Research Article

Azemeraw Wubalem*, Gashaw Tesfaw, Zerihun Dawit, Belete Getahun, Tamrat Mekuria, and Muralitharan Jothimani

Comparison of statistical and analytical hierarchy process methods on flood susceptibility mapping: In a case study of the Lake Tana sub-basin in northwestern Ethiopia

https://doi.org/10.1515/geo-2020-0329
received January 22, 2021; accepted December 05, 2021

Abstract: The flood is one of the frequently occurring natural hazards within the sub-basin of Lake Tana. The flood hazard within the sub-basin of Lake Tana causes damage to cropland, properties, and a fatality every season. Therefore, flood susceptibility modeling in this area is significant for hazard reduction and management purposes. Thus, the analytical hierarchy process (AHP), bivariate (information value [IV] and frequency ratio [FR]), and multivariate (logistic regression [LR]) statistical methods were applied. Using an intensive field survey, historical document, and Google Earth Imagery, 1,404-flood locations were determined, classified into 70% training datasets and 30% testing flood datasets using a subset within the geographic information system (GIS) environment. The statistical relationship between the probability of flood occurrence and 11 flood-driving factors was performed using the GIS tool. The flood susceptibility maps of the study area were developed by summing all weighted aspects using a raster calculator. It is classified into very low, low, moderate, high, and very high susceptibility classes using the natural breaks method. The accuracy and performance of the models were evaluated using the area under the curve (AUC). As the result indicated, the FR model has better performance (AUC = 99.1%) compared to the AHP model (AUC = 86.9%), LR model (AUC = 81.4%), and IV model (AUC = 78.2%). This research finds out that the applied methods are quite worthy for flood susceptibility modeling within the study area. In flood susceptibility modeling, method selection is not a serious challenge; the care should tend to the input parameter quality. Based on the AUC values, the FR model is comparatively better, followed by the AHP model for regional land use planning, flood hazard mitigation, and prevention purposes.

Keywords: Gumara, Ribb, inventory, geographic information system

1 Introduction

A flood is an overflow of water that usually submerges dry land. It also can occur in rivers or lakes when the flow exceeds the rivers channel's capacity, particularly at the bends or meanders within the waterway and backflow from the Lakes [1]. Although flood is one of the natural parts of the hydrological cycle, it has been increasing in both frequency and magnitude from year to year [1]. This is often due to the over-change of climate and land degradation in the world due to anthropogenic intervention. This can reduce the water-retention capacity of the catchments and raise the rate of soil erosions due to the cleanup of forests for a special purpose. Overflows from rivers and impermeable land surfaces can cause flooding in the surrounding area or flood plains [2,3]. According to recent studies, dangerous floods occur over the world because of climate and land use land cover change [4,5]. According to Charlton et al. [6], changes in land use or cover pattern have a major impact on the frequency and amplitude of floods in a given region when the land becomes impermeable, increasing flow velocity. Other parameters such as slope angle, curvature, flow accumulation, normalized vegetation index, stream density,

* **Corresponding author: Azemeraw Wubalem,** Department of Geology, College of Natural and Computational Sciences, University of Gondar, P.O. BOX 196, Maraki Street, Ethiopia,
e-mail: alubelw@gmail.com
Gashaw Tesfaw, Zerihun Dawit, Belete Getahun, Tamrat Mekuria, Muralitharan Jothimani: Department of Geology, College of Natural and Computational Sciences, University of Gondar, P.O. BOX 196, Maraki Street, Ethiopia

distance to a river, groundwater level, soil texture, and elevation can also play a role in flood occurrence [7,8].

