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

Zakaria Shams Siam[a,e], Rubyat Tasnuva Hasan[a], Soumik Sarker Anik[a], Fahima Noor[a], Mohammed Sarfaraz Gani Adnan[b,d] *, Rashedur M. Rahman[a], Ashraf Dewan[c]

[a] *Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh*
[b] *Department of Urban and Regional Planning, Chittagong University of Engineering and Technology (CUET), Chattogram 4349, Bangladesh*
[c] *Spatial Sciences Discipline, School of Earth and Planetary Sciences, Curtin University, Perth 6102, Australia*
[d] *Environment Change Institute, School of Geography and the Environment, University of Oxford, OX1 3QY, United Kingdom*
[e] *Department of Electrical and Computer Engineering, Presidency University, Dhaka, Bangladesh*

* Corresponding author at: Mohammed Sarfaraz Gani Adnan; Department of Urban and Regional Planning, Chittagong University of Engineering and Technology (CUET), Chattogram 4349, Bangladesh; sarfarazadnan@cuet.ac.bd

This is a peer-reviewed, author's accepted manuscript of the following research article: 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

National-Scale Flood Risk Assessment Using GIS and Remote Sensing-Based Hybridized Deep Neural Network and Fuzzy Analytic Hierarchy Process: A Case of Bangladesh

Assessing flood risk is challenging due to complex interactions among flood susceptibility, hazard, exposure, and vulnerability parameters. This study presents a novel flood risk assessment framework by utilizing a hybridized deep neural network (DNN) and fuzzy analytic hierarchy process (AHP) models. Bangladesh was selected as a case study region, where limited studies examined flood risk at a national scale. The results exhibited that hybridized DNN and fuzzy AHP models can produce the most accurate flood risk map while comparing among 15 different models. About 20.45% of Bangladesh are at flood risk zones of moderate, high, and very high severity. The northeastern region, as well as areas adjacent to the Ganges–Brahmaputra–Meghna rivers, have high flood damage potential, where a significant number of people were affected during the 2020 flood event. The risk assessment framework developed in this study would help policymakers formulate a comprehensive flood risk management system.

Keywords: Flood Risk Assessment; Flood Susceptibility Mapping; Hybridized Deep Neural Network; Hybridized Support Vector Regression; Genetic Algorithm; Fuzzy Analytic Hierarchy Process; Random Forest

Introduction

Flooding is known to be one of the most common yet devastating natural hazards (Stefanidis and Stathis 2013, Dewan 2015, Rahmati et al. 2020). Floods caused direct economic losses of USD 386 billion worldwide since 2001 (Wang et al. 2011, Rahmati et al. 2020). Economic damages caused by floods negatively impact human wellbeing, promoting long-term poverty in flood-affected regions (Adnan et al. 2020a, Barbour et al. 2022). An upsurge in population growth, exorbitant poverty, and climate change have increased flood risk in developing countries, especially in South Asia (Rahman et al. 2021a). Locating in an active deltaic region and crisscrossed by many large river channels, Bangladesh is frequently affected by floods of different magnitudes primarily due to high discharge in the Ganges, Brahmaputra, and Meghna (GBM) rivers caused by an excessive amount of rainfall in upstream regions(Chowdhury and Hassan 2017, Leon et al. 2020, Rahman et al. 2021b). The country is generally affected by four distinct types of floods: riverine or fluvial, flash or rainwater, urban or pluvial, and coastal floods (Adnan et al. 2019b). Heavy monsoon rainfall in the upstream river catchments leads to recurring riverine floods in Bangladesh (Rahman et al. 2021a). Various extreme riverine flood events, especially those that occurred in 1988, 1998, and 2004, killed many lives and caused extensive property damages, causing significant losses to the national economy (Dewan 2015). Most recently (in 2020), about a quarter of the country’s lands was inundated by monsoon flooding, affecting over four million people (NASA 2020).

Since flooding is the outcome of extremely complex and intricate dynamic processes, it is nearly impossible to prevent it from occurring (Pappenberger et al. 2006). Hence, flood risk reduction has become one of the major challenges worldwide (Rahmati et al. 2020). Reducing the detrimental effects of flooding depends on a quick and accurate assessment of risk, which helps to formulate risk management plans (Mojaddadi et al. 2017). The emergence of various remote sensing and the geospatial techniques has enabled researchers and practitioners to assess flood risk more accurately (Dewan and Kankam-Yeboah 2006, Pradhan 2010, Thirumurugan and Krishnaveni 2019, Rahman et al. 2021b). Evaluation of flood risk includes investigating flood risk-prone zones where the flood potentials are very high (Mojaddadi et al. 2017). A comprehensive flood risk assessment plays a vital role in the overall flood risk management system, which requires quantification of flood hazard, exposure, and vulnerability (Meyer et al. 2009, Pham et al. 2021a, Pham et al. 2021b). Various studies indicated that an accurate flood susceptibility model (FSM) can be translated into a flood hazard model by integrating factors such as flood depth, flood duration, and rainfall (Mojaddadi et al. 2017, Rahman et al. 2019, Pham et al. 2021a, Pham et al. 2021b, Rahman et al. 2021a).

Several studies conducted flood risk assessments both at the local and national scales around the world with the aid of remote sensing and GIS techniques, traditional statistical models, and multi-criteria decision analysis (MCDA) methods (Wang et al. 2011, Rincón et al. 2018, Luu et al. 2019, Akay and Baduna Koçyiğit 2020, Akay 2021, Ekmekcioğlu et al. 2021). However, the results produced by those methods could be affected by the nonlinear and dynamic nature of flooding (Tehrany et al. 2015), scarcity of necessary data especially in developing countries (Darabi et al. 2019), and restricted applicability of the models at multiple scales (de Moel et al. 2015). The limitations of various statistical flood models have prompted researchers to apply different machine learning (ML) algorithms in assessing flood risk (Rahmati et al. 2020). Recent studies applied different standalone as well as hybridized ML models. For instance, hybridized support vector machine (SVM) (Mojaddadi et al. 2017, Ma et al. 2019b) including SVM based on the radial basis function (SVM-RBF) (Ngo et al. 2021, Siam et al. 2021a) and SVM with the convolutional neural network (CNN) (Wang et al. 2020), standalone and hybridized decision table models (Pham et al. 2021b), hybridized decision tree (DT) (Chen et al. 2021) and others (Darabi et al. 2019). Tehrany et al. (2015) examined the efficacy of SVM in flood susceptibility mapping by comparing the performance of such models with four distinct kernels: linear, polynomial, RBF, and sigmoid. All these studies reported that hybridized ML models potentially produce more accurate results compared to standalone models (Rahmati et al. 2020, Siam et al. 2021a, Siam et al. 2021b). Also, to address the uncertainties related to the classical MCDA approaches, a few studies exploited the fuzzy MCDA approach (Akay 2021, Costache et al. 2021, Vilasan and Kapse 2021).

The application of deep learning (DL) algorithms has proved to be very efficient in quantifying flood probability (Ma et al. 2019a). Recently, several studies have been conducted using various deep neural network (DNN) architectures for FSM, with various combinations of algorithms. The latest DNN-based flood susceptibility models include the use of (1) DNN in combination with the manta ray foraging optimization algorithm (Nguyen et al. 2021), (2) combined the multilayer perceptron (MLP) and autoencoder models (Ahmadlou et al. 2021), (3) CNN and recurrent neural network (RNN) (Panahi et al. 2021),

(4) standalone and hybridized CNN architectures (Wang et al. 2020). However, all these studies were limited to flood susceptibility assessment. Consequently, little is known regarding the applicability of hybridized models in assessing flood risk. Only a few studies utilized DNN models (Chen et al. 2021) in combination with the MCDA approach for flood risk modeling (Pham et al. 2021a, Pham et al. 2021b). Still, the use of the hybridized DNN architectures is underexplored in flood risk studies. Besides, in the context of Bangladesh, only a few studies carried out flood susceptibility assessment at a national scale (Rahman et al. 2019, Rahman et al. 2021a, Rahman et al. 2021b, Siam et al. 2021a), while no study has attempted to quantify country-level flood risk.

In response to the above-discussed research gaps, this study aims to present a flood risk assessment framework by utilizing a hybridized DNN and fuzzy analytic hierarchy process (AHP) models. This study hypothesized that the integration of hybridized DNN model with the fuzzy AHP method can potentially produce more realistic results than the classical AHP method. Unlike previous studies on risk assessment framework to flood, we have modeled a hybridized DNN-based flood susceptibility model as a principal operator in developing a flood hazard map. The framework has been applied in assessing flood risk at the national scale in Bangladesh.

Materials and methods

The study was conducted in five steps. First, various flood conditioning factors were identified for developing a flood susceptibility model. Second, flood susceptibility models were developed based on different standalone and hybridized DNN and SVR models, as well as other conventional ML models (e.g., conditional inference tree, KNN, and MLP). Third, based on several evaluation metrics, the best-performing method was chosen for mapping the flood susceptibility. Fourth, flood hazard, exposure, and vulnerability maps were developed using the fuzzy AHP method, where the best-performing flood susceptibility map was used to model flood hazards. Finally, a flood risk map was developed by integrating flood hazard, exposure, and vulnerability maps. Figure 1 shows a brief methodological overview of the present study.

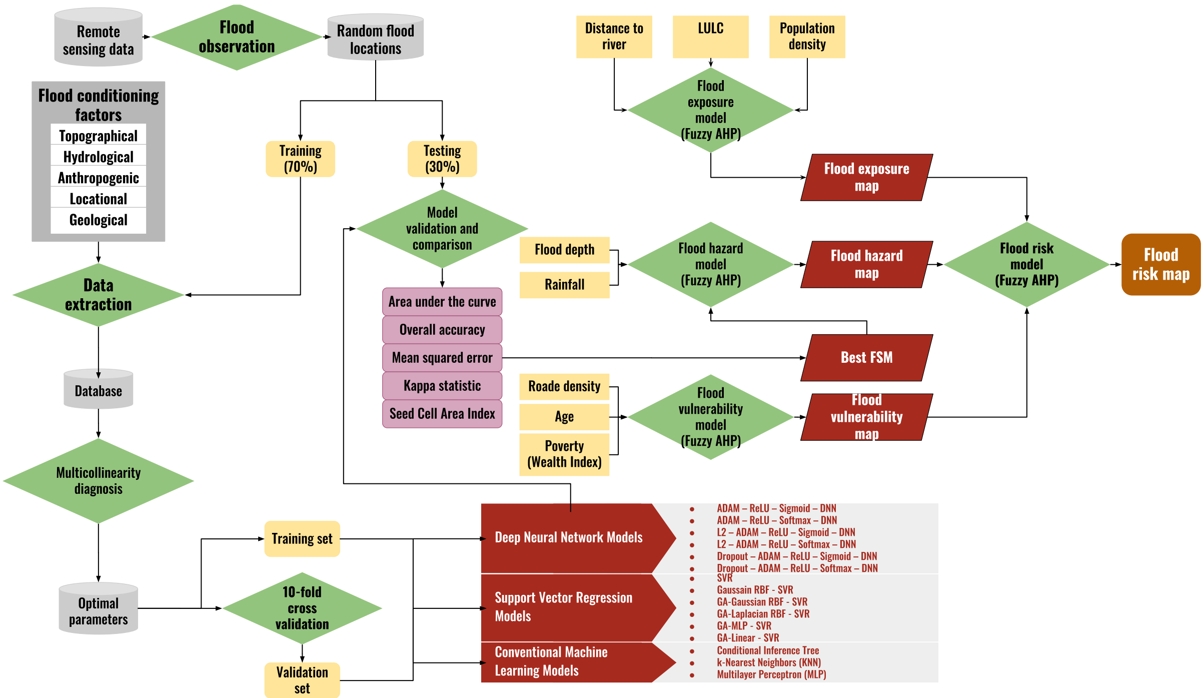


Figure 1. Flowchart of this study

Study area

The present study focused on Bangladesh (Figure 2). Geographically, the country is located in South Asia, between the latitudes of 20°34′ and 26°38′ to the north and longitudes of 88°01′ and 92°41′ to the east (Hasan et al. 2017, Rahman et al. 2019). More than 162.7 million people inhabit the country, with an annual population growth rate of 1.37%, within an area of 1,47,570 km^2. Thus, Bangladesh has the highest population density in the world, with a density of approximately 1,063 people per km^2 (Hasan et al. 2017, Rahman et al. 2019). The country is characterized by five topographic regions — Chittagong, Tippera-Comilla, north Bengal, northeastern, and southwestern regions — comprising 64 districts, eight divisions, and 492 subdistricts (Islam and Sado 2000). It includes three major river systems: the Ganges, Meghna, and Brahmaputra, with numerous distributaries and tributaries. The geographical location, flat topography, and tropical climatic conditions of Bangladesh make it one of the world's most flood-prone areas. The yearly average precipitation generally ranges between 2200 and 2500 mm. Annual mean temperature ranges between 25 °C and 35 °C. Almost 80% of the total landmass of Bangladesh is fertile alluvial lowlands. The rest of the country slightly elevated older plains and small hilly regions (Rahman et al. 2019).

