (48.23%, 309.68 km²) and Charrajibpur (25.76%, 165.85 km²). Moderate flood risk expands considerably in Kurigram Sadar, accounting for 17.90% (79.17 km²). High flood risk is most extensive in Nageshwari, covering 50.17% (221.10 km²) of the upazila area. Very high flood risk is again dominated by Kurigram Sadar, representing 45.48% (205.79 km²) of its total area.

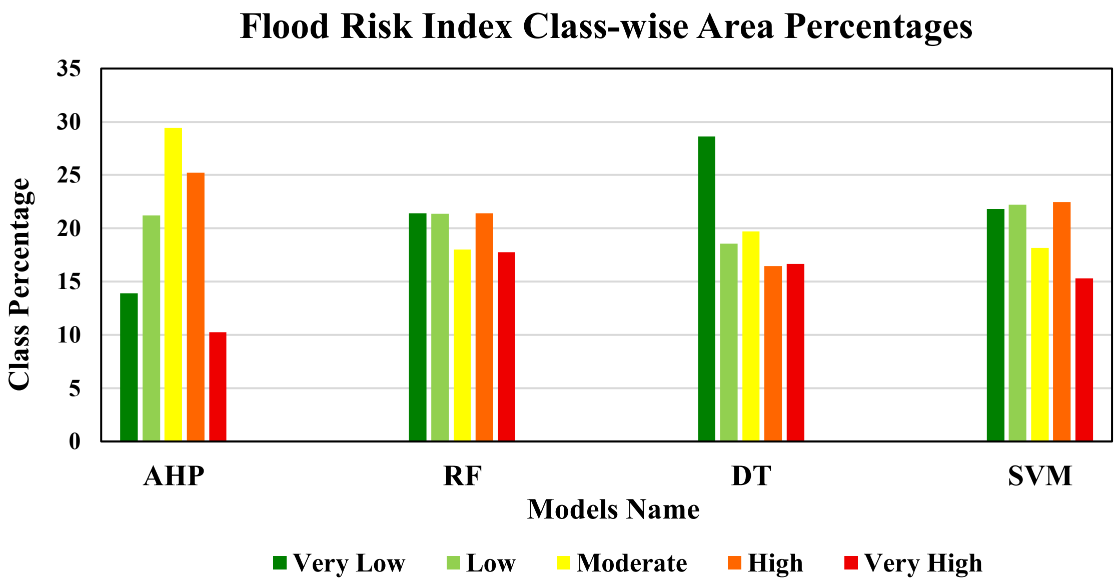

Figure 8. Class-wise Area Percentage Distributions of the Flood Risk Index (FRI) models

The SVM-based flood risk map (Figures 7d and 8) shows the sharpest spatial boundaries and a more distinct delineation of extreme flood-risk zones. Very low flood risk is concentrated in Charrajibpur (36.79%, 180.42 km²) and Ulipur (45.43%, 88.79 km²). Low-risk areas are primarily observed in Ulipur (39.57%, 199.73 km²) and Raomari (25.95%, 129.28 km²). Moderate flood risk zones are substantial in Nageshwari (29.29%, 119.28 km²), while high flood risk remains dominant in the same upazila (38.38%, 154.16 km²). Very high flood risk is most pronounced in Kurigram Sadar, where it accounts for 60.47% (207.03 km²), clearly identifying this upazila as the principal flood hotspot under the SVM framework. The RF and SVM models provided a more refined delineation of high-risk zones in Kurigram Sadar compared to the coarser output of the AHP model, highlighting the superior capability of data-driven models in identifying spatially refined flood-risk hotspots for disaster management and flood planning.

Validation Using AUC-ROC

The performance of the flood risk models was evaluated using Receiver Operating Characteristic (ROC) curves (Figure 9), which show the balance between True Positive Rate (TPR) and False Positive Rate (FPR). Among the models, Random Forest (RF) performed the best with an Area Under the Curve (AUC) of 0.93, meaning it can very well distinguish between flood-prone and safe areas. Support Vector Machine (SVM) came next with an AUC of 0.91, showing strong predictive ability. Decision Tree (DT) had moderate performance (AUC = 0.86), and the Analytical Hierarchy Process (AHP) model was the lowest (AUC = 0.82), but still reasonably reliable. Overall, the ROC curves show that machine learning models like RF and SVM are better than the traditional AHP method, correctly identifying more flood-prone areas while reducing false alarms. These results support using RF and SVM for practical flood risk management.