The flood hazard has been causing damage to crops, infrastructures, engineering structures, properties, and loss of human and animal lives worldwide, including Ethiopia. As reported by earlier studies [1,9,10], the flood has resulted in a risk to a human being, properties, communication systems, cultural heritage, and ecosystems. Flood hazards in Ethiopia resulted in huge damages to properties, crops, farmlands, infrastructures, and loss of life. For instance, within the last 2 years, 2019–2020, flood hazards caused to displace more than 500,000 people and damaged wide cultivated lands (more than 25,000 ha cultivated lands), damaged various engineering structures, destructed more than 35 houses, and loss of lives in Amhara, Somali, Afar, Southern Nations, Nationality and People, Dire Dawa, and Oromia regions of Ethiopia. These flood damages can reduce by identifying susceptibility areas and undertaking prevention measures [7,11]. Floods are the most common and destructive natural hazards worldwide, prompting academics to examine and understand flood hazards to develop efficient flood prevention and mitigation strategies [2,12–15]. Identifying flood-prone areas aids in well-informed flood management and helps to develop adequate flood management measures to reduce risks and losses [15–21].

Due to the complexity of flood occurrences, interactions of many geo-environmental and anthropogenic factors, and a lack of precise data, accurate prediction of areas susceptible to flooding and generation of credible susceptibility maps is multifarious and time consuming [17,18,22–24]. Advancement of geographical information systems (GIS) [25] and computational techniques has resulted in the development of various spatially explicit modeling approaches that rely on expert experience, physically based hydraulic models, and data-driven methods; these could be used for flood susceptibility modeling [17,18,22–24,26]. These methods are grouped into qualitative (geomorphology analysis and topographic wetness index (TWI)), semiquantitative (analytical hierarchy process (AHP)), quantitative (machine learning and statistical), and hydrological-based methods.

The hydrological methods are very simple and perform based on a nonlinear concept. They are less effective in modeling complex features like catchments [21]. Nowadays, these traditional methods are replaced by automated and rule-based methods that are more suitable for flood hazard mapping [21]. Soil & water assessment tool and WetSpass methods are examples of hydrological methods used to produce spatial flood susceptibility models by integrated GIS and remote sensing tools. Qualitative methods are an expert-driven approach, which requires field experience specialists [27]. Relying on the experience and professional background of experts and subjectivity is that the drawback of those methods.

An AHP is an example of a semiquantitative method used to produce a flood susceptibility model supported by a multicriteria analysis framework [28]. In 1991, the approach was initially utilized in an ArcGIS context [29]. Its inability to determine the inherent ambiguities in spatial decision-making procedures frequent when selecting, comparing, and, perhaps most significantly, rating various criteria are one of its shortcomings [30]. However, for regional research, AHP is still a viable option [31,32]. Flood mapping has made extensive use of AHP, as shown by earlier studies [11,33–38].

Machine learning techniques are advanced methods utilized in flood susceptibility mapping. However, a substantial time interval, the need of getting high-performance computing systems and specific software, and strict selection criteria for input parameters make machine-learning methods less usable for a wide range of users [7]. TWI may be a commonly used very simple traditional flood susceptibility mapping method. However, this method used considering only the slope and flow accumulation in a region. Nevertheless, flood hazard is the result of the combination of several factors.

Statistical methods are indirect methods widely or routinely used to evaluate the correlation between flood-driving factors and floods based on mathematical expression [38–42]. Statistical methods are imperative to utilize a quick, understandable, and accurate flood susceptibility model. It has no specific requirements regarding input data, software, and computer capacity. The statistical methods are often further divided into multivariate and bivariate statistical methods. They supply reliable results [43–53]. The bivariate statistical methods were used to evaluate the relationship between flood governing factors and past flooding. Frequency ratio (FR), certainty factor, IV, and weight of evidence are examples of bivariate statistical methods. They are simple and produce reliable models. They also help to measure the consequences of a flood at an element class level that is impossible in data mining or multivariate methods. However, it requires quality flood factor data, past flood data, and lacking to measure of the relationship among flood governing factors. Multivariate statistical methods were used to examine the relationship between two and above dependent and set of independent variables [45,46,50,54,55]. Logistic regression (LR) and discriminant analysis are examples of multivariate statistical methods that are frequently used in the flood susceptibility modeling and provide reliable results

[39,45,46,48,50,55]. However, it is incapable of examining each factor class's contribution for flood probability such as data mining, unlike bivariate methods.