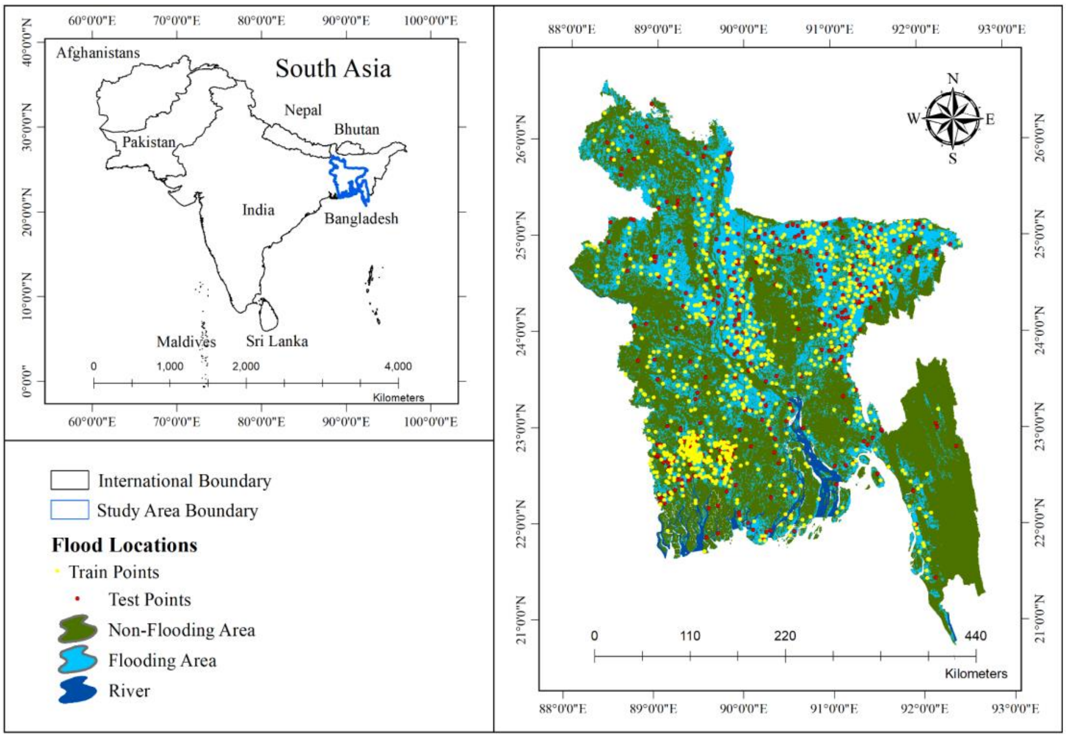


Figure 2. Map of Bangladesh with sample flood locations

Flood inventory mapping

The flood inundation areas of historical flooding events are typically used as a dependent variable for modeling flood susceptibility (Rahman et al. 2019, Pham et al. 2021a). Inundation data, of three different periods (July 12-21, July 23-27, and July 29-August 02) in monsoon season of 2020, were collected from the United Nations Institute for Training and Research (UNITAR). The UNITAR used Sentinel-1 satellite data to detect inundated areas (UNITAR 2020). The obtained inundation vector files were then converted to raster layers at 30 m resolution to ensure agreement with the digital elevation model (DEM) used in this study. The inundation raster layer was binarized — non-flood and flood locations were labeled as 0 and 1, respectively (equation 1).

$$Flood\ Inventory, y = \begin{cases} 1;\ if\ flooding \\ 0;\ if\ non-flooding \end{cases} \tag{1}$$

The combined flood inundation map was utilized to produce sample flood and non-flood points. A total of 2,766 sample points (flood points – 1408 and non-flood points – 1358) were created using the stratified random sampling technique. The stratified random sampling technique divides a population into smaller homogeneous subgroups known as strata. The strata are constructed depending on the members' shared characteristics or attributes. This technique has been widely used in flood modeling due to its ability to reduce bias in the sample (Adnan et al. 2020a, Adnan et al. 2020b). Based on the previous studies (Pham et al. 2021a, Pham et al. 2021b), the sample points were split into two groups: 70% of the total sample points (983 flood points, 953 non-flood points) was considered to train the flood susceptibility model while the other 30% sample (425 flood points, 405 non-flood points) was employed to test the model. To reduce model overfitting, this study applied a 10-

fold cross-validation technique to further divide the train set (70% sample points) into train and validation sets.

Flood conditioning factors

An important component of preparing FSMs is to choose appropriate flood conditioning factors that contribute to the occurrence of flooding in an area (Pham et al. 2021a). There is no universal method to identify appropriate flood conditioning factors as different studies used various combinations (Rahman et al. 2019, Wang et al. 2019, Costache et al. 2020a, Rahmati et al. 2020, Talukdar et al. 2020, Pham et al. 2021a). However, factors should be identified according to the environmental conditions of the study area (Adnan et al. 2020b). In this study, initially, thirteen flood causative factors were chosen based on the topographical, hydrological, locational, geological, and anthropogenic characteristics of the study area. Selected factors include slope, aspect, curvature, elevation, Stream Power Index (SPI), flow accumulation, Topographic Wetness Index (TWI), soil permeability, soil texture, land use/land cover (LULC), geology, distance to rivers, and drainage density. The thematic maps for all thirteen flood causative factors were developed at a spatial resolution of 30 m (Figure 3).

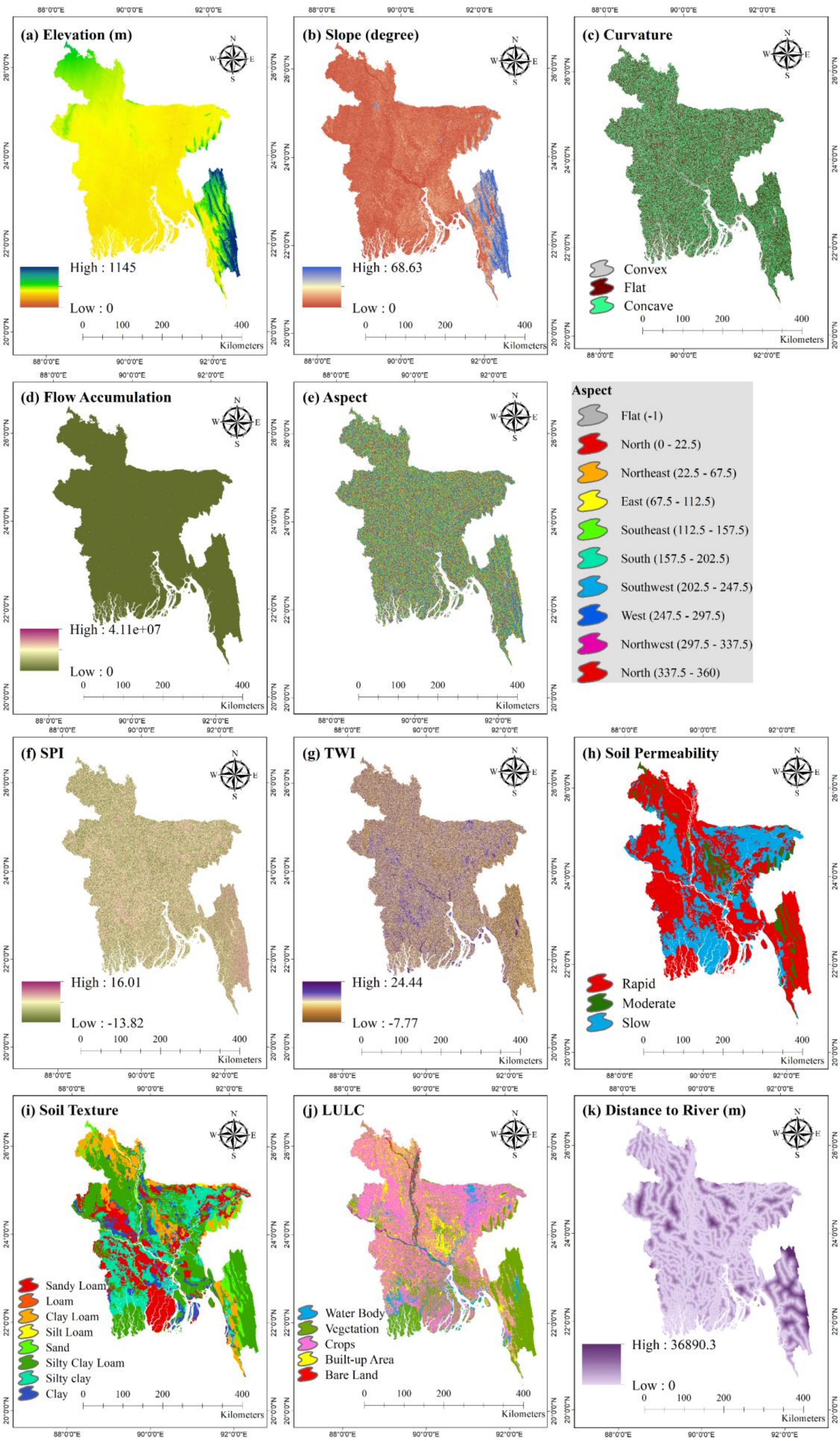
Topographical factors considered for flood susceptibility modeling include elevation, slope, aspect, and curvature. Surface elevation is an important factor accountable for flooding (Bui et al. 2020a, Sarkar and Mondal 2020, Islam et al. 2021). Generally, elevation is negatively associated with flooding, as areas with lower elevation tend to be highly susceptible to flooding (Rahman et al. 2021b). In this study, a raster elevation layer was prepared using the Advanced Land Observing Satellite (ALOS) Digital Elevation Model (DEM) at 30 m resolution (JAXA 2015). Other topographical factors like slope, curvature, and aspect are computed from DEM. Slope determines the runoff velocity after a rainfall event (Talukdar et al. 2020). Flood potential is higher in areas with a lower slope and vice versa (Adnan et al. 2020b). Aspect is another important topographical factor that indicates slope directions (Adnan et al. 2020b). Generally, aspect denotes the magnitude of rainfall and sunshine that an area would receive, influencing the water balance of an area (Tehrany et al. 2017). Curvature indicates geomorphological features of an area (Paul et al. 2019). Surfaces with flat or concave characteristics are usually susceptible to flooding (Adnan et al. 2020b).

Flow accumulation is an important hydrological factor that impacts the flood susceptibility of an area. The raster layer of flow accumulation was derived from DEM by developing a continuing network of drainage systems (Planchon and Darboux 2002). Pixel-wise flow accumulation value denotes accumulated water flowing in the downslope direction (Adnan et al. 2020b). The flow accumulation layer was used to identify drainage channels (Adnan et al. 2019a), which was later used to develop a drainage density layer. Other hydrological factors such as SPI and TWI indicate drainage characteristics of the study area. SPI typically exhibits the erosive power of flowing water (Talukdar et al. 2020), indicating the rate of sediment that could relocate to natural drainage channels (Adnan et al. 2020b). On the other hand, TWI denotes the amount of water that is accumulated in every pixel size (Islam et al. 2021). TWI explains the possibility of a wet surface. An area with higher SPI and TWI is highly likely to be flooded (Bannari et al. 2017). SPI and TWI were computed using equations (2) and (3).

$$SPI = A_s \times tan\beta \qquad (2)$$

$$TWI = ln\left(\frac{A_s}{\beta}\right) \qquad (3)$$

where, A_s is the fixed catchment region (m^2/m) and β is the slope gradient.


(a) Elevation (m)
High : 1145
Low : 0
(b) Slope (degree)
High : 68.63
Low : 0
(c) Curvature
Convex
Flat
Concave
(d) Flow Accumulation
High : 4.11e+07
Low : 0
(e) Aspect
Aspect
Flat (-1)
North (0 - 22.5)
Northeast (22.5 - 67.5)
East (67.5 - 112.5)
Southeast (112.5 - 157.5)
South (157.5 - 202.5)
Southwest (202.5 - 247.5)
West (247.5 - 297.5)
Northwest (297.5 - 337.5)
North (337.5 - 360)
(f) SPI
High : 16.01
Low : -13.82
(g) TWI
High : 24.44
Low : -7.77
(h) Soil Permeability
Rapid
Moderate
Slow
(i) Soil Texture
Sandy Loam
Loam
Clay Loam
Silt Loam
Sand
Silty Clay Loam
Silty clay
Clay
(j) LULC
Water Body
Vegetation
Crops
Built-up Area
Bare Land
(k) Distance to River (m)
High : 36890.3
Low : 0
0 100 200 400
Kilometers
88°0'0"E
90°0'0"E
92°0'0"E
20°0'0"N
22°0'0"N
24°0'0"N
26°0'0"N

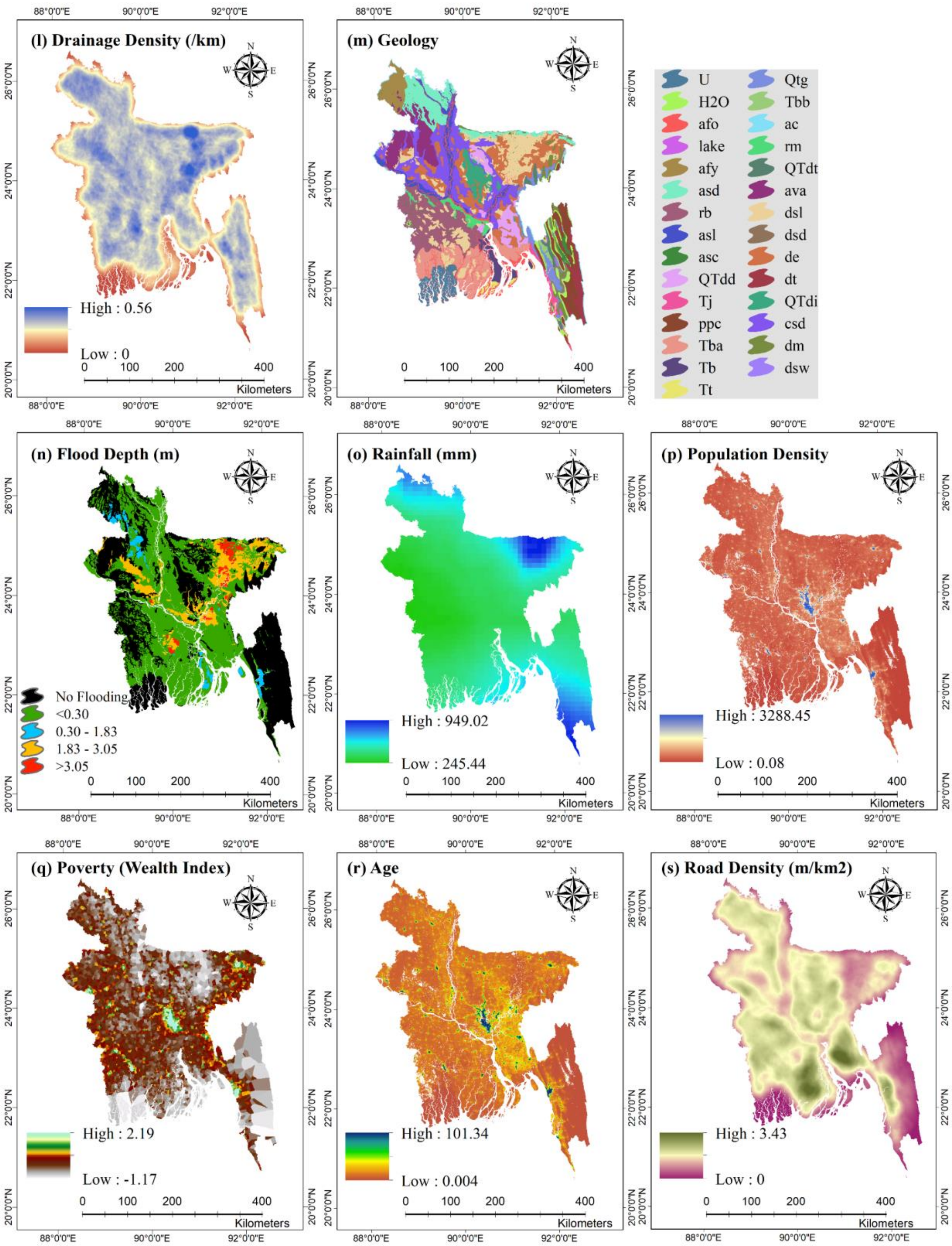


Figure 3. Thematic layers of various indicators for modeling flood risk

This study also used three geological factors: geology, soil permeability, and soil texture. Soil texture controls infiltration rate as well as surface runoff, hence, it is considered a significant flood conditioning factor (Rahman et al. 2021b). The raster layer of soil texture was taken from Bangladesh Agricultural Research Council (BARC) database (BARC 2014).