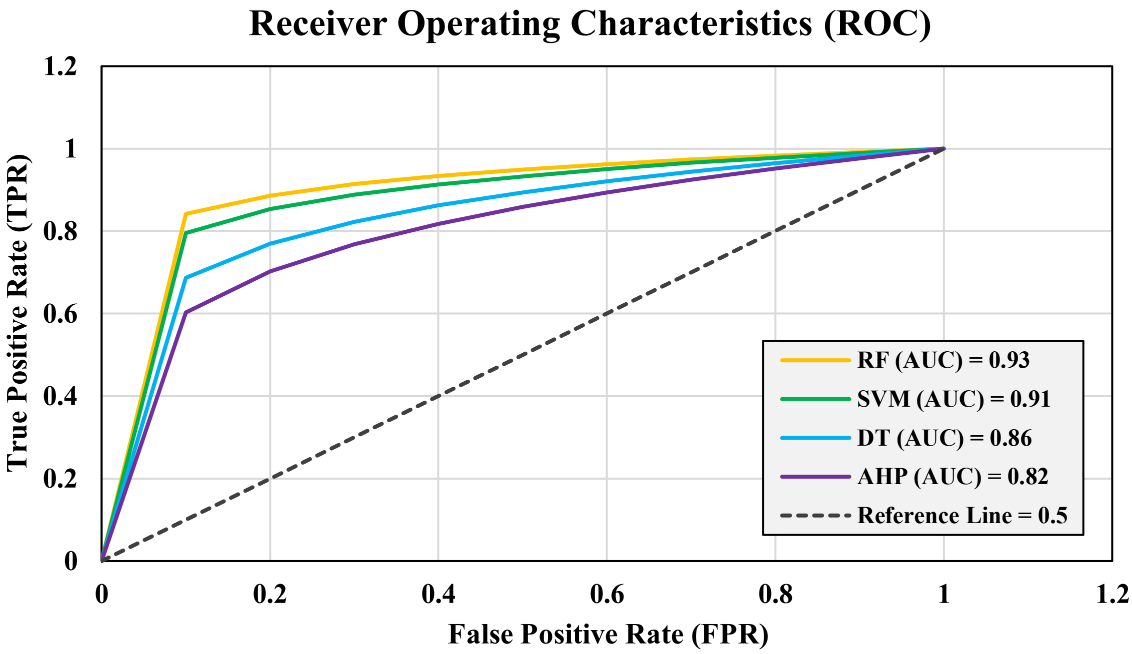

Figure 9. Comparison analysis of models' performance through ROC Curve and AUC Scores

DISCUSSION

This study presents an integrated flood risk assessment framework combining physical flood hazard, socio-economic vulnerability, and machine learning–based validation to address the complex nature of flood risk in Kurigram District, Bangladesh. The results demonstrate that flood risk is not solely controlled by hydrological or geomorphological conditions, but also strongly influenced by socio-economic sensitivity, consistent with contemporary disaster risk theory (Wisner et al., 2023; UNDRR, 2020). Although several parts of Kurigram exhibit moderate physical hazard, some upazilas experience high flood risk due to dense population, limited infrastructure accessibility, and low adaptive capacity, reinforcing previous findings that social vulnerability substantially amplifies flood impacts in Bangladesh (Haque & Zaman, 1993; Hoq et al., 2021; Alam et al., 2022). Despite the use of multi-source datasets from different years, the integrated datasets provide a reliable regional-scale representation of flood hazard and vulnerability conditions within the study area.

The Flood Hazard Index (FHI) identified rainfall, distance to rivers, drainage density, and elevation as the dominant drivers of flood occurrence, reflecting the strong influence of monsoon rainfall and active fluvial processes within the Brahmaputra–Teesta river system. Low elevation and dense drainage networks contribute to localized water accumulation and prolonged inundation, which are characteristic features of deltaic floodplains (Mirza, 2022; Rana et al., 2024). The inclusion of remote sensing–derived indices such as NDVI and MNDWI further improved hazard characterization by capturing vegetation and surface moisture dynamics (Costache et al., 2020; Antzoulatos et al., 2022).