Although the model's suitability depends on physical parameters, data quality, and availability, expert and technological advancement, comparison among different natural hazard mapping methods is one solution to choose appropriate approaches. Hence, each method has its limitation; using different techniques for landslide or flood susceptibility mapping is extremely important to fill the gap among the methods. For instance, the LR model can perform multivariate statistical analysis between a variable and a group of independent variables. However, it is incapable of analyzing the impacts of internal classes of flood governing factors individually on flood occurrence. This limitation can be solved using FR and IV methods. For instance, FR and IV methods have often extracted the influence of each flood governing factor class on flood occurrence, but it cannot consider the relationship between these flood controlling factors and flood occurrence. AHP method is used to evaluate the effects of flood factor class and individual factors, but it depends on expert judgment partially. Therefore, the combination use of AHP, bivariate (IV and FR), and multivariate (LR) statistical methods is used in this study to fill the limitation of every applied method. Based on the concerns stated overhead, the main objective of this study is (1) to generate a flood susceptibility map of the study area; (2) to compare and evaluate the performance of the FR, IV, LR, and AHP methods to work out flood-prone areas; and (3) to measure the relationship between flood factors and flood probability as well as flood factor class and flood occurrence. The nobility of this study lies in the following: (1) for the first time, the rigorous flood susceptibility methods such as statistical methods were conducted within the sub-basin of Lake Tana to get flood susceptibility model and (2) the comparison among the IV, FR, LR, and AHP methods has not performed yet. The resulting map will help the regional and local authorities and policymakers to mitigate flood hazards.

2 Study area

The study area is located in the sub-basin of Lake Tana, northwestern Ethiopia, characterized by wide flat to gently sloping plains and somehow raged topography. Its elevation ranges from 1,774 to 4,037 m above mean sea level (Figure 1). It is bound between 330,000–410,000E and 1,280,000–1,350,000N. Annual rainfall varies between 600 and 2,600 mm (Figure 2a). The seasonal rainfall follows a unimodal distribution, with July being the heaviest month. The average annual rainfall is 1394.73 mm, with mean monthly values ranging from 0.6 mm in January to 415.8 mm in July (Figure 2b). The area's mean monthly temperature is around 19°C, with a monthly maximum temperature of 27.3°C and a monthly minimum