Soil permeability data can explain runoff patterns and drainage processes. It indicates the hydraulic activity of unsaturated soils and is an important factor influencing streamflow (Singh et al. 2020). The soil permeability data were also obtained from BARC (2014). This study also considered the geological characteristics of Bangladesh. The geology of an area influences the formation and construction of drainage patterns (Islam and Sado 2000, Bui et al. 2019), leading to the generation and development of floodplains. Typically, areas with a mostly impenetrable surface geology are highly susceptible to flood (Islam and Sado 2000). The digital geological data of Bangladesh was taken from the United States Geological Survey (USGS) (Persits et al. 2001).

LULC is a crucial flood conditioning factor since it directs the initiation as well as infiltration of the surface runoff and transportation of sediment (Adnan et al. 2020b). It directly impacts some parameters in the hydrological cycle such as interception and concentration (Rahman et al. 2019). Generally, built-up areas are more prone to flooding compared to the forest and open spaces due to low infiltration rates and high surface runoff (Talukdar et al. 2020). LULC data of 2020 was collected from the Environmental Systems Research Institute (Esri), which is developed using Sentinel-2 imagery (Karra et al. 2021).

Rivers are considered as the main paths of water flow causing flood events (Rahmati et al. 2020). This study incorporated a layer explaining distance to river as a locational factor (Mojaddadi et al. 2017). Areas that are close to the river are generally more susceptible (Costache et al. 2020b, Talukdar et al. 2020). The distance to river layer was derived from a river network database, collected from Water Resources Planning Organization (WARPO) (WARPO 2018) using the Euclidean distance algorithm. Table 1 shows a summary of the sources and spatial resolution of flood causative factors.

Table 1. Indicators used for flood susceptibility, hazard, exposure and vulnerability modeling

No.	Factors	Spatial resolution	Variable type	Sources
1	Elevation	30 m	Numeric	(JAXA 2015)
2	Slope	"	Numeric	Derived from DEM
3	Aspect	"	Categorical	"
4	Curvature	"	Categorical	"
5	Flow Accumulation	"	Numeric	"
6	SPI	"	Numeric	"
7	TWI	"	Numeric	"
8	Soil Permeability	"	Categorical	(BARC 2014)
9	Soil Texture	"	Categorical	(BARC 2014)
10	LULC	10 m	Categorical	(Karra et al. 2021)
11	Geology	30 m	Categorical	(Persits et al. 2001)
12	Distance to River	"	Numeric	(WARPO 2018)
13	Drainage Density	"	Numeric	Derived from DEM
14	Flood Depth	"	Categorical	(BARC 2014)
15	Rainfall	11.1 km	Numeric	(Huffman et al. 2019)
16	Population Density	100 m	Numeric	(WorldPop 2020)

	(Population per Cell)			
17	Age (Less than 14 and Greater than 60)	100 m	Categorical	(Bondarenko et al. 2020)
18	Poverty (Wealth Index)	60 m – 5 km	Numeric	(Steele et al. 2017)
19	Road Density	30 m		(WARPO 2018)

Flood risk components

Flood hazard

This study considered flood susceptibility (Pham et al. 2021a), flood depth (Pham et al. 2021a), and rainfall (David and Schmalz 2020) to develop a flood hazard map of Bangladesh (Table 1). Rainfall is a crucial hydrological factor for flood hazard mapping (Lu et al. 2020). In Bangladesh, both short-term heavy rainfall and long-term low to moderate rainfall are accountable for flooding (Adnan et al. 2019b). Rainfall can cause hydrostatic pressure, promoting a higher water level in the major rivers (Rahman et al. 2019). Satellite-derived gridded precipitation data of July and August 2020, collected from Huffman et al. (2019), were used to develop a layer of the average monthly total rainfall. A thematic layer of flood depth was collected from BARC (2014) (Figure 3 (n)).

Flood exposure

Three indicators were used for developing a flood exposure map: distance to river, LULC, and population density (Table 1). Previous studies considered population density as an important indicator for modeling flood exposure (Zou et al. 2013, Pham et al. 2021a). Flood-prone areas with a high population density are more vulnerable to flooding than areas with a low density. In this study, population density data of 2020 was collected from WorldPop (2020) (Figure 3 (p)). As described in section 2.3, areas near the river are identified from DEM, and LULC data are collected from Karra et al. (2021).

Flood vulnerability

Flood vulnerability is typically correlated with the type of infrastructures as well as characteristics of the communities in flood-prone areas. Flood vulnerability was estimated based on three indicators: road density (Ronco et al. 2015, Pham et al. 2021a), age (Brito et al. 2018), and poverty (wealth index) (Pham et al. 2021a) (Table 1). Generally, flood-prone areas with a high road density are vulnerable to flooding (Pham et al. 2021a). A raster road density layer was derived from road network data collected from WARPO (2018). The population age structure is also a useful flood vulnerability indicator (Brito et al. 2018). A high percentage of children and older people increase flood vulnerability of an area (Brito et al. 2018). The age distribution data was retrieved from the WorldPop (Bondarenko et al. 2020), where the total number of people aged less than 14 and greater than 60 was estimated for Bangladesh for the year 2020. Also, an area with a high poverty ratio becomes vulnerable to flooding (Adnan et al. 2020a, Pham et al. 2021a). The wealth index data was retrieved from Steele et al. (2017) to analyze poverty scenarios. Flood vulnerability indicator maps are shown in Figure 3 (q–s).

Flood risk assessment

We estimated flood risk to be the product of flood hazard, exposure, and vulnerability (equation 4) (Pham et al. 2021a, Pham et al. 2021b).

$$Flood\ Risk = Flood\ Hazard \times Flood\ Exposure \times Flood\ Vulnerability \qquad (4)$$

Flood susceptibility modeling

Flood susceptibility modeling was considered as a component of flood hazard mapping. Pixel-wise flood susceptibility scores (FS) were estimated using equation (5) (Rahman et al. 2019, Siam et al. 2021a).

$$FS = \sum_{j=1}^{n} w_j x_j \qquad (5)$$

where n denotes the number of flood conditioning factors used for FSM, x_j indicates selected flood conditioning factors and w_j represents the weight of every factor. To find the optimal weight of every factor for flood susceptibility modeling, a total of six standalone and hybridized DNN models were established: adaptive moment estimation (ADAM) – rectified linear unit (ReLU) – Softmax – DNN, ADAM – ReLU – Sigmoid – DNN, L2 regularization (L2) – ADAM – ReLU – Softmax – DNN, L2 – ADAM – ReLU – Sigmoid – DNN, Dropout – ADAM – ReLU – Softmax – DNN and Dropout – ADAM – ReLU – Sigmoid – DNN. Also, a total of six standalone and hybridized SVR models were investigated such as standalone SVR, Gaussian Radial Basis Function Kernel (Gaussian RBF) – SVR using grid search technique, GA – Gaussian RBF – SVR, GA – laplacian RBF kernel (Laplacian RBF) – SVR, GA – sigmoid or multilayer perceptron kernel (MLP) – SVR and GA – linear kernel (Linear) – SVR. Besides, three conventional ML models (e.g., conditional inference tree, k-nearest neighbor (KNN), and MLP) were established. All standalone and hybridized deep neural network models were developed using the 'keras' package in the R programming language. The conditional inference tree, k-nearest neighbor, and multilayer perceptron models were established using the 'ctree' function of 'party' package, 'knnreg' function of 'caret' package, and 'neuralnet' function of 'neuralnet' package in R, respectively.

i. Multicollinearity analysis for optimizing features

In the present study, multicollinearity among flood causative factors was diagnosed by estimating the variance inflation factors (VIF) (Midi et al. 2010), using the 'Car' package in R, to remove factors that are subject to multicollinearity. VIF for each factor should be <2.5 to circumvent the model bias (Midi et al. 2010). If the value is >10, it denotes the presence of multicollinearity (Midi et al. 2010). After investigating multicollinearity, the flood susceptibility model includes a total of eleven flood conditioning factors whose VIF values were less than 2.5 (Bai et al. 2011). TWI and flow accumulation layers were discarded since the addition of these two layers increased VIF values (Table 2).

Table 2. VIF values, indicating multicollinearity of selected factors

Factors	VIF (Iteration 1)	VIF (Iteration 2)
Aspect	1.029	1.007
Distance to River	1.162	1.160
Drainage Density	1.189	1.182
Elevation	2.496	2.472

Flow Accumulation	4.119	-
Geology	1.487	1.481
LULC	1.255	1.254
Curvature	1.251	1.182
Slope	3.673	1.833
Soil Permeability	1.997	1.988
Soil Texture	2.461	2.459
SPI	6.410	1.284
TWI	7.621	-

ii. Feature scaling

Since we exploited gradient descent as well as distance-based models, all continuous variables such as slope, drainage density, distance to river, elevation and SPI were scaled using z-score normalization technique (equation 6).

$$z = \frac{x - \mu}{\sigma} \tag{6}$$

where, x is the feature value, μ and σ are mean and standard deviation of that feature, respectively. After feature scaling, values of eleven flood conditioning factors were extracted corresponding to flood and non-flood points.

iii. Standalone and hybridized DNN models

We developed and applied six standalone and hybridized DNN models for mapping flood susceptibility. In the DNN model, we experimented with three hidden layers consistent with the study by Bui et al (Bui et al. 2019). A total of eleven nodes (i.e., 11 flood conditioning factors) were taken in the input layer and one node (sample flood points) in the output layer. We set the number of nodes to eight in each of the three consecutive hidden layers since the number of nodes in each hidden layer is suggested to be in between the number of input nodes and output nodes (Bui et al. 2020b). We used rectified linear activation function (ReLU) in each of the three hidden layers. However, in the output layer, we used the sigmoid activation function and the softmax activation function separately. For the sigmoid activation function, we used the binary cross-entropy loss function. For the softmax activation function, we applied one-hot encoded the output variable. Therefore, the number of output nodes became two instead of one in the case of the softmax activation function. For the loss function, we used the categorical cross-entropy function for the softmax activation function.

We initialized the weights setting the parameters of mean to 0, the standard deviation to 0.05, and the biases with the values of zero. For gradient descent optimization, we used the ADAM optimizer that integrates the gradient descent with momentum technique with the root mean square propagation (RMSprop) method. In the model, the number of epochs and mini-batches was set to 50 and 32, respectively. To circumvent the model overfitting issue with the train set, we further divided the train set (70% sample points) into train and validation sets implementing a 10-fold cross-validation technique so that the prediction accuracy on the test set (30% sample points) gets maximized.

We hybridized two DNN models: ADAM – ReLU – Sigmoid – DNN and ADAM – ReLU – Softmax – DNN, using two approaches that are L2 regularization and dropout technique to reduce the high variance in the models. For L2 regularization, we specified regularization as the parameter in each of the three hidden layers and set the value of λ to 0.001. For dropout, we added an extra layer after each of the three hidden layers and set the value of κ to 0.6.

iv. Standalone and hybridized SVR models

We developed and evaluated six standalone and hybridized SVR models for predicting flood susceptibility. First, the baseline SVR model was developed and combined with four different kernel functions (e.g., linear, gaussian RBF, laplacian RBF, and MLP kernels) separately. The grid search algorithm and GA were used for hyperparameter tuning and hybridization.

The objective of SVR is to generate function, describing correlation between input and output mentioned in equation (7).

$$f(x) = w^T \psi(x) + bias \tag{7}$$

where, $x \epsilon R^n$ indicates flood conditioning features, $w \in R^n$ represents weight vector, and non-linear mapping function is denoted by $\psi(x)$. The final solution to the constrained optimization problem in SVR using Lagrangian formulation is described in equation (8).

$$f(x) = \sum_{j=1}^{n} (\alpha_j - \alpha_j{}^*) k(x, x_j) + bias \tag{8}$$

where, α_j and $\alpha_j{}^*$ denote the Lagrangian multipliers and $k(x_m, x_n) = < \psi(x_m), \psi(x_n) >$ indicates the kernel function. Various types of kernel functions could be employed (Rahmati et al. 2020). The linear, gaussian RBF, laplacian RBF and MLP kernels can be described in equations (9)-(12), respectively.

$$k(x, x_j) = sum(x.x_j) \tag{9}$$

$$k(x, x_j) = e^{-\gamma \|x - x_j\|^2} \tag{10}$$

$$k(x, x_j) = e^{-\frac{\|x - x_j\|}{\gamma}} \tag{11}$$

$$k(x, x_j) = tanh\ (Ax^T x_j + B) \tag{12}$$

where, γ is an optimizing hyperparameter indicating the spread of the kernel. A is the scale value and B is the offset value. The prediction accuracy of SVR model also depends on other parameters, that are, epsilon, ε representing approximation quality and the cost value that determines the tradeoff between model complexity and training error.