The Flood Vulnerability Index (FVI) revealed that demographic pressure, gender-sensitive exposure, and limited accessibility to roads and healthcare facilities significantly increase flood risk, even in areas with moderate hazard intensity (Shuaibu et al., 2022; Mokhtari et al., 2023). In several locations, socio-economic vulnerability contributed more substantially to overall flood risk than physical hazard intensity alone, highlighting the critical role of social sensitivity and adaptive capacity in shaping flood risk patterns across Kurigram District. These findings indicate that effective flood risk management requires interventions beyond structural flood control, including improvements in social infrastructure, education, and accessibility.

A key contribution of this study lies in the comparative evaluation of AHP and machine learning models for flood risk mapping. While the AHP-based approach provides a transparent structure for integrating expert knowledge, its predictive performance (AUC = 0.82) was lower than that of the machine learning models. Among the evaluated algorithms, Random Forest (RF) achieved the highest predictive accuracy (AUC = 0.93), followed by Support Vector Machine (AUC = 0.91) and Decision Tree (AUC = 0.86). The superior performance of RF is primarily related to its ability to capture complex nonlinear relationships among flood-conditioning factors while reducing overfitting through ensemble averaging (Breiman, 2001; Chen et al., 2021; Adnan et al., 2023). These findings suggest that machine learning approaches are more effective in modelling complex spatial interactions compared to conventional weighted overlay techniques.

From a policy perspective, the strong agreement between GIS–AHP-derived flood risk patterns and ML-based validation confirms the robustness and practical applicability of the proposed hybrid framework. The identification of high-risk hotspots such as Kurigram Sadar and Nageshwari provides valuable insights for disaster risk reduction, land-use planning, and resilience-building initiatives. Given its methodological flexibility and moderate data requirements, the proposed framework offers a practical decision-support tool for improving climate resilience and sustainable flood risk management in flood-prone and data-scarce regions.

CONCLUSION

This study developed an integrated GIS-based flood risk assessment framework by combining multi-criteria decision analysis and machine learning techniques to effectively capture the complex interactions between physical flood hazards and socio-economic vulnerability in Kurigram District, Bangladesh. By constructing a Flood Hazard Index (FHI) using ten environmental variables and a Flood Vulnerability Index (FVI) based on five key socio-economic indicators, the study provides a comprehensive and multidimensional representation of flood risk that extends beyond conventional hazard-focused approaches (Saaty, 1980; Hu et al., 2017; Alam Chyon et al., 2023). The AHP-based weighting scheme demonstrated strong internal consistency, reinforcing the robustness and reliability of factor prioritization in flood risk modeling (Abu El-Magd et al., 2020; Mokhtari et al., 2023). The spatial distribution of flood risk reveals that moderate to high-risk zones occupy a substantial portion of Kurigram District, with risk patterns closely associated with population concentration, gender-sensitive exposure, accessibility to critical infrastructure, and educational attainment. Importantly, several areas characterized by only moderate physical hazard levels exhibit high overall flood risk due to elevated socio-economic sensitivity and limited adaptive capacity. This finding underscores the necessity of integrating vulnerability indicators into flood risk assessments and supports earlier evidence that flood risk in Bangladesh is strongly shaped by social and infrastructural dimensions rather than hydrological factors alone (Haque & Zaman, 1993; Hossain & Islam, 2021; Alam et al., 2022; Shahriar, 2021). The incorporation of machine learning algorithms further strengthened the assessment by validating the spatial reliability of the flood risk patterns. Among the applied models, the Random Forest (RF) algorithm achieved the highest predictive performance with an AUC value of 0.93, demonstrating superior capability in capturing nonlinear relationships among flood-conditioning factors compared to the SVM, DT, and conventional AHP approaches (Breiman, 2001; Costache et al., 2020; Chen et al., 2021; Adnan et al., 2023). The strong agreement between GIS–AHP-derived flood risk outputs and ML-based validation confirms the methodological robustness and practical applicability of the proposed hybrid framework. Therefore, this study contributes a scientifically rigorous and policy-relevant flood risk assessment approach that can support targeted disaster risk reduction, efficient resource allocation, and resilient land-use planning in one of Bangladesh's most flood-affected regions. Given its flexibility and moderate data requirements, the proposed framework may also be transferable to other flood-prone and data-scarce regions for supporting climate resilience and sustainable flood management planning.