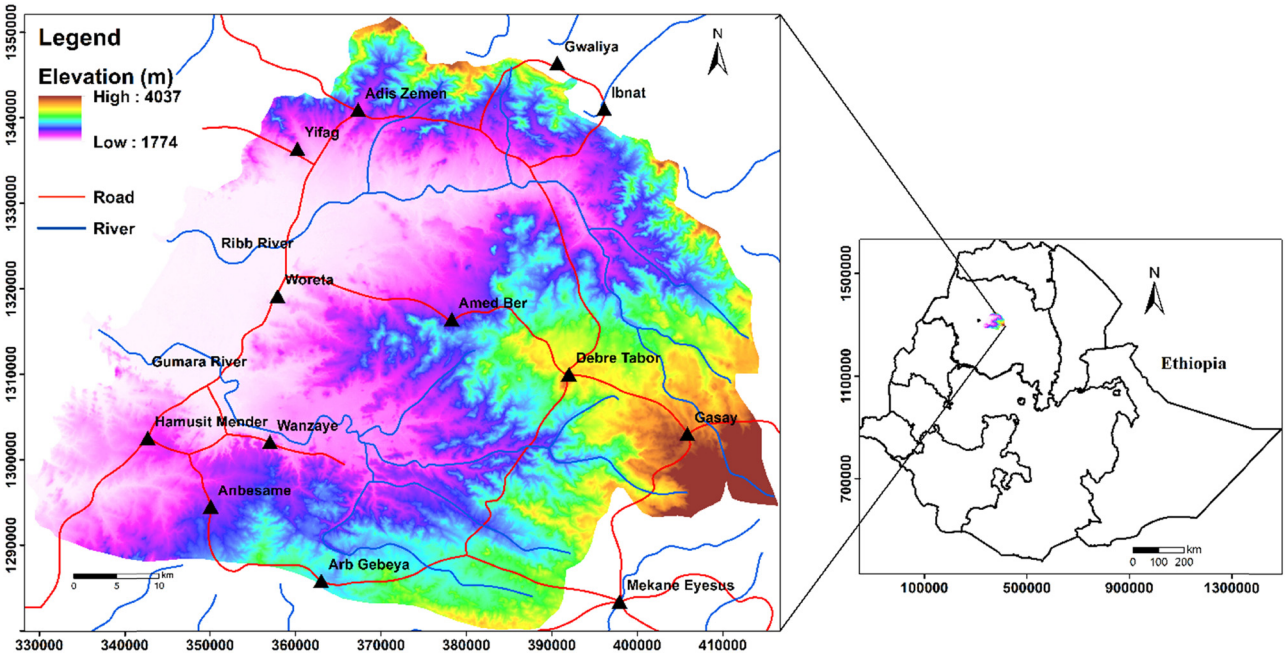

Figure 1: Location map of the study area (projection: WGS 1984 UTM Zone 37N).