In the standalone SVR model, we have set epsilon to 0.1, cost to 1, and gamma to 0.1. For gaussian RBF – SVR, we optimized gamma and cost using the grid search technique in combination with the 10-fold cross-validation technique while setting epsilon to 0.1. We searched from 0.1 to 2 (interval = 0.1) to find the optimal value of gamma. The optimal value of cost was searched from 0.1 till 10 (interval = 0.1) using a grid search algorithm. This resulted in generating and training a total of 2000 SVR models with different values of gamma and cost. The optimal parameter values derived from the grid search technique produce the least mean squared error (MSE) on the test dataset. Using GA, we optimized the parameters of GA – Linear – SVR (i.e., epsilon and cost), GA – Gaussian RBF – SVR and GA – Laplacian RBF – SVR (i.e., epsilon, cost, and gamma), and GA – MLP – SVR (i.e., epsilon, cost, scale, and offset). The negative quantity of the MSE on the test set prediction was defined as the objective function of GA as we maximized the objective function. Again, a 10-fold cross-validation technique was employed while training all the SVR models on the train set to reduce overfitting.

v. Conventional ML models

This study also developed three conventional ML models: conditional inference tree, KNN, and MLP models. The conditional inference tree is a distinct type of decision tree model that employs recursive partitioning of the dependent variables depending on the correlation values to avoid biasing. This model exploits a significance test to choose the input variables rather than choosing the variable maximizing the information measure. We set the values of the minimum criterion and split to 0.95 and 200, respectively. KNN is a supervised ML model that assumes the similarity or resemblance between the novel case and the known or available cases and consequently puts the novel case into the class or category most similar to the available classes or categories (Costache et al. 2020a). We experimented with different values for *k* in the KNN model. However, the model performed better for a *k* value of five. MLP is another supervised ML model that provides a very fundamental feedforward neural network architecture utilized for both classification and regression-based problems (Ahmadlou et al. 2021). In the architecture of MLP, we used two hidden layers with the first layer containing a total of ten nodes and the second layer containing a total of three nodes. We set the values of the threshold to 0.1 and the maximum steps for training to 10^6. We used RPROP+ as the learning algorithm for MLP.

vi. Validation and comparison of models

For identifying the best performing flood susceptibility model, this study estimated values of various cutoff-dependent and cutoff-independent validation indicators using the ‘roc’ and ‘plot.roc’ functions of ‘pROC’ package in R. The indices include receiver operating characteristic (ROC) and area under the receiver operating characteristic (AUROC) curves, kappa statistic, overall accuracy (OA), positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, and MSE. We used Youden’s index for estimating the optimal cutoff point (Youden 1950) and binarized the predicted flood susceptibility scores by the models (Adnan et al. 2020b). We also estimated the seed cell area index (SCAI) (Akay 2021) values for validation and comparison of flood susceptibility, hazard, exposure, vulnerability, and risk models.

vii. Flood susceptibility map

Applying the best-performing flood prediction model, a flood susceptibility map of Bangladesh was developed using the ArcGIS 10.8 software. The susceptibility values were normalized on a 0-1 scale. The resultant flood susceptibility map was categorized into five classes using the equal interval method in GIS: Very Low (0 – 0.2), Low (0.2 – 0.4), Medium (0.4 – 0.6), High (0.6 – 0.8), and Very High (0.8 – 1) (Rahman et al. 2019).

Flood hazard modeling

Flood hazard in the study area was estimated using equation (13) (Pham et al. 2021a, Pham et al. 2021b).

$$Flood\ Hazard\ Score = A_1 \times Flood\ Susceptibility\ Score + B_1 \times Flood\ Depth + C_1 \times Rainfall \tag{13}$$

where, A_1, B_1, and C_1 are the weights of flood susceptibility, flood depth, and rainfall, respectively. Although previous studies reported the efficacy of the classical AHP tool in modeling flood hazards (Pham et al. 2021a, Pham et al. 2021b), this study utilized a fuzzy AHP model (Zadeh 1996) due to its higher prediction accuracy (Büyüközkan and Feyzıoğlu 2004). First, fuzzy pairwise comparison matrices of the criteria and sub-criteria were developed using the triangular fuzzy numbers (TFN) of the scale of Saaty on relative importance (Ekmekcioğlu et al. 2021). Then weights of different criteria and the local weights of their sub-criteria were generated (Liou and Wang 1992). We also conducted a pairwise comparison of each alternative against every sub-criterion. Global weights of all sub-criterion were estimated by multiplying the weight of each criterion by their local weights. The flood susceptibility parameter was given the most importance, followed by rainfall and flood depth (Pham et al. 2021a). The higher values of all these three criteria indicate a higher flood hazard score. The validity of the weights was checked by ensuring a consistency ratio of less than 10%, where the consistency ratio is defined in equations (14)-(15) (Liou and Wang 1992).

$$Consistency\ Index = \frac{\lambda_{max} - k}{k - 1} \tag{14}$$

$$Consistency\ Ratio = \frac{Consistency\ Index}{Random\ Index} \tag{15}$$

where λ_{max} denotes the highest eigenvalue that belongs to the decision matrix and k is the number of criteria. We set a random index value consistent with the study of Saaty and Tran (2007). The optimism index was set to 80%. Finally, a weighted sum method was employed in equation (13) to estimate a flood hazard score.

Flood exposure modeling

The flood exposure score was estimated using equation (16) (Pham et al. 2021a, Pham et al. 2021b).

$$Flood\ Exposure\ Score = A_2 \times Distance\ to\ River + B_2 \times LULC + C_2 \times Population\ Density \tag{16}$$

where A_2, B_2, and C_2 are the weights of distance to river, LULC, and population density, respectively. For designing fuzzy pairwise comparison matrices of criteria and sub-criteria for flood exposure modeling, the population density parameter was prioritized for its positive association with exposure (Pham et al. 2021b), followed by LULC and distance to river (Pham et al. 2021b).

Flood vulnerability modeling

The flood vulnerability score can be defined in equation (17) (Brito et al. 2018, Pham et al. 2021a, Pham et al. 2021b).

$$Flood\ Vulnerability\ Score = A_3 \times Road\ Density + B_3 \times Age + C_3 \times Poverty\ (Wealth\ Index) \tag{17}$$

where A_3, B_3, and C_3 are the generated weights of road density, age, and poverty (wealth index) respectively utilizing the fuzzy AHP model. Here, poverty (wealth index) was given the highest preference (Pham et al. 2021a), followed by age and road density.

Flood risk modeling

After estimating flood hazard, exposure, and vulnerability scores using fuzzy AHP models, we normalized their scores on a 0-1 scale. Finally, the flood risk map of Bangladesh was derived using equation (4) in GIS. In this study, all fuzzy AHP models were established using MATLAB R2020a software.

Sensitivity analysis of flood causative factors

This study performed a sensitivity analysis of all the flood causative factors in modeling flood susceptibility, hazard, exposure, vulnerability, and risk by estimating their importance rank using the random forest (RF) function. The %IncMSE and IncNodePurity indicators were exploited to rank the flood causative factors, estimated using the 'randomForest' package in R. The %IncMSE measures the upsurge in the MSE value of model prediction when the values of a feature are randomly permuted. The IncNodePurity indicates the total reduction of node impurities estimated by the Gini Index from variable splitting averaged over all the decision trees. The higher the values of %IncMSE and IncNodePurity suggest greater importance of a feature in the model — a greater sensitivity (Rahmati et al. 2020, Siam et al. 2021a).

Results

Flood susceptibility assessment

Standalone and hybridized DNN models

Figure 4 shows the variation of the loss and accuracy metrics over the progression of 50 epochs in each of the six DNN models on the train and validation datasets. The L2 – ADAM – ReLU – Softmax – DNN model is found to be the best-performed model for the train set, with an accuracy value of 0.8892. However, the ADAM – ReLU – Sigmoid – DNN model yielded the highest accuracy (0.8196) with validation data.

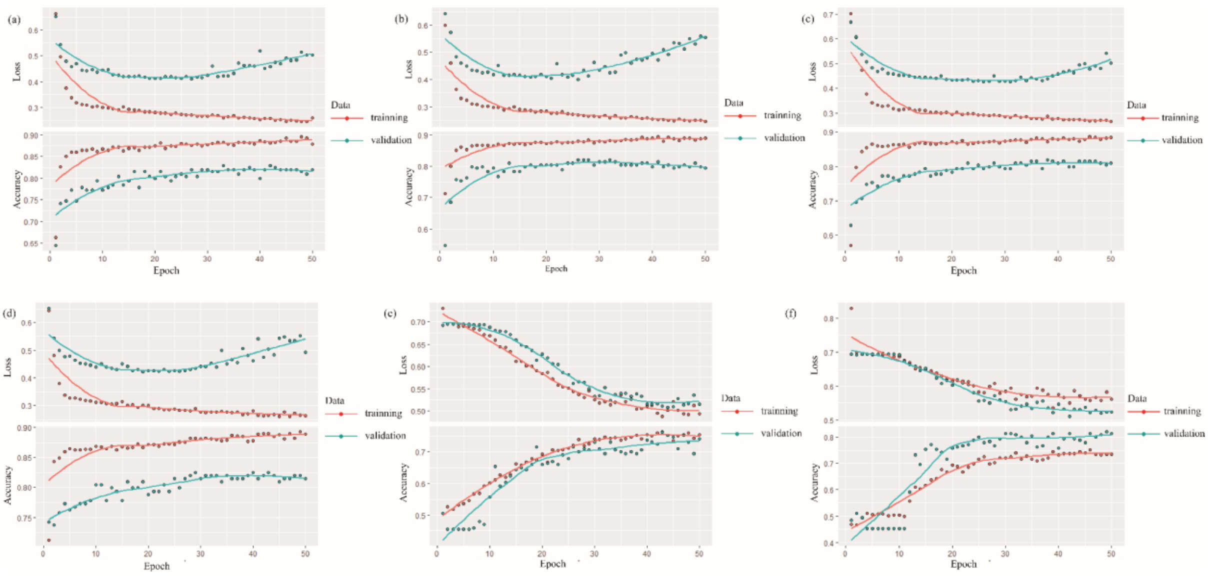


Figure 4. Variation of train loss and accuracy, validation loss and accuracy over the number of epochs for: (a) ADAM – ReLU – Sigmoid – DNN, (b) ADAM – ReLU – Softmax – DNN, (c) L2 – ADAM – ReLU – Sigmoid – DNN, (d) L2 – ADAM – ReLU – Softmax – DNN, (e) Dropout – ADAM – ReLU – Sigmoid – DNN and (f) Dropout – ADAM – ReLU – Softmax – DNN models.

Standalone and hybridized SVR models

In this study, the number of support vectors is found to be 1650, 1266, 1086, 1335, 756, and 978 for standalone SVR, Gaussian RBF – SVR, GA – Gaussian RBF – SVR, GA – Laplacian RBF – SVR, GA – MLP – SVR, and GA – Linear – SVR models, respectively, during the training phase. This indicates that the MLP kernel reduces the complexity of the SVR model more compared to other kernels. The algorithm settings and solutions of GA – Gaussian RBF – SVR, GA – Laplacian RBF – SVR, GA – MLP – SVR, and GA – Linear – SVR are shown in Table 3.

Table 3. Settings and results of hybridized SVR models

	Criteria →			**GA-Gaussian RBF-SVR**	**GA-Laplacian RBF-SVR**	**GA-MLP-SVR**	**GA-Linear-SVR**
GA Settings	Type			Real value	Real value	Real value	Real value
	Population size			50	50	50	50
	Number of generations			100	100	100	100
	Elitism			2	2	2	2
	Crossover probability			0.8	0.8	0.8	0.8
	Mutation probability			0.1	0.1	0.1	0.1
	Search domain	Epsilon	Lower	0	0	0	0
			Upper	1	1	1	1
		Gamma	Lower	0.0010	0.0010	-	-
			Upper	2.0000	2.0000	-	-
		Cost	Lower	0.0001	0.0001	0.0001	0.0001
			Upper	10	10	10	10
		Scale	Lower	-	-	0.00001	-
			Upper	-	-	1	-
		Offset	Lower	-	-	-10	-
			Upper	-	-	-0.00001	-
GA Results	Iterations			100	100	100	100
	Fitness function value			-0.0922	-0.0904	-0.1164	-0.1186

Solution	Epsilon	0.1737	0.1020	0.6959	0.4982
	Gamma	0.1149	0.2989	-	-
	Cost	1.4114	9.3438	8.1556	5.1641
	Scale	-	-	0.1435	-
	Offset	-	-	-3.2567	-

Figure 5 (a) shows the performance of all the trained Gaussian – RBF – SVR models using grid search in a contour plot where values of gamma are shown along the *x*-axis and values of cost are in the *y*-axis while the *z*-axis shows corresponding MSE. The optimal value of cost is 1.10 while the optimal gamma value is 0.10 for the best Gaussian RBF – SVR model with an MSE of 0.0925 from the grid search result. In the best Gaussian RBF – SVR model, weight values of slope, distance to river, drainage density, elevation, SPI, soil texture, soil permeability, LULC, geology, curvature, and aspect are -30.03, -40.58, 8.15, -47.18, 10.20, 18.70, 25.02, -2.52, 2.96, 0.09 and -0.08, respectively, where the bias is 0.39. The fitness values of the other four hybridized SVR models are shown in Figure 5 (b–e).