Although this study provides a comprehensive flood risk assessment framework integrating GIS, AHP, and machine learning techniques, several limitations should be acknowledged. The analysis relied on multi-source datasets acquired from different years due to data availability constraints, which may introduce minor uncertainties in the spatial representation of flood risk. The flood inventory developed from Sentinel-1 SAR imagery may also contain uncertainties associated with radar shadow effects, mixed pixels, and seasonal variability. In addition, the AHP method involves expert judgment during weight assignment, while several potentially important variables, such as river discharge, soil moisture, groundwater conditions, and real-time rainfall intensity, could not be incorporated because of data limitations.

Future research should focus on integrating high-resolution topographic and hydrological datasets, real-time flood monitoring information, and climate change scenarios to improve flood prediction accuracy. The application of advanced deep learning approaches and the inclusion of community-level socio-economic indicators and infrastructure resilience factors may further enhance flood risk assessment and support more effective disaster management strategies.

AUTHOR'S CONTRIBUTIONS

Pranta Bhowmik conceptualized the study, developed the methodology, collected and curated the data, performed GIS, remote sensing, and machine learning analyses, prepared the visualizations, and wrote the original manuscript draft. Md. Aminul Haque Laskor contributed to methodology development, supervision, validation, and critical review of the manuscript. Md. Mohaimenul Islam assisted with data interpretation, literature review, and manuscript revision. Ananta Bhowmik contributed to machine learning implementation, data processing, technical review, and manuscript editing. All authors read and approved the final version of the manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest regarding the publication of this manuscript.

REFERENCES

Abu El-Magd, S. A., Amer, R. A., & Embaby, A. (2020). Multi-criteria decision-making for the analysis of flash floods: A case study of Awlad Toq-Sherq, Southeast Sohag, Egypt. *Journal of African Earth Sciences, 162*, 103709. https://doi.org/10.1016/j.jafrearsci.2019.103709

Adnan, M. S. G., Siam, Z. S., Kabir, I., Kabir, Z., Ahmed, M. R., Hassan, Q. K., Rahman, R. M., & Dewan, A. (2023). A novel framework for addressing uncertainties in machine learning-based geospatial approaches for flood prediction. *Journal of Environmental Management, 326*(Pt B), 116813. https://doi.org/10.1016/j.jenvman.2022.116813

Alam Chyon, M. S., Biswas, S., Mondal, M. S., Roy, B., & Rahman, A. (2023). Integrated assessment of flood risk in the Arial Khan floodplain of Bangladesh under changing climate and socioeconomic conditions. *Journal of Flood Risk Management*. https://doi.org/10.1111/jfr3.12876

Alam, M., Rahman, M., & Karim, R. (2022). Riverine flooding and its socio-economic impacts in northern Bangladesh. *Journal of Flood Risk Management, 15*(4), e12875. https://doi.org/10.1111/jfr3.12875

Antzoulatos, G., Kouloglou, I.-O., 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. https://doi.org/10.3390/su14063251

Ali, R., Sajjad, H., Saha, T. K., Rahaman, M. H., Sharma, A., & Ganaie, D. B. (2025). Assessment of flood risk using remote sensing data and machine learning models in Upper Jhelum Subcatchment, India. *Human and Ecological Risk Assessment: An International Journal, 31*(7–8), 993–1019. https://doi.org/10.1080/10807039.2025.2516586

Asiatic Society of Bangladesh. (2003). Kurigram District. In *Banglapedia: National Encyclopedia of Bangladesh*. http://en.banglapedia.org/index.php?title=Kurigram_District

Bangladesh Bureau of Statistics. (2022). *Population and housing census 2022: Preliminary report* (pp. viii, 28, 30, 39, 44, 46). Government of Bangladesh. ISBN 978-984-35-2977-0

Bangladesh National Portal. (2018). Geographical introduction. Archived from the original on March 21, 2023. https://www.kurigram.gov.bd

Beven, K. J., & Kirby, M. J. (1979). A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin*, *24*(1), 43–69. https://doi.org/10.1080/02626667909491834

Breiman, L. Random Forests. *Machine Learning* **45**, 5–32 (2001). https://doi.org/10.1023/A:1010933404324

Can, Y., Liu, L.-L., Huang, F., Huang, L., & Wang, X.-M. (2023). Machine learning-based landslide susceptibility assessment with optimized ratio of landslide to non-landslide samples. *Gondwana Research, 123*, 198–216. https://doi.org/10.1016/j.gr.2022.05.012

Chen, J., Huang, G., & Chen, W. (2021). Towards better flood risk management: Assessing flood risk and investigating the potential mechanism based on machine learning models. *Journal of Environmental Management, 293*, 112810. https://doi.org/10.1016/j.jenvman.2021.112810