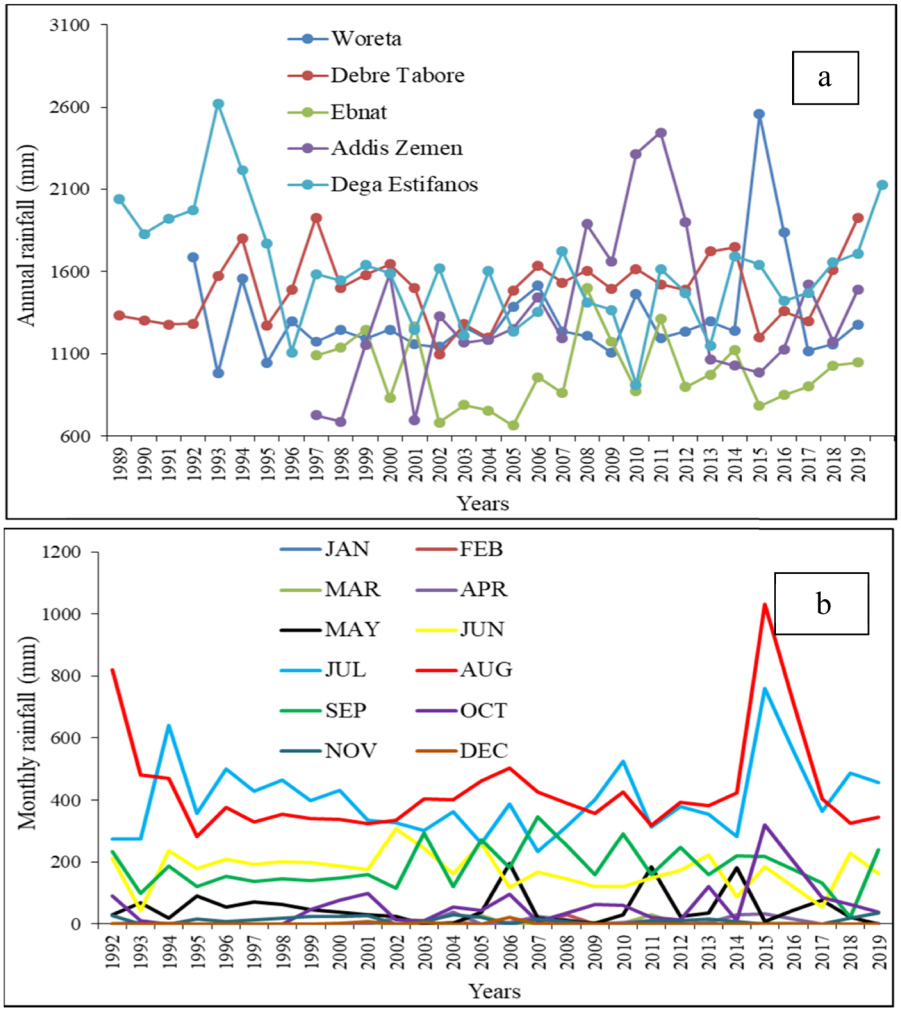

Figure 2: (a) Dark blue, light blue, red, green, and dark pink color dash lines indicated annual rainfall distribution trends from 1989 to 2019 in Woreta, Dega Estifanos, Debre Tabor, Ebnat, and Addis Zemen Rainguage stations, respectively. (b) Monthly rainfall of the study area: red color showed high rainfall in August. Light blue color dash line indicated high rainfall in the month of June next to August, but the green and yellow color dash line indicated moderate rainfall trends in September and March, respectively. But the dark pink, black, dark blue, brown, and light red color indicated very low rainfall trend in the month of October, May, January, December, February, etc., respectively.

temperature of 11.5°C. The flood hazards have been frequently affecting the study area during heavy and prolonged rainfall seasons (June–September) since 1996 Ethiopian calender. The 97% of the population are farmers who settled densely in the low-lying lands of Ribb and Gumara river outlet in which the areas are suitable for farming due to high soil fertility and very flatlands. This makes the area vulnerable to flood incidences. The flooding in the study area, leading to the loss of life, properties, destruction of homes, roads, and more than 7,000 ha farmlands covered by various crops within the area. These show that the massive economic loss caused by flooding hazards that retard the sustainable development of the economy of the resident and the country. The study area has many tributaries that drained to the two major rivers called Gumara and Ribb that also drained to Lake Tana, which are the parts of the Abay basin.

3 Materials and methods

3.1 Materials

3.1.1 Flood inventory map

A flood inventory map is one of the critical elements, which can be prepared using various techniques such as an aerial photograph or Google Earth Imagery, field investigation, and evaluation of archived data coupled

with GIS tool. Evaluating and recognizing the correlation between flood-driving factors and flood incidence requires quality data [55–58]. In this study, 1,404 flood inventory data were collected after flooding on August 15–25, 2019, and from November 25 to December 15, 2019. The flood inventory data of the study area were collected from historical records, time-series Google Earth Imagery, and extensive fieldwork. In the literature, several suggestions were provided regarding the size of flood samples to be used for modeling and model verification [59]. Some scholars classified the flood data into 80%/20%, but 70%/30% classification is commonly used in flood susceptibility mapping [27,49,50]. Therefore, based on a literature review [37,60,61], the flood inventory data were classified into 70% (983) flood for the training dataset and 30% (421) for testing datasets keeping their spatial distribution using subset in ArcGIS 10.1 [58,60,61] as shown in Figure 3. The training flood datasets were used for model preparation, whereas the testing flood datasets were used for the evaluation of model prediction accuracy. The same numbers of flood and nonflood points were chosen for the LR analysis.

3.1.2 Flood-driving factors

The selection of flood factors depends on the study area's physical and natural characteristics and data availability [62,63]. The factors in the study area are selected based on the study area's geo-environmental condition, data availability, and a literature review [58,59]. The slope angle, slope curvature, land use, soil texture, distance to stream/river, stream density, normalized vegetation index, flow accumulation, groundwater depth, rainfall, and elevation are considered to examine the spatial relationship between