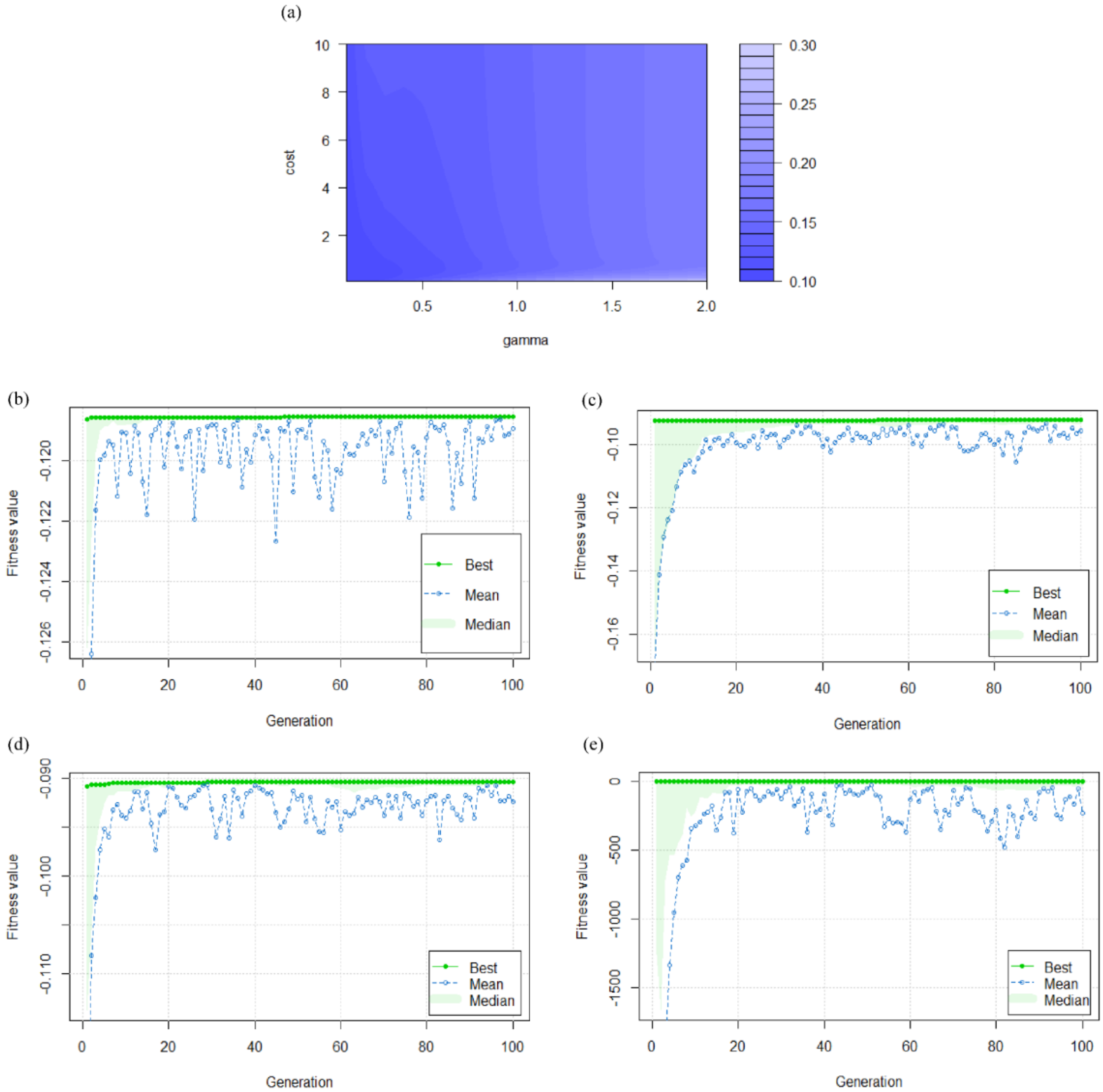


Figure 5. (a) Performance of all the trained Gaussian RBF – SVR models. Variation of fitness value over the number of generations in (b) GA – Linear – SVR, (c) GA – Gaussian RBF – SVR, (d) GA – Laplacian RBF – SVR and (e) GA – MLP – SVR models.

Conventional ML models

Among the other three ML models employed, the conditional inference tree performed better than KNN and MLP models in terms of fitting the train data more accurately. Figure 6 illustrates the fitted conditional inference tree on the train set.

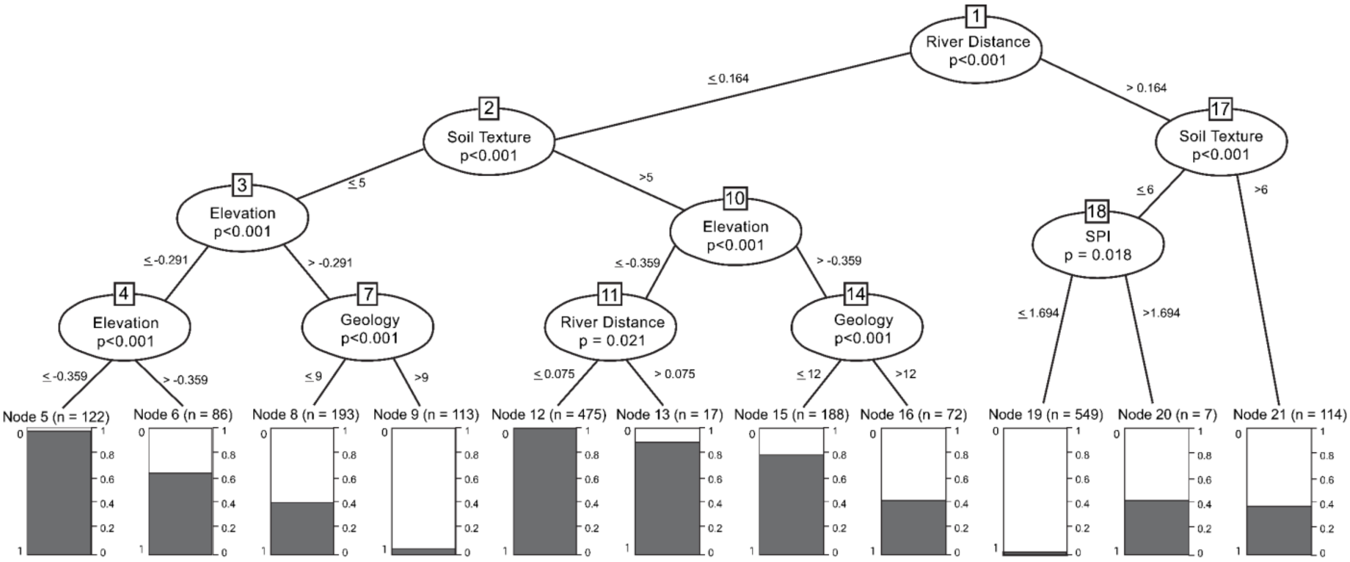


Figure 6. Conditional inference tree based on train set

Model validation and comparison

This study compares all fifteen ML models to select the best-performed model for flood susceptibility mapping in Bangladesh. Figure 7 illustrates the ROC curves of all models based on the test set.

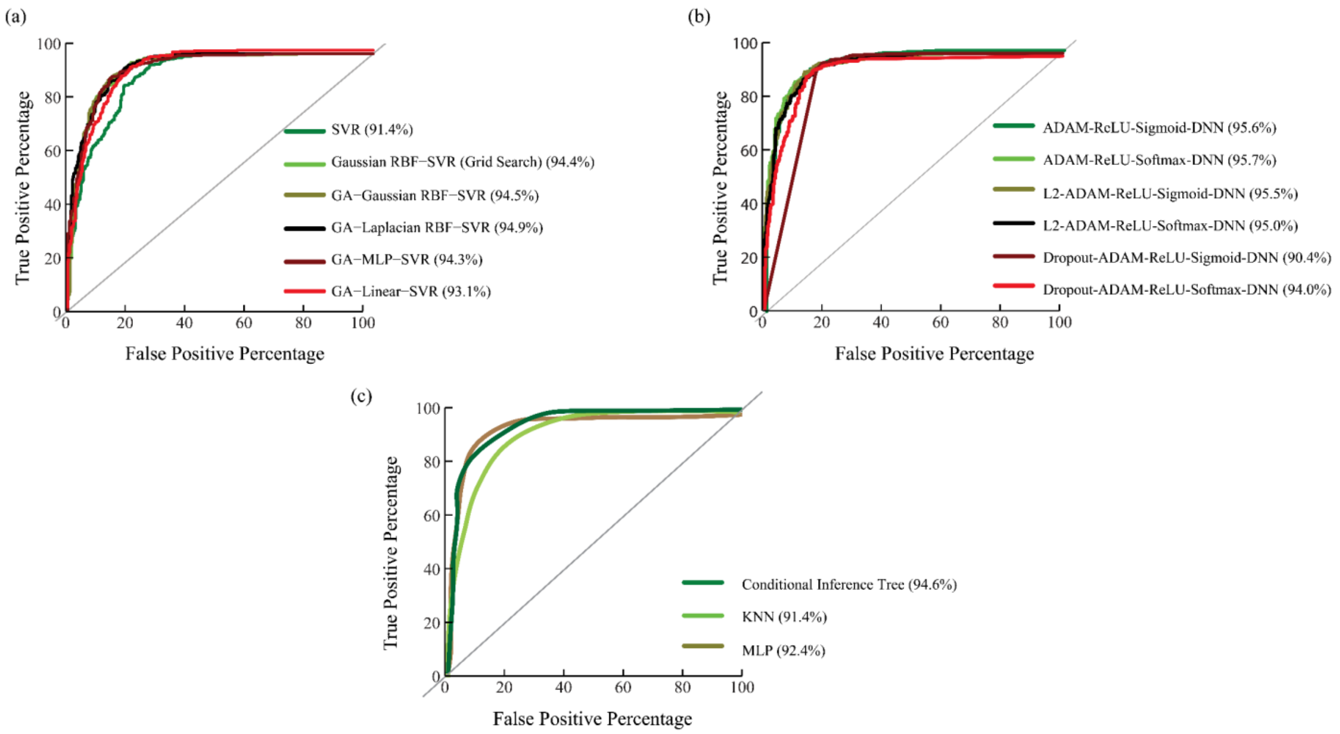


Figure 7. Validation of (a) the standalone and hybridized SVR, (b) standalone and hybridized DNN and, (c) other machine learning models using the ROC curves

The ADAM – ReLU – Softmax – DNN model yields the highest prediction accuracy, with an AUROC value of 95.7%, followed by the ADAM – ReLU – Sigmoid – DNN model (AUROC - 95.6%) and the L2 – ADAM – ReLU – Sigmoid – DNN model (AUROC - 95.5%) (Figure 7 (b)). A total of four DNN models have an AUROC greater than or equal to 95%. Contrarily, SVR models have relatively a lower prediction accuracy, where the GA – Laplacian RBF – SVR model obtained the highest AUROC value of 94.9% (Figure 7 (a)). In the case of conventional ML models, the conditional inference tree obtained the highest AUROC value of 94.6% (Figure 7 (c)). Model comparison results indicate a higher efficacy of the DNN models over the other models in estimating flood susceptibility. Table 4 presents the outcomes of performance assessment of different models.

Table 4. Model performance using different statistical indices

Models	Cutoff	AUROC	OA	Kappa	Sensitivity	Specificity	PPV	NPV	MSE
ADAM-ReLU-Sigmoid-DNN	0.697	0.956	0.893	0.785	0.911	0.874	0.884	0.903	0.087
ADAM-ReLU-Softmax-DNN	0.507	0.957	0.894	0.788	0.929	0.857	0.872	0.920	0.083
L2- ADAM-ReLU-Sigmoid-DNN	0.603	0.955	0.898	0.795	0.927	0.867	0.880	0.919	0.084
L2- ADAM-ReLU-Softmax-DNN	0.848	0.950	0.883	0.766	0.894	0.872	0.880	0.887	0.108
Dropout- ADAM-ReLU-Sigmoid-DNN	0.618	0.904	0.887	0.773	0.960	0.810	0.841	0.951	0.117

Dropout- ADAM-ReLU-Softmax-DNN	0.429	0.940	0.892	0.783	0.941	0.840	0.860	0.932	0.140
SVR	0.554	0.914	0.847	0.693	0.878	0.815	0.833	0.864	0.126
Gaussian RBF-SVR	0.572	0.944	0.879	0.759	0.913	0.844	0.860	0.902	0.093
GA-Gaussian RBF-SVR	0.582	0.945	0.884	0.768	0.906	0.862	0.873	0.897	0.092
GA-Laplacian RBF-SVR	0.394	0.949	0.881	0.761	0.944	0.815	0.842	0.932	0.090
GA-MLP-SVR	0.525	0.943	0.883	0.766	0.908	0.857	0.869	0.899	0.116
GA-Linear -SVR	0.496	0.931	0.866	0.732	0.934	0.795	0.827	0.920	0.119
Conditional Inference Tree	0.639	0.946	0.869	0.740	0.812	0.931	0.925	0.825	0.087
KNN	0.600	0.914	0.842	0.684	0.873	0.810	0.828	0.859	0.114
MLP	0.633	0.924	0.879	0.759	0.915	0.842	0.859	0.905	0.108

The L2 – ADAM – ReLU – Sigmoid – DNN model obtains the highest OA value of 0.898 and a kappa statistic of 0.795, followed by the ADAM – ReLU – Softmax – DNN (OA = 0.894 and kappa = 0.788) and ADAM – ReLU – Sigmoid – DNN (OA = 0.893 and kappa = 0.785) models. However, the ADAM – ReLU – Softmax – DNN model achieves the lowest MSE value of 0.083, followed by the L2 – ADAM – ReLU – Sigmoid – DNN (MSE = 0.084) and ADAM – ReLU – Sigmoid – DNN (MSE = 0.087) models. Based on the AUROC, OA, kappa statistic, and MSE metrics together, this study identifies the L2 – ADAM – ReLU – Sigmoid – DNN and the ADAM-ReLU-Softmax-DNN models as the best two models for flood susceptibility mapping. However, the estimated SCAI values (Table 7) of flood susceptibility indicate that the hybridized L2 – ADAM – ReLU – Sigmoid – DNN model outperforms the ADAM-ReLU-Softmax-DNN model. Therefore, this study uses the hybridized L2 – ADAM – ReLU – Sigmoid – DNN model for mapping flood susceptibility in Bangladesh.

Flood hazard assessment

Figure 8 (b) shows the resultant flood hazard map. Among three criteria of flood hazard, flood susceptibility received the highest weight, followed by rainfall and flood depth (Table 5). About 20% of the total area is estimated to be flood hazard-prone zones of moderate to very high levels of severity. Southwestern and northeastern Bangladesh, as well as areas adjacent to major rivers, are high hazard zones (Figure 8 (b)). The SCAI of high and very high classes in the hazard map is the lowest, with values of 0.53 and 0.59, respectively (Table 6).