Chowdhury, M. E., Islam, A. K. M. S., Zzaman, R. U., & Khadem, S. (2025). A machine learning-based approach for flash flood susceptibility mapping considering rainfall extremes in the northeast region of Bangladesh. *Advances in Space Research, 75*(2), 1990–2017. https://doi.org/10.1016/j.asr.2024.10.047

Costache, R., Pham, Q. B., Sharifi, E., Linh, N. T. T., Abba, S. I., Vojtek, M., Vojteková, J., Nhi, P. T. T., & Khoi, D. N. (2020). Flash-flood susceptibility assessment using multi-criteria decision making and machine learning supported by remote sensing and GIS techniques. *Remote Sensing, 12*(1), 106. https://doi.org/10.3390/rs12010106

Dhaka Tribune. (2024). Rising water levels trigger flooding in Kurigram. https://www.dhakatribune.com

Duy, N., Tuan Pham, L., Xuan Linh, N., Van Truong, T., Dang, D. K., Quang Hai, T., & Bui, Q.-T. (2024). Flood risk assessment using machine learning, hydrodynamic modelling, and the analytic hierarchy process. *Journal of Hydroinformatics, 26*(8), 1852–1882. https://doi.org/10.2166/hydro.2024.033

Ehteram, M., Ahmed, A. N., Chow, M. F., Latif, S. D., Chau, K. W., Chong, K. L., & El-Shafie, A. (2022). Optimal operation of hydropower reservoirs under climate change. *Environment, Development and Sustainability, 25*, 10627–10659. https://doi.org/10.1007/s10668-022-02497-y

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation, 24*(1), 38–49. https://doi.org/10.1017/S0376892997000088

Hagos, Y. G., Andualem, T. G., Yibeltal, M., et al. (2022). Flood hazard assessment and mapping using GIS integrated with multi-criteria decision analysis in the Upper Awash River basin, Ethiopia. *Applied Water Science, 12*, 148. https://doi.org/10.1007/s13201-022-01674-8

Haque, C. E., & Zaman, M. Q. (1993). Human responses to riverine hazards in Bangladesh: A proposal for sustainable floodplain development. *World Development, 21*(1), 93–107. https://doi.org/10.1016/0305-750X(93)90139-Z

Hoq, M. S., Raha, S. K., & Hossain, M. I. (2021). Livelihood vulnerability to flood hazard: Understanding from the flood-prone Haor ecosystem of Bangladesh. *Environmental Management, 68*(2), 270–289. https://doi.org/10.1007/s00267-021-01441-6

Hossain, S., & Islam, M. (2021). Flood vulnerability assessment in the Brahmaputra floodplain of Bangladesh. *Natural Hazards, 108*(2), 1507–1524. https://doi.org/10.1007/s11069-021-04682-9

Hu, S., Cheng, X., Zhou, D., & Zhang, H. (2017). GIS-based flood risk assessment in suburban areas: A case study of the Fangshan District, Beijing. *Natural Hazards, 87*(3), 1525–1543. https://doi.org/10.1007/s11069-017-2828-0

Huang, Y., & Zhao, L. (2018). Review of landslide susceptibility mapping using support vector machines. *CATENA, 165*, 520–529. https://doi.org/10.1016/j.catena.2018.03.003

Islam, I., Rahman, M. A., Adham, M. I., Majumder, A. A. S., & Zaman, A. (2025). Assessment and zonation of flood susceptibility in Sylhet Division, Bangladesh using GIS and Analytic Hierarchy Process (AHP). *Journal of Flood Risk Management*. https://doi.org/10.1111/jfr3.70121

Kabir Uddin, S., Hossain, S., Bajracharya, B., Ahmed, B., & Islam, M. K. (2025). Rapid flood inundation mapping for effective management: A machine learning and pixel-based classification approach in Feni District, Bangladesh. *Journal of Flood Risk Management*. https://doi.org/10.1111/jfr3.70087

Kader, Z., Islam, M. R., Aziz, M. T., & Wan Jaafar, W. (2024). GIS and AHP-based flood susceptibility mapping: A case study of Bangladesh. *Sustainable Water Resources Management, 10*, 170. https://doi.org/10.1007/s40899-024-01150-y

Khosravi, K., Pham, B. T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., Prakash, I., & Bui, D. T. (2018). A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Science of The Total Environment, 627*, 744–755. https://doi.org/10.1016/j.scitotenv.2018.01.266