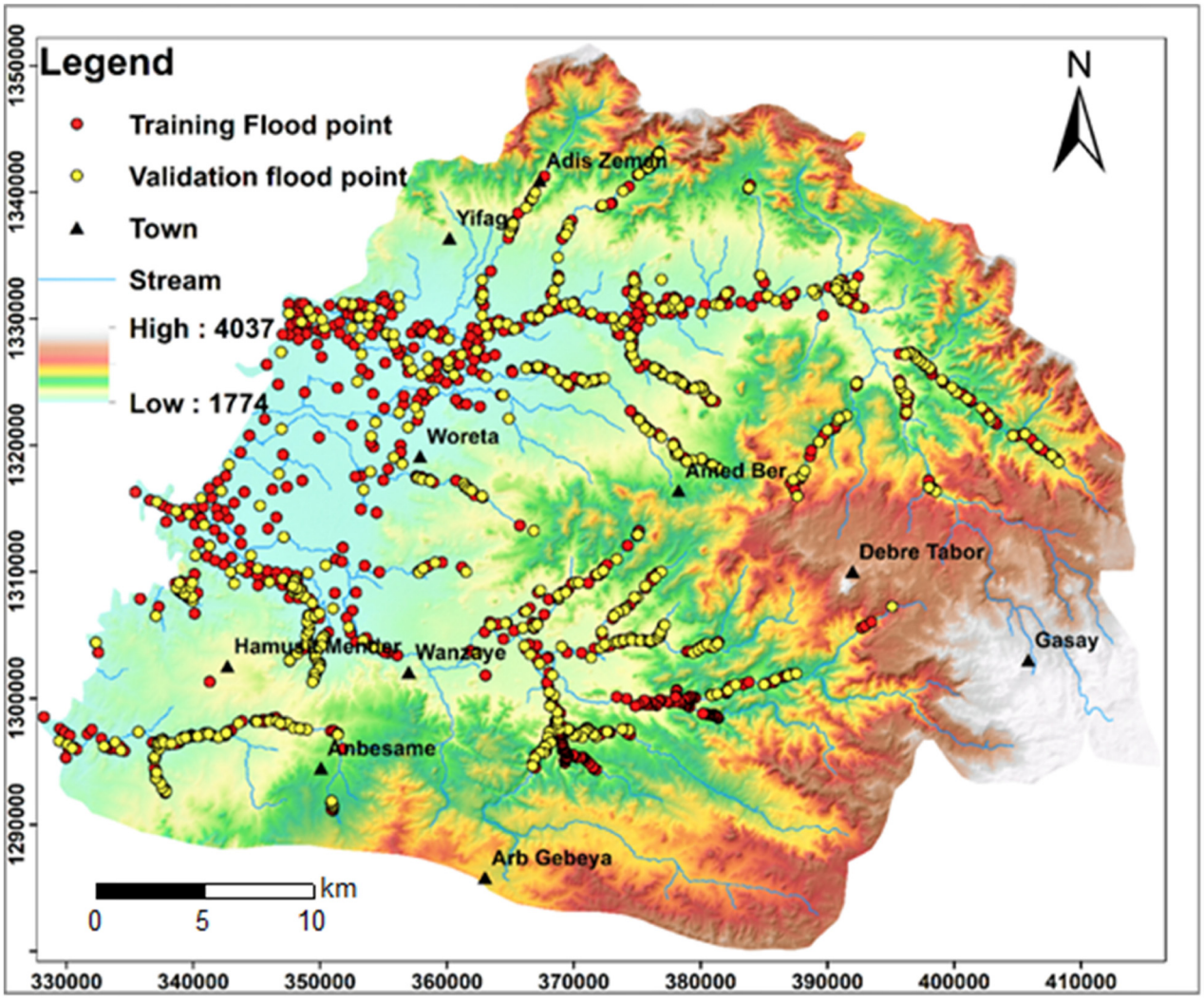

Figure 3: Flood location map of the study area (projection: WGS 1984 UTM Zone 37N): red points are training flood datasets and yellow points are validation flood datasets.