Flood exposure assessment

Figure 8 (c) shows the flood exposure map of Bangladesh. About 40% of the country is categorized as moderate to very high magnitudes. Among the three variables (distance to river, LULC, and population density), the estimated weight for population density is the highest (Table 5). Unsurprisingly, areas characterized by high population density are highly exposed to flooding.

Flood vulnerability assessment

The flood vulnerability map is shown in Figure 8 (d). Results show that about 69% of Bangladesh is vulnerable (moderate to very high) to flooding. The highest weight for the parameter wealth index (WI) (Table 5) indicates that the economic status of the people is one of the major determining flood vulnerability factors. Areas characterized by a low wealth index are highly vulnerable to flooding.

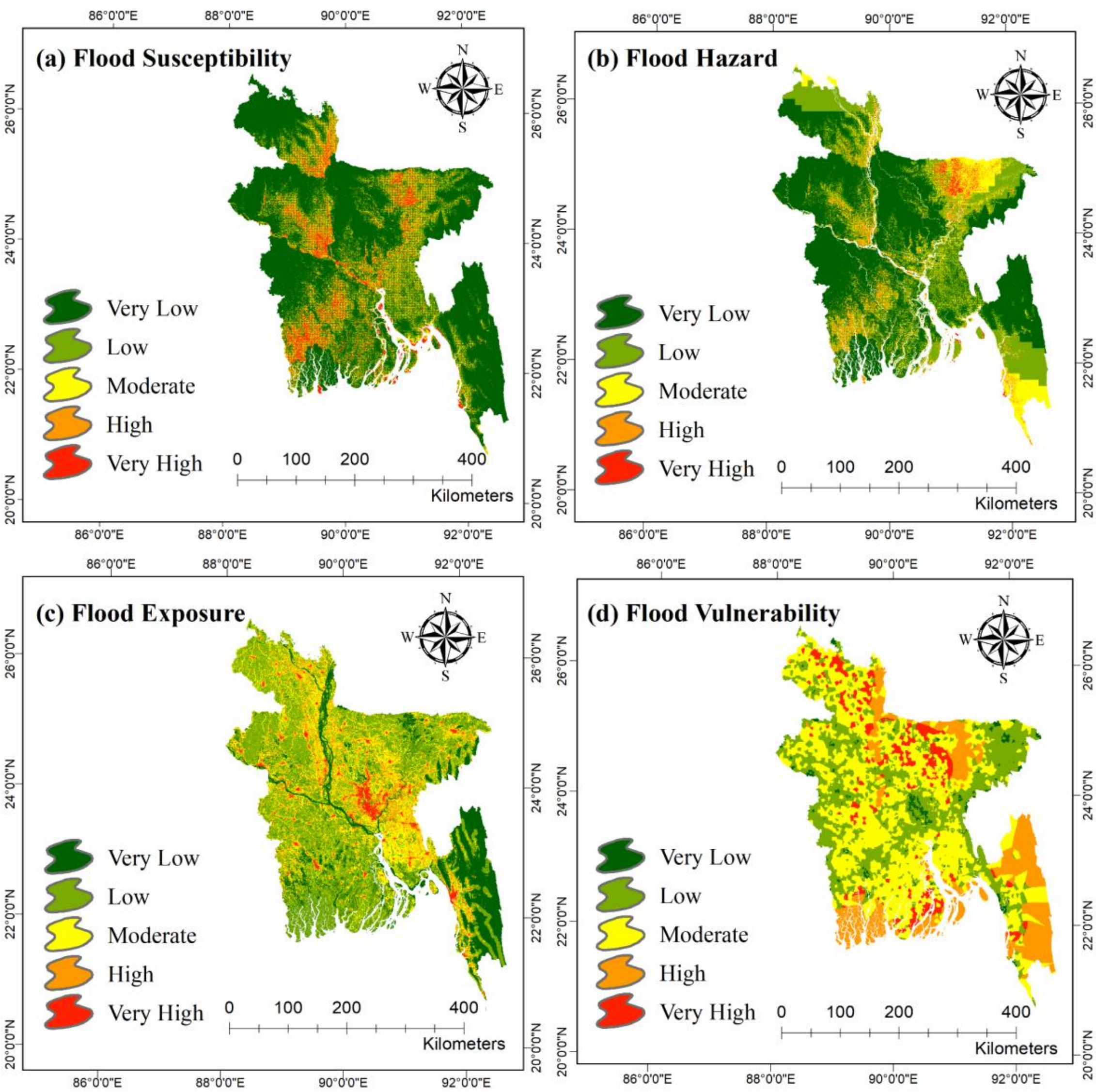


Figure 8. (a) Flood susceptibility, (b) flood hazard, (c) flood exposure and (d) flood vulnerability maps of Bangladesh

Flood Risk Assessment

Table 5 exhibits weights of criteria as well as sub-criteria for flood hazard, exposure, and vulnerability. Local weights indicate the type of association that exists between floods and various risk indicators. For instance, flood susceptibility, flood depth, rainfall, population density, road density, and age are positively associated with flood risk. On the other hand, distance to river and wealth index are negatively correlated. In the case of LULC, built-up areas and croplands are highly prone to flood risk, particularly in areas with high flood potentials. In the case of the SCAI results, moderate to very high flood risk zones yield relatively low SCAI values. These results indicate a good agreement between the observed flood locations and modeled flood risk zones.

Table 5. Weights of criteria as well as sub-criteria generated by fuzzy AHP method

Component	Criteria	Weight	Class	Sub – criteria	Local Weight	Global Weight
Flood Hazard	Flood susceptibility	0.6037	0 - 0.2	Very Low	0.0309	0.0187
			0.2 - 0.4	Low	0.0843	0.0509
			0.4 - 0.6	Moderate	0.1698	0.1025
			0.6 - 0.8	High	0.2870	0.1733
			0.8 - 1	Very High	0.4280	0.2584
	Flood depth	0.1003	No Flooding	1	0.0412	0.0041
			<0.30	2	0.0757	0.0076
			0.30 - 1.83	3	0.1223	0.0123
			1.83 - 3.05	4	0.2950	0.0296
			>3.05	5	0.4658	0.0467
	Rainfall	0.2960	245.4 - 333.7	1	0.0475	0.0141
			333.8 - 435.8	2	0.0870	0.0258
			435.9 - 560	3	0.1408	0.0417
			560.1 - 725.5	4	0.2770	0.0820
			725.6 – 949.02	5	0.4476	0.1325
Flood Exposure	Distance to river	0.0918	0 - 432	1	0.4199	0.0385
			432 - 1297	2	0.2597	0.0238
			1297 - 2594	3	0.1922	0.0176
			2594 - 4899	4	0.0937	0.0086
			4899 - 36890	5	0.0345	0.0032
	LULC	0.3727	Water	1	0.0321	0.0120
			Bare Land	2	0.0871	0.0325
			Vegetation	3	0.2213	0.0825
			Crops	4	0.2897	0.1080
			Built Area	5	0.3698	0.1378
	Population density (Population per cell)	0.5355	0-1	1	0.0298	0.0160
			1-2	2	0.1104	0.0591
			2-3	3	0.1579	0.0846
			3-6	4	0.2785	0.1491
			6-370	5	0.4234	0.2267
Flood Vulnerability	Road density	0.0859	0-0.9	1	0.0375	0.0032
			0.9 - 1.3	2	0.1046	0.0090
			1.3 - 1.6	3	0.1601	0.0138
			1.6 - 1.9	4	0.2287	0.0196
			1.9 - 3.4	5	0.4692	0.0403

	Age (< 14 and > 60)	0.2643	0-1	1	0.0395	0.0104
			1-2	2	0.1110	0.0293
			2-3	3	0.1700	0.0449
			3-6	4	0.2433	0.0643
			6-101	5	0.4362	0.1153
	Poverty (Wealth index)	0.6498	-1.2 - -0.61	1	0.4141	0.2691
			-0.6 - -0.3	2	0.2492	0.1619
			-0.29 - 0.07	3	0.1797	0.1168
			0.071 - 0.64	4	0.1237	0.0804
			0.65 - 2.2	5	0.0334	0.0217

Table 6: SCAI measurements of flood susceptibility, exposure, hazard, vulnerability and risk maps

Class	**Flood Susceptibility** (L2-ADAM-ReLU-Sigmoid-DNN)	**Flood Susceptibility** (ADAM-ReLU-Softmax-DNN)	**Flood Exposure**	**Flood Hazard**	**Flood Vulnerability**	**Flood Risk**
Very Low	1.56	3.24	1.56	1.49	1.20	1.40
Low	0.60	2.09	0.85	0.81	0.93	0.96
Moderate	0.63	1.92	0.97	0.66	0.91	0.66
High	0.53	1.89	0.97	0.53	1.55	0.59
Very High	0.56	0.68	3.17	0.59	1.04	0.67

Table 7 represents the consistency ratio for each component and criteria which is less than 10% i.e., acceptable in each case.

Table 7. Consistency ratio for flood risk components

Component	**Consistency ratio (%)**	**Criteria**	**Consistency ratio (%)**
Flood hazard	8.70	Flood susceptibility	9.88
		Flood depth	8.93

		Rainfall	6.17
Flood exposure	8.11	Distance to river	8.02
		LULC	7.79
		Population density	9.94
Flood vulnerability	4.19	Road density	7.61
		Age	5.88
		Poverty (Wealth index)	7.99

Figure 8 (a–d) illustrates the predicted flood susceptibility, flood hazard, flood exposure, and flood vulnerability maps of Bangladesh. The flood risk map obtained in this study is shown in Figure 9. About 20.45% of the area is categorized as flood risk zones, where the percentages of moderate, high, and very high flood risk-prone zones are 13.37%, 5.44%, and 1.64%, respectively. The northeastern region of Bangladesh, as well as areas near the GBM rivers, have high flood damage potential.

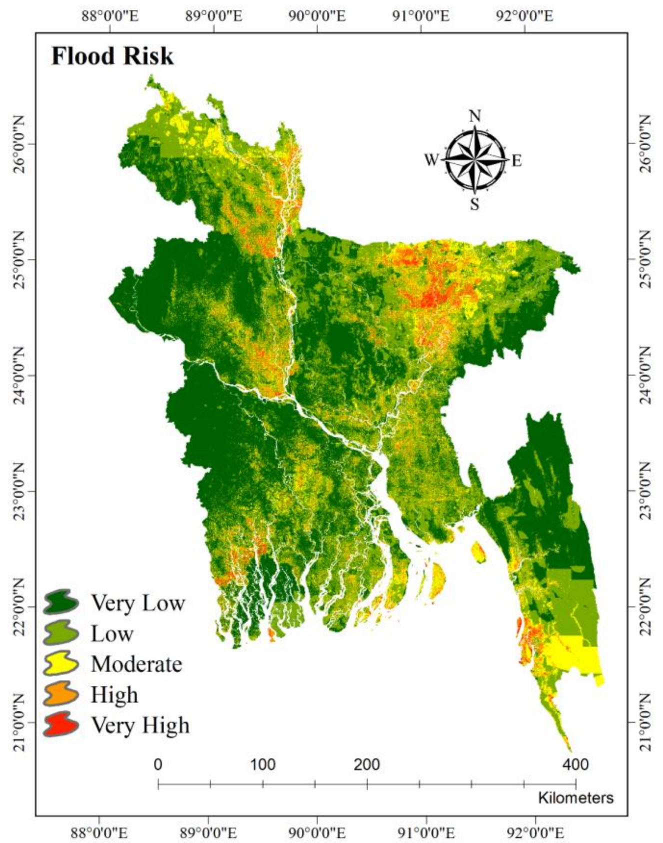


Figure 9. Flood risk map of Bangladesh

Figure 10 shows the percent of flood risk areas in a few districts where floods affected a significant number of people in 2020. For instance, in the Kurigram district, a total of 227,440 people (10.4% of the total population of Kurigram) were affected during monsoon flooding in 2020. This study found that about 52.95% of the total area of Kurigram district is a flood risk zone of moderate to very high severity. Similarly, in other northern districts such as Gaibandha, Nilphamari, and Ranpur, a significant number of people were flood-affected.

This study also found highly risk-prone regions. In the case of northeastern Bangladesh, districts such as Sunamganj and Netrakona are in this risk zone, with damage potential of 64.43% and 65.38%, respectively. In these two districts, a total of 113,237 and 84,300 people were inflicted by floods in 2020 (CARE 2020).

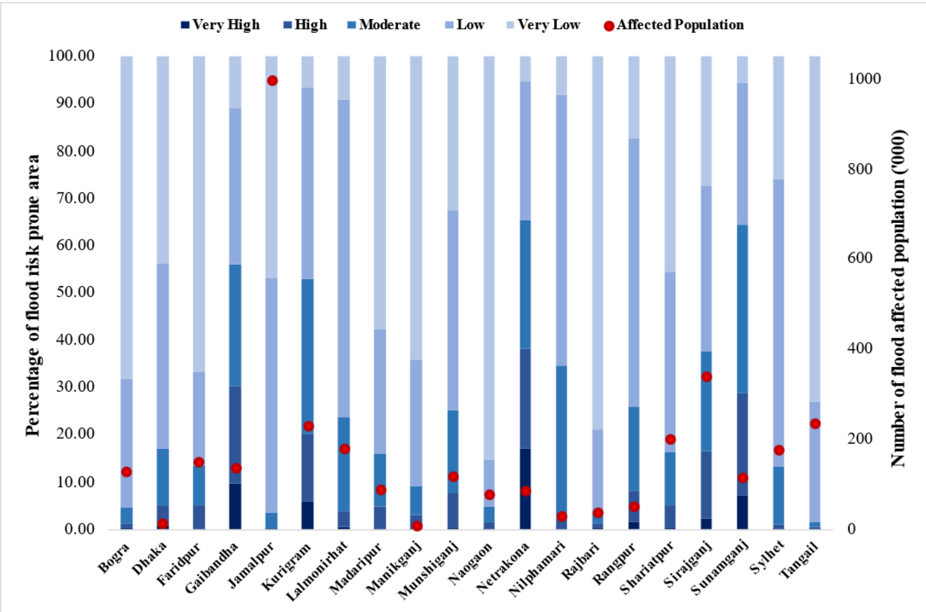


Figure 10. Percentage of flood risk prone areas in different districts

Sensitivity analysis results

This study estimates the sensitivity of all corresponding factors in modeling flood susceptibility, hazard, exposure, vulnerability, and risk with respect to %IncMSE and IncNodePurity scores provided by RF. The flood susceptibility model is highly sensitive to factors such as elevation and distance to rivers (Figure 11 (a - b)). In the case of flood hazard, flood susceptibility is the most significant parameter (Figure 11(c - d)). LULC and population density are the imporant factors determining flood exposure (Figure 11(e - f)). In the case of flood vulnerability, poverty is the most influential factor (Figure 11(g - h)). Finally, this study notes that flood risk is sensitive to flood hazard (Figure 11(i - j)). A recent study (Adnan et al. 2020a) validates the results of flood exposure, vulnerability, and risk.