Lai, V., Huang, Y. F., Koo, C. H., Ahmed, A. N., & El-Shafie, A. (2022). Conceptual Sim-Heuristic optimization algorithm to evaluate the climate impact on reservoir operations. *Journal of Hydrology, 614*(A), 128530. https://doi.org/10.1016/j.jhydrol.2022.128530

Liu, Y., Liu, L., Sun, H., et al. (2025). Flood risk assessment combining machine learning with multi-criteria decision analysis in Jiangxi Province, China. *International Journal of Disaster Risk Science, 16*, 858–869. https://doi.org/10.1007/s13753-025-00669-8

Luu, C., & Von Meding, J. (2018). A flood risk assessment of Quang Nam, Vietnam using spatial multicriteria decision analysis. *Water, 10*(4), 461. https://doi.org/10.3390/w10040461

Masood, M., Takeuchi, K., 2012. Assessment of flood hazard, vulnerability, and risk of mid-eastern Dhaka using DEM and 1D hydrodynamic model. Nat. Hazards 61, 757–770. https://doi.org/10.1007/s11069-011-0060-x

Meraj, G., Romshoo, S.A., Yousuf, A.R. *et al.* Assessing the influence of watershed characteristics on the flood vulnerability of the Jhelum basin in the Kashmir Himalaya. *Nat Hazards* **77**, 153–175 (2015). https://doi.org/10.1007/s11069-015-1605-1

Ming, X., Liang, Q., Dawson, R., Xia, X., & Hou, J. (2022). A quantitative multi-hazard risk assessment framework for compound flooding considering hazard inter-dependencies and interactions. *Journal of Hydrology, 607*, 127477. https://doi.org/10.1016/j.jhydrol.2022.127477

Mirza, M. M. Q. (2002). Global warming and changes in the probability of occurrence of floods in Bangladesh and its implications. *Global Environmental Change, 12*(2), 127–138. https://doi.org/10.1016/S0959-3780(02)00002-X

Mokhtari, E., Mezali, F., Abdelkebir, B., & Engel, B. (2023). *Flood risk assessment using analytical hierarchy process: A case study from the Cheliff–Ghrib watershed, Algeria*. Journal of Water and Climate Change, 14(3), 694–711. https://doi.org/10.2166/wcc.2023.316

Pham, B. T., Luu, C., Dao, D. V., Phong, T. V., Nguyen, H. D., Le, H. V., von Meding, J., & Prakash, I. (2021). Flood risk assessment using deep learning integrated with multi-criteria decision analysis. *Knowledge-Based Systems, 219*, 106899. https://doi.org/10.1016/j.knosys.2021.106899

Pham, B. T., Shirzadi, A., Shahabi, H., Omidvar, E., Singh, S. K., Sahana, M., Talebpour Asl, D., Bin Ahmad, B., Quoc, N. K., & Lee, S. (2019). Landslide susceptibility assessment by novel hybrid machine learning algorithms. *Sustainability, 11*(16), 4386. https://doi.org/10.3390/su11164386

Polong, F., Deng, K., Pham, Q. B., Linh, N. T. T., Abba, S. I., Ahmed, A. N., Anh, D. T., Khedher, K. M., & El-Shafie, A. (2023). Separation and attribution of impacts of changes in land use and climate on hydrological processes. *Theoretical and Applied Climatology, 151*, 1337–1353. https://doi.org/10.1007/s00704-022-04351-7

Rahman, M., Chowdhury, T., & Sultana, N. (2023). Climate change and flood risk management in Bangladesh. *Environmental Science & Policy, 140*, 210–222. https://doi.org/10.1016/j.envsci.2023.04.012

Rana, S. M. S., Habib, S. A., Sharifee, M. N. H., Sultana, N., & Rahman, S. H. (2024). Flood risk mapping of the flood-prone Rangpur division of Bangladesh using remote sensing and multi-criteria analysis. *Natural Hazards Research, 4*(1), 20–31. https://doi.org/10.1016/j.nhres.2023.09.012

Rowell, A. (2017, August 22). Severe flooding in South Asia has caused more than 1,200 deaths this summer. *Time*. https://time.com/4921340/south-asia-floods-india-mumbai-bangladesh-nepal/

Rudra, R. R., & Sarkar, S. K. (2023). Artificial neural network for flood susceptibility mapping in Bangladesh. *Heliyon, 9*(6), e16459. https://doi.org/10.1016/j.heliyon.2023.e16459

Saaty, T. L. (1980). *The analytic hierarchy process: Planning, priority setting, resource allocation*. McGraw-Hill.

Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. J. Math. Psychol. 15, 234–281. https://doi.org/10.1016/0022-2496(77)90033-5

Seydi, S. T., Kanani-Sadat, Y., Hasanlou, M., Sahraei, R., Chanussot, J., & Amani, M. (2023). Comparison of machine learning algorithms for flood susceptibility mapping. *Remote Sensing, 15*(1), 192. https://doi.org/10.3390/rs15010192

Shahriar, S. (2021). Impacts of floods and riverbank erosions on the rural lives and livelihoods strategies in Bangladesh: Evidence from Kurigram. In W. Leal Filho, J. Luetz, & D. Ayal (Eds.), *Handbook of climate change management* (pp. 1–15). Springer, Cham. https://doi.org/10.1007/978-3-030-22759-3_150-1

Shuaibu, A., Hounkpè, J., Bossa, Y. A., & Kalin, R. M. (2022). Flood risk assessment and mapping in the Hadejia River Basin, Nigeria, using a hydro-geomorphic approach and multi-criterion decision-making method. *Water, 14*(22), 3709. https://doi.org/10.3390/w14223709

Siam, Z. S., Hasan, R. T., Anik, S. S., Noor, F., Adnan, M. S. G., Rahman, R. M., & Dewan, A. (2022). National scale flood risk assessment using GIS and remote sensing based hybridized deep neural network and fuzzy analytic hierarchy process models: A case of Bangladesh. *Geocarto International, 37*(26), 12119–12148. https://doi.org/10.1080/10106049.2022.2063411

Sinha, R., Bapalu, G.V., Singh, L.K. *et al.* Flood risk analysis in the Kosi river basin, north Bihar, using a multi-parametric approach of Analytical Hierarchy Process (AHP). *J Indian Soc Remote Sens* 36, 335–349 (2008). https://doi.org/10.1007/s12524-008-0034-y

Suman, A. S., & Bhattacharya, B. (2015). Flood characterisation of the Haor region of Bangladesh using flood index. *Hydrology Research, 46*(5), 824–835. https://doi.org/10.2166/nh.2014.065

Termeh, S. V. R., Kornejady, A., Pourghasemi, H. R., & Keesstra, S. (2018). Flood susceptibility mapping using novel ensembles of adaptive neuro-fuzzy inference systems and metaheuristic algorithms. *Science of The Total Environment, 615*, 438–451. https://doi.org/10.1016/j.scitotenv.2017.09.262

The Daily Star. (2024). 1.5 lakh stranded in Kurigram as flood situation worsens. https://www.thedailystar.net

Tsangaratos, P., Ilia, I. Landslide susceptibility mapping using a modified decision tree classifier in the Xanthi Perfection, Greece. *Landslides* 13, 305–320 (2016). https://doi.org/10.1007/s10346-015-0565-6

UNDRR. (2020). *Human cost of disasters: An overview of the last 20 years (2000–2019)*. United Nations Office for Disaster Risk Reduction. https://www.undrr.org/publication/human-cost-disasters-2000-2019

UNICEF. (2017, September 2). 16 million children are affected by massive flooding in South Asia, with millions more at risk. *UNICEF*. https://www.unicef.ca/en/press-release/16-million-children-affected-massive-flooding-south-asia-millions-more-risk

Wisner, B., Blaikie, P., Cannon, T., Davis, I., 2003. At Risk: Natural Hazards, People's Vulnerability and Disasters, second ed. Routledge, United States.

Yariyan, P., Janizadeh, S., Van Phong, T. *et al.* Improvement of Best First Decision Trees Using Bagging and Dagging Ensembles for Flood Probability Mapping. *Water Resour Manage* 34, 3037–3053 (2020). https://doi.org/10.1007/s11269-020-02603-7

Zanotti, C., Rotiroti, M., Sterlacchini, S., Cappellini, G., Fumagalli, L., Stefania, G. A., Nannucci, M. S., Leoni, B., & Bonomi, T. (2019). Choosing between linear and nonlinear models and avoiding overfitting for short and long-term groundwater level forecasting in a linear system. *Journal of Hydrology*, *578*, 124015. https://doi.org/10.1016/j.jhydrol.2019.124015