them and flood occurrence. The flood factors, derived from digital elevation model (DEM 12.5 m × 12.5 m resolution), are distance to stream, slope angle, flow accumulation, stream density, elevation, and slope curvature maps (Figure 4). The study area's soil map was prepared through digitization from a 1:250,000 textural soil map of the Amhara Region, which has four classes (silty sand, sandy loam, clay, and loam). Land use/land cover (LULC) and normalized difference vegetation index (NDVI) maps of the study area were prepared from Sentinel 2 satellite image analysis using ArcGIS with a high-resolution Google Earth image interpretation. The LULC has eight classes including grazing land, agricultural land/cropland, barren land, residential/settlement, river zone/water body, forest, and wetland (Figure 4), whereas NDVI has five classes (Figure 4). The rainfall and groundwater depth maps were constructed using ArcGIS 10.1 from annual mean rainfall and well data.

Elevation and slope are two important factors that influence flood incidence. Floods occur in low-lying and flat areas, whereas flooding at the summits of mountains is not conceivable [58]. Steep slopes accelerate surface runoff and reduce the amount of time the soil has to absorb the water. In the case of the aspect factor, this component influences the amount of rainfall and sunlight received by the terrain [58]. Curvature is divided into three categories: flat, convex, and concave and is an important factor in flood studies [58]. The LULC factor has a significant impact on terrain infiltration, runoff speed, and extent [37,58]. Rainfall is a major factor in the incidence of floods [53]. Flooding occurs when there is a lot of rain in a short period, usually less than 6 h [37,62]. When flow accumulation increases, flood susceptibility increases [37]. Distance from rivers is another significant determining factor. The Euclidean distance tool in ArcGIS software was used to calculate distance from rivers in this study. When it comes to river network density, more density often indicates more surface runoff and, as a result, a higher risk of flooding. Another aspect to consider is the NDVI, which is a useful