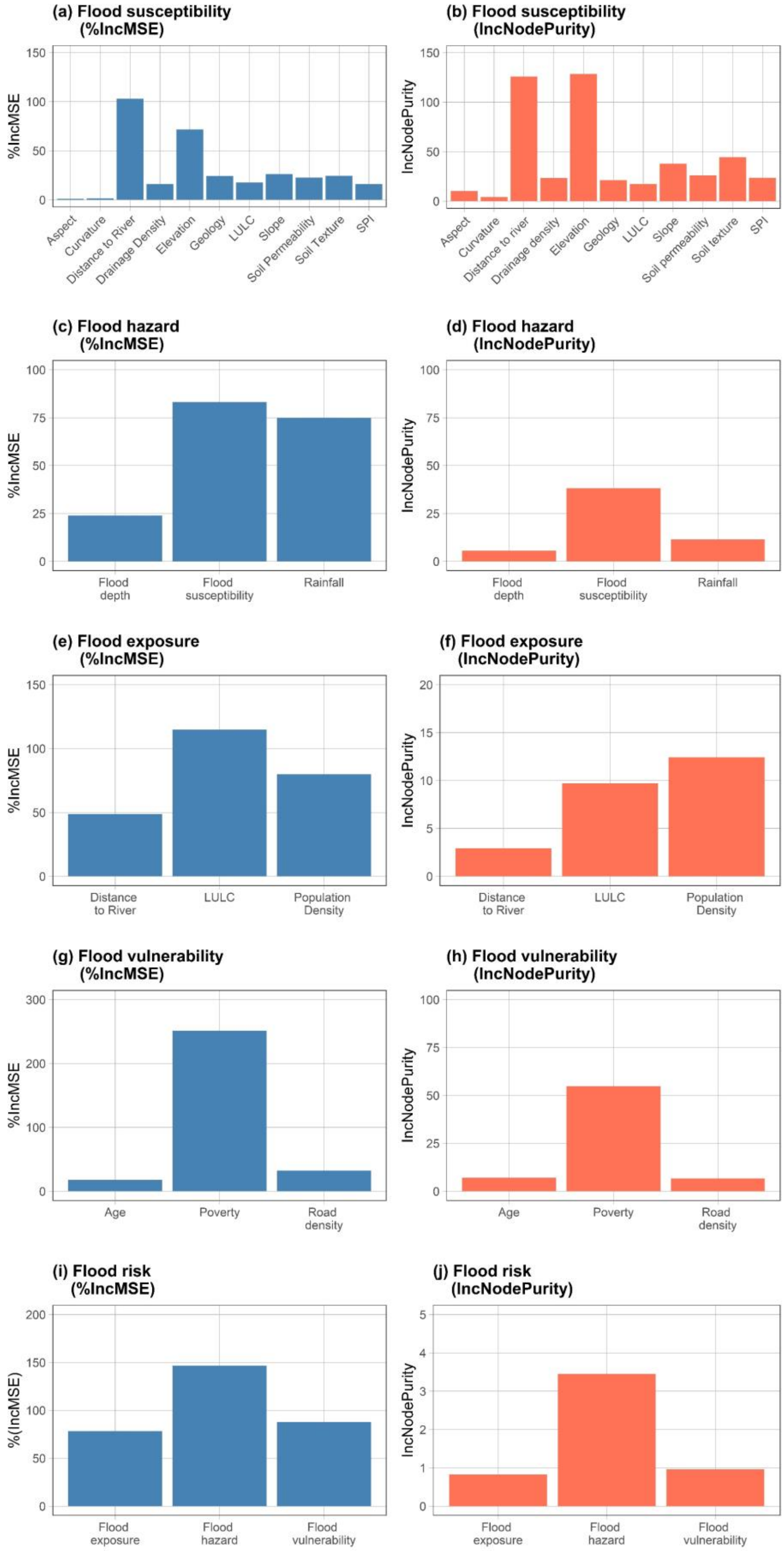


Figure 11. Sensitivity analysis of flood causative factors in modeling flood susceptibility, hazard, exposure, vulnerability and risk based on %IncMSE and IncNodePurity.

Discussion

This study aimed to present a flood risk assessment framework using hybridized DNN and fuzzy AHP models, hypothesizing that the use of hybridized models would improve the accuracy of flood risk models. Hence, we developed and evaluated the performance of fifteen models including twelve standalone and hybridized ML models and three conventional ML models. The results exhibit the efficacy of the hybridized DNN architectures over all other models. This is a first attempt to combine hybridized DNN architectures with fuzzy AHP models to assess flood risk in a complex flood regime like deltaic Bangladesh.

In the case of flood susceptibility, elevation and distance to river were found as the most influential factors influencing flood potentials. Both these findings are supported by other recent studies (Wang et al. 2019, Rahmati et al. 2020, Chou et al. 2021, Pham et al. 2021a, Pham et al. 2021b). This study established a total of fifteen flood susceptibility models that produced an AUC value of more than 90%, indicating an excellent prediction accuracy (Arabameri et al. 2019). Flood susceptibility map produced using the hybridized L2 – ADAM – ReLU – Sigmoid – DNN model (Figure 8 (a)) yielded the highest prediction accuracy, resulting in a good agreement with the flood inundation map of Bangladesh in 2020.

The flood susceptibility map produced in this study showed that the northeastern part of Bangladesh is highly susceptible, including Netrokona, Sunamganj, Kishoreganj, and Mymensingh. These districts are also in high-risk-prone zones. All these districts include large water bodies (locally known as "*Haor*") and faced severe flooding in the last couple of years. These districts are also characterized by a low slope and elevation. A recent study reported that areas with a lower slope and elevation have greater flood damage potential (Adnan et al. 2020b). On the contrary, districts in the southeastern zone such as Khagrachori and Banderbans are characterized by high elevation areas and low-density population; hence, pose a relatively low risk. These districts mostly remained inundation-free during the flood events of 2020 (Figure 2). This finding is in accord with other studies that noted that elevation has an inverse relationship with flooding in general (Rahman et al. 2021b). The flood risk map produced in this study showed that several districts in northern and northeastern parts of Bangladesh are located in a high-risk zone, where a significant number of people were affected during the 2020 flood event. Previous studies also reported that the flood potentials of these districts are very high primarily due to their proximity to major rivers (Rahman et al. 2019, Siam et al. 2021a). This finding is also consistent with studies that mentioned that areas closer to the rivers are highly at risk of flood disaster (Talukdar et al. 2020). This study also noted that flood hazard, vulnerability, and risk models are sensitive to flood susceptibility, poverty, and flood hazard, respectively. Several recent studies (Adnan et al. 2020a, Adnan et al. 2020b, Siam et al. 2021a) validates the results of flood hazard, vulnerability, and risk.

Although the proposed framework resulted in a very high flood risk prediction accuracy, several limitations and uncertainties can be anticipated. First, this study considered only one flood event due to the unavailability of long-term flood observation data at the national level. Second, flood susceptibility, hazard, exposure, and vulnerability indicators' data had differing spatial resolutions. For these reasons, the independent and dependent variables used in this study might be subject to label noise. A recent study has observed negative effects of label noise on the performance of ML-based flood susceptibility modeling (Siam et al. 2021b). Future research can address these limitations by establishing label noise-tolerant standalone and hybridized ML models.

Conclusion

In the present study, a novel approach to flood risk assessment in Bangladesh was developed, combining hybridized DNN and fuzzy AHP methods. Based on various model performance assessment indices, the hybridized L2 – ADAM – ReLU – Sigmoid – DNN model was selected as the best-performed flood susceptibility model. The resultant flood susceptibility map was used to develop a flood hazard map utilizing the fuzzy AHP model. Finally, the flood risk map of Bangladesh was developed by integrating flood hazard, exposure, and vulnerability maps. Despite some uncertainties and limitations, the study promotes the use of hybridized DNN model for spatial flood risk modeling to achieve a country-scale flood risk map. The proposed flood risk assessment framework is expected to be useful for policymakers to better manage flood risk. For future research, this study can be extended to appraise spatiotemporal flood risk assessment using hybridized DNN models.

Acknowledgements

This work is supported by the ICT Innovation Fund (2020-21) provided by the ICT division, Ministry of Post, Telecommunication and Information Technology of the People's Republic of Bangladesh.

Data and code availability statement

The data and codes that support the findings of this study are available from the corresponding author upon reasonable request.

References

Adnan, M.S.G., Abdullah, A.Y.M., Dewan, A. & Hall, J.W., 2020a. The effects of changing land use and flood hazard on poverty in coastal bangladesh. *Land Use Policy,* 99, 104868.

Adnan, M.S.G., Dewan, A., Zannat, K.E. & Abdullah, A.Y.M., 2019a. The use of watershed geomorphic data in flash flood susceptibility zoning: A case study of the karnaphuli and sangu river basins of bangladesh. *Natural Hazards,* 99 (1), 425-448.

Adnan, M.S.G., Haque, A. & Hall, J.W., 2019b. Have coastal embankments reduced flooding in bangladesh? *Science of the total environment,* 682, 405-416.

Adnan, M.S.G., Talchabhadel, R., Nakagawa, H. & Hall, J.W., 2020b. The potential of tidal river management for flood alleviation in south western bangladesh. *Science of The Total Environment,* 731, 138747 %@ 0048-9697.

Ahmadlou, M., Al-Fugara, A.K., Al-Shabeeb, A.R., Arora, A., Al-Adamat, R., Pham, Q.B., Al-Ansari, N., Linh, N.T.T. & Sajedi, H., 2021. Flood susceptibility mapping and assessment using a novel deep learning model combining multilayer perceptron and autoencoder neural networks. *Journal of Flood Risk Management,* 14 (1), e12683.

Akay, H., 2021. Flood hazards susceptibility mapping using statistical, fuzzy logic, and mcdm methods. *Soft Computing,* 25 (14), 9325-9346.

Akay, H. & Baduna Koçyiğit, M., 2020. Flash flood potential prioritization of sub-basins in an ungauged basin in turkey using traditional multi-criteria decision-making methods. *Soft Computing,* 24 (18), 14251-14263.

Arabameri, A., Rezaei, K., Cerdà, A., Conoscenti, C. & Kalantari, Z., 2019. A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in northern iran. *Science of the Total Environment,* 660, 443-458.

Bai, S., Lü, G., Wang, J., Zhou, P. & Ding, L., 2011. Gis-based rare events logistic regression for landslide-susceptibility mapping of lianyungang, china. *Environmental Earth Sciences,* 62 (1), 139-149.

Bannari, A., Ghadeer, A., El-Battay, A., Hameed, N. & Rouai, M., 2017. Detection of areas associated with flash floods and erosion caused by rainfall storm using topographic attributes, hydrologic indices, and gis. *Global changes and natural disaster management: Geo-information technologies.* Springer, 155-174.

Barbour, E.J., Adnan, M.S.G., Borgomeo, E., Paprocki, K., Khan, S.A., Salehin, M. & Hall, J.W., 2022. The unequal distribution of water risks and adaptation benefits in coastal bangladesh. *Nature Sustainability*.

Barc, 2014. Land resource database. *In* (Barc), B.a.R.C. ed. Dhaka, Bangladesh

Bondarenko, M., Kerr, D., Sorichetta, A. & Tatem, A.J., 2020. Estimates of 2020 total number of people per grid square, adjusted to match the corresponding unpd 2020 estimates and broken down by gender and age groupings, produced using built-settlement growth model (bsgm) outputs. UK: University of Southampton.

Brito, M.M.D., Evers, M. & Almoradie, A.D.S., 2018. Participatory flood vulnerability assessment: A multi-criteria approach. *Hydrology and Earth System Sciences,* 22 (1), 373-390.

Bui, D.T., Hoang, N.-D., Martínez-Álvarez, F., Ngo, P.-T.T., Hoa, P.V., Pham, T.D., Samui, P. & Costache, R., 2020a. A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. *Science of The Total Environment,* 701, 134413.

Bui, D.T., Ngo, P.-T.T., Pham, T.D., Jaafari, A., Minh, N.Q., Hoa, P.V. & Samui, P., 2019. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *Catena,* 179, 184-196.

Bui, Q.-T., Nguyen, Q.-H., Nguyen, X.L., Pham, V.D., Nguyen, H.D. & Pham, V.-M., 2020b. Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. *Journal of Hydrology,* 581, 124379.

Büyüközkan, G. & FeyzıoğLu, O., 2004. A fuzzy-logic-based decision-making approach for new product development. *International journal of production economics,* 90 (1), 27-45.

Care, 2020. *Monsoon floods 2020 - coordinated preliminary impact and needs assessment*.

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.

Chou, T.Y., Hoang, T.V., Fang, Y.M., Nguyen, Q.H., Lai, T.A., Pham, V.M., Vu, V.M. & Bui, Q.T., 2021. Swarm-based optimizer for convolutional neural network: An application for flood susceptibility mapping. *Transactions in GIS,* 25 (2), 1009-1026.

Chowdhury, E.H. & Hassan, Q.K., 2017. Use of remote sensing data in comprehending an extremely unusual flooding event over southwest bangladesh. *Natural Hazards,* 88 (3), 1805-1823.

Costache, R., Ali, S.A., Parvin, F., Pham, Q.B., Arabameri, A., Nguyen, H., Crăciun, A. & Anh, D.T., 2021. Detection of areas prone to flood-induced landslides risk using certainty factor and its hybridization with fahp, xgboost and deep learning neural network. *Geocarto International*, 1-36.

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., 2020a. Flash-flood susceptibility assessment using multi-criteria decision making and machine learning supported by remote sensing and gis techniques. *Remote Sensing,* 12 (1), 106.

Costache, R., Popa, M.C., Bui, D.T., Diaconu, D.C., Ciubotaru, N., Minea, G. & Pham, Q.B., 2020b. Spatial predicting of flood potential areas using novel hybridizations of fuzzy decision-making, bivariate statistics, and machine learning. *Journal of Hydrology,* 585, 124808.

Darabi, H., Choubin, B., Rahmati, O., Haghighi, A.T., Pradhan, B. & Kløve, B., 2019. Urban flood risk mapping using the garp and quest models: A comparative study of machine learning techniques. *Journal of hydrology,* 569, 142-154.

David, A. & Schmalz, B., 2020. Flood hazard analysis in small catchments: Comparison of hydrological and hydrodynamic approaches by the use of direct rainfall. *Journal of Flood Risk Management,* 13 (4), e12639.

De Moel, H., Jongman, B., Kreibich, H., Merz, B., Penning-Rowsell, E. & Ward, P.J., 2015. Flood risk assessments at different spatial scales. *Mitigation and Adaptation Strategies for Global Change,* 20 (6), 865-890.

Dewan, A.M. & Kankam-Yeboah, K., 2006. Using synthetic aperture radar (sar) data for mapping river water flooding in an urban landscape: A case study of greater dhaka, bangladesh. *Journal of Japan Society of Hydrology and Water Resources,* 19 (1), 44-54.

Dewan, T.H., 2015. Societal impacts and vulnerability to floods in bangladesh and nepal. *Weather and Climate Extremes,* 7, 36-42.

Ekmekcioğlu, Ö., Koc, K. & Özger, M., 2021. District based flood risk assessment in istanbul using fuzzy analytical hierarchy process. *Stochastic Environmental Research and Risk Assessment,* 35 (3), 617-637.

Hasan, S.S., Deng, X., Li, Z. & Chen, D., 2017. Projections of future land use in bangladesh under the background of baseline, ecological protection and economic development. *Sustainability,* 9 (4), 505.

Huffman, G.J., Stocker, E.F., Bolvin, D.T., Nelkin, E.J. & Tan, J., 2019. Gpm imerg final precipitation l3 half hourly 0.1 degree x 0.1 degree v06, greenbelt, md, goddard earth sciences data and information services center (ges disc). The National Aeronautics and Space Administration (NASA).

Islam, A.R.M.T., Talukdar, S., Mahato, S., Kundu, S., Eibek, K.U., Pham, Q.B., Kuriqi, A. & Linh, N.T.T., 2021. Flood susceptibility modelling using advanced ensemble machine learning models. *Geoscience Frontiers,* 12 (3), 101075.

Islam, M.M. & Sado, K., 2000. Flood hazard assessment in bangladesh using noaa avhrr data with geographical information system. *Hydrological Processes,* 14 (3), 605-620.

Jaxa, 2015. Alos global digital surface model “alos world 3d-30m (aw3d30)”.

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J., Mathis, M. & Brumby, S., Year. Global land use / land cover with sentinel 2 and deep learninged.^eds. *IGARSS 2021-2021 IEEE International Geoscience and Remote Sensing Symposium*, United States: IEEE.

Leon, M.A., Barua, P., Sarker, P., Kumar, P. & Hasan, M., 2020. *Annual flood report 2019*.

Liou, T.-S. & Wang, M.-J.J., 1992. Ranking fuzzy numbers with integral value. *Fuzzy sets and systems,* 50 (3), 247-255.

Lu, Y., Bookman, R., Waldmann, N. & Marco, S., 2020. A 45 kyr laminae record from the dead sea: Implications for basin erosion and floods recurrence. *Quaternary Science Reviews,* 229, 106143.

Luu, C., Von Meding, J. & Mojtahedi, M., 2019. Analyzing vietnam's national disaster loss database for flood risk assessment using multiple linear regression-topsis. *International Journal of Disaster Risk Reduction,* 40, 101153.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G. & Johnson, B.A., 2019a. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS journal of photogrammetry and remote sensing,* 152, 166-177.

Ma, M., Liu, C., Zhao, G., Xie, H., Jia, P., Wang, D., Wang, H. & Hong, Y., 2019b. Flash flood risk analysis based on machine learning techniques in the yunnan province, china. *Remote Sensing,* 11 (2), 170.

Meyer, V., Scheuer, S. & Haase, D., 2009. A multicriteria approach for flood risk mapping exemplified at the mulde river, germany. *Natural hazards,* 48 (1), 17-39.

Midi, H., Sarkar, S.K. & Rana, S., 2010. Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics,* 13 (3), 253-267.

Mojaddadi, H., Pradhan, B., Nampak, H., Ahmad, N. & Ghazali, A.H.B., 2017. Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and gis. *Geomatics, Natural Hazards and Risk,* 8 (2), 1080-1102.

Nasa, 2020. Intense flooding in bangladesh. The National Aeronautics and Space Administration (NASA).

Ngo, P.-T.T., Pham, T.D., Hoang, N.-D., Tran, D.A., Amiri, M., Le, T.T., Hoa, P.V., Van Bui, P., Nhu, V.-H. & Bui, D.T., 2021. A new hybrid equilibrium optimized sysfor based geospatial data mining for tropical storm-induced flash flood susceptible mapping. *Journal of Environmental Management,* 280, 111858.

Nguyen, H.D., Nguyen, Q.-H., Du, Q.V.V., Nguyen, T.H.T., Nguyen, T.G. & Bui, Q.-T., 2021. A novel combination of deep neural network and manta ray foraging optimization for flood susceptibility mapping in quang ngai province, vietnam. *Geocarto International*, 1-25.

Panahi, M., Jaafari, A., Shirzadi, A., Shahabi, H., Rahmati, O., Omidvar, E., Lee, S. & Bui, D.T., 2021. Deep learning neural networks for spatially explicit prediction of flash flood probability. *Geoscience Frontiers,* 12 (3), 101076.

Pappenberger, F., Matgen, P., Beven, K.J., Henry, J.-B. & Pfister, L., 2006. Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in water resources,* 29 (10), 1430-1449.

Paul, G.C., Saha, S. & Hembram, T.K., 2019. Application of the gis-based probabilistic models for mapping the flood susceptibility in bansloi sub-basin of ganga-bhagirathi river and their comparison. *Remote Sensing in Earth Systems Sciences,* 2 (2), 120-146.

Persits, F.M., Wandrey, C.J., Milici, R.C. & Manwar, A., 2001. Digital geologic and geophysical data of bangladesh: U.S. Geological survey open-file report 97-470-h. U.S. Geological Survey.

Pham, B.T., Luu, C., Van Dao, D., Van Phong, T., Nguyen, H.D., Van Le, H., Von Meding, J. & Prakash, I., 2021a. Flood risk assessment using deep learning integrated with multi-criteria decision analysis. *Knowledge-Based Systems,* 219, 106899.

Pham, B.T., Luu, C., Van Phong, T., Nguyen, H.D., Van Le, H., Tran, T.Q., Ta, H.T. & Prakash, I., 2021b. Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in quang nam province, vietnam. *Journal of Hydrology,* 592, 125815.

Planchon, O. & Darboux, F., 2002. A fast, simple and versatile algorithm to fill the depressions of digital elevation models. *Catena,* 46 (2-3), 159-176.

Pradhan, B., 2010. Flood susceptible mapping and risk area delineation using logistic regression, gis and remote sensing. *Journal of Spatial Hydrology,* 9 (2).

Rahman, M., Chen, N., Elbeltagi, A., Islam, M.M., Alam, M., Pourghasemi, H.R., Tao, W., Zhang, J., Shufeng, T. & Faiz, H., 2021a. Application of stacking hybrid machine learning algorithms in delineating multi-type flooding in bangladesh. *Journal of Environmental Management,* 295, 113086.

Rahman, M., Ningsheng, C., Islam, M.M., Dewan, A., Iqbal, J., Washakh, R.M.A. & Shufeng, T., 2019. Flood susceptibility assessment in bangladesh using machine learning and multi-criteria decision analysis. *Earth Systems and Environment,* 3 (3), 585-601.

Rahman, M., Ningsheng, C., Mahmud, G.I., Islam, M.M., Pourghasemi, H.R., Ahmad, H., Habumugisha, J.M., Washakh, R.M.A., Alam, M. & Liu, E., 2021b. Flooding and its relationship with land cover change, population growth, and road density. *Geoscience Frontiers,* 12 (6), 101224.

Rahmati, O., Darabi, H., Panahi, M., Kalantari, Z., Naghibi, S.A., Ferreira, C.S.S., Kornejady, A., Karimidastenaei, Z., Mohammadi, F. & Stefanidis, S., 2020. Development of novel hybridized models for urban flood susceptibility mapping. *Scientific Reports,* 10 (1), 1-19.

Rincón, D., Khan, U.T. & Armenakis, C., 2018. Flood risk mapping using gis and multi-criteria analysis: A greater toronto area case study. *Geosciences,* 8 (8), 275 Available from: https://www.mdpi.com/2076-3263/8/8/275.

Ronco, P., Bullo, M., Torresan, S., Critto, A., Olschewski, R., Zappa, M. & Marcomini, A., 2015. Kulturisk regional risk assessment methodology for water-related natural hazards–part 2: Application to the zurich case study. *Hydrology and Earth System Sciences,* 19 (3), 1561-1576.

Saaty, T.L. & Tran, L.T., 2007. On the invalidity of fuzzifying numerical judgments in the analytic hierarchy process. *Mathematical and Computer Modelling,* 46 (7-8), 962-975.

Sarkar, D. & Mondal, P., 2020. Flood vulnerability mapping using frequency ratio (fr) model: A case study on kulik river basin, indo-bangladesh barind region. *Applied Water Science,* 10 (1), 1-13.

Siam, Z.S., Hasan, R.T., Anik, S.S., Noor, F., Adnan, M.S.G. & Rahman, R.M., Year. Study of hybridized support vector regression based flood susceptibility mapping for bangladeshed.^eds. *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*Springer, 59-71.

Siam, Z.S., Hasan, R.T. & Rahman, R.M., 2021b. Effects of label noise on regression performances and model complexities for hybridized machine learning based spatial flood susceptibility modelling. *Cybernetics and Systems*, 1-18.

Singh, V.K., Kumar, D., Kashyap, P., Singh, P.K., Kumar, A. & Singh, S.K., 2020. Modelling of soil permeability using different data driven algorithms based on physical properties of soil. *Journal of Hydrology,* 580, 124223.

Steele, J.E., Sundsøy, P.R., Pezzulo, C., Alegana, V.A., Bird, T.J., Blumenstock, J., Bjelland, J., Engø-Monsen, K., De Montjoye, Y.-A. & Iqbal, A.M., 2017. Mapping poverty using mobile phone and satellite data. *Journal of The Royal Society Interface,* 14 (127), 20160690.

Stefanidis, S. & Stathis, D., 2013. Assessment of flood hazard based on natural and anthropogenic factors using analytic hierarchy process (ahp). *Natural hazards,* 68 (2), 569-585.

Talukdar, S., Ghose, B., Salam, R., Mahato, S., Pham, Q.B., Linh, N.T.T., Costache, R. & Avand, M., 2020. Flood susceptibility modeling in teesta river basin, bangladesh using novel ensembles of bagging algorithms. *Stochastic Environmental Research and Risk Assessment,* 34 (12), 2277-2300.

Tehrany, M.S., Pradhan, B., Mansor, S. & Ahmad, N., 2015. Flood susceptibility assessment using gis-based support vector machine model with different kernel types. *Catena,* 125, 91-101.

Tehrany, M.S., Shabani, F., Neamah Jebur, M., Hong, H., Chen, W. & Xie, X., 2017. Gis-based spatial prediction of flood prone areas using standalone frequency ratio, logistic regression, weight of evidence and their ensemble techniques. *Geomatics, Natural Hazards and Risk,* 8 (2), 1538-1561.

Thirumurugan, P. & Krishnaveni, M., 2019. Flood hazard mapping using geospatial techniques and satellite images—a case study of coastal district of tamil nadu. *Environmental monitoring and assessment,* 191 (3), 1-17.

Unitar, 2020. Satellite detected water extent in bangladesh. Geneva, Switzerland: United Nations Institute for Training and Research (UNITAR)

Vilasan, R. & Kapse, V.S., 2021. Evaluation of the prediction capability of ahp and f_ahp methods in flood susceptibility mapping of ernakulam district (india).

Wang, Y., Fang, Z., Hong, H. & Peng, L., 2020. Flood susceptibility mapping using convolutional neural network frameworks. *Journal of Hydrology,* 582, 124482.

Wang, Y., Hong, H., Chen, W., Li, S., Panahi, M., Khosravi, K., Shirzadi, A., Shahabi, H., Panahi, S. & Costache, R., 2019. Flood susceptibility mapping in dingnan county (china) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm. *Journal of environmental management,* 247, 712-729.

Wang, Y., Li, Z., Tang, Z. & Zeng, G., 2011. A gis-based spatial multi-criteria approach for flood risk assessment in the dongting lake region, hunan, central china. *Water resources management,* 25 (13), 3465-3484.

Warpo, 2018. National water resources database (nwrd). . Dhaka, Bangladesh.: Water Resources Planning Organization (WARPO).

Worldpop, 2020. The spatial distribution of population in 2020, bangladesh University of Southampton.

Youden, W.J., 1950. Index for rating diagnostic tests. *Cancer,* 3 (1), 32-35.

Zadeh, L.A., 1996. Fuzzy sets. *Fuzzy sets, fuzzy logic, and fuzzy systems: Selected papers by lotfi a zadeh.* World Scientific, 394-432.

Zou, Q., Zhou, J., Zhou, C., Song, L. & Guo, J., 2013. Comprehensive flood risk assessment based on set pair analysis-variable fuzzy sets model and fuzzy ahp. *Stochastic Environmental Research and Risk Assessment,* 27 (2), 525-546.