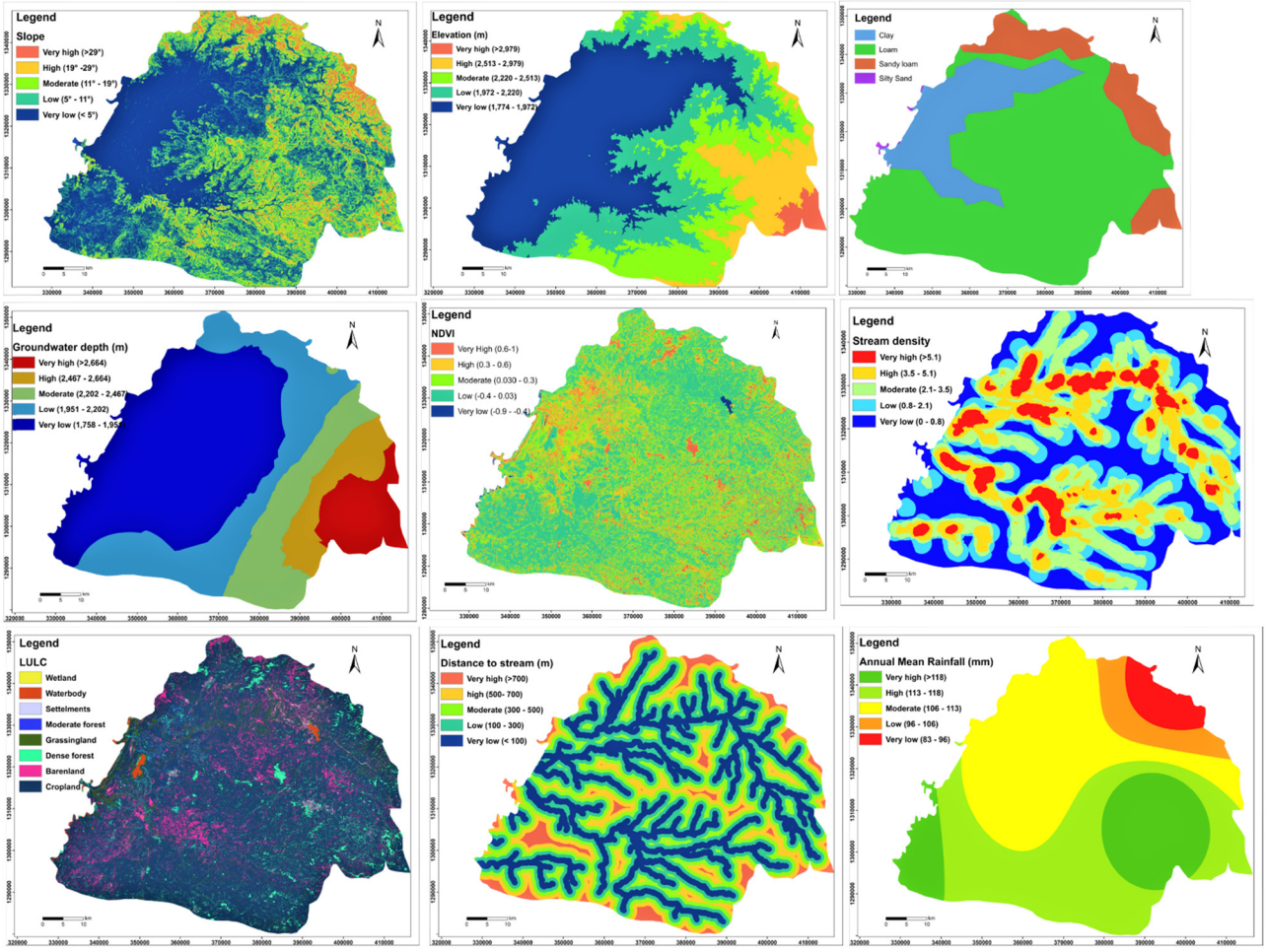

Figure 4: Flood governing factor maps of the study area (slope, elevation, soil texture, groundwater depth, NDVI, stream density, LULC, and distance to stream and rainfall) produced using ArcGIS (projection: WGS 1984 UTM Zone 37N).

index for determining vegetation covering and its impact on flooding in a basin, which varies from −1 to +1 [63].

4 Methods

4.1 Research procedure

To achieve the goal of this study, various activities and steps were used. These are data collection, flood inventory mapping, database creation, flood susceptibility modeling, and model validation. Moreover, appropriate data, including a topographic map, borehole data, DEM with 12.5 m resolution, historical flood events, soil map, and meteorological data were collected. These data were collected from the United States Geological Survey, Amhara Water Well Drilling Enterprise (AWWDE), field survey, Google Earth Imagery, Ethiopian National Meteorological Agency, and the Geological Survey of Ethiopia. Geodatabase building is one of the most fundamental elements in flood susceptibility mapping. Therefore, four databases were built for IV, LR, FR, and AHP models. The FR, IV, and AHP databases contain flood inventory and flood-driving factors. In contrast, the LR database contains flood and no flood points with 11-weighted flood-driving factors. After the databases were built, 11 flood factor maps were reclassified into subclass and overlaid with reclassified training flood datasets. Weight ratings for all flood factor classes were assigned statistically using Excel and ArcGIS. These weighted maps are rasterized using lookup in spatial analyst. After rasterizing the factor maps, the flood susceptibility index (FSI) maps were generated by sum-up of all raster maps using a raster calculator in Map Algebra. These maps (FSIs) are classified into a fivefold classification scheme: very low, low, moderate, high, and very high susceptibility classes using the natural breaks method (Figure 6). In the LR method, the study area is classified as flood and nonflood points using GIS. Then, the weight of 11 factors was extracted to generate LR coefficients of each flood factor in Statistical Package for the Social Sciences (SPSS). Finally, the FSI of the area was generated using the logistic flood probability equation (equation (8)) and GIS tools. Finally, the four models' accuracy was evaluated using the prediction rate curve based on observed testing flood datasets (Figure 7). In general, the general flow chart for the study approach is summarized in Figure 5.

4.2 Modeling approaches

4.2.1 IV model

The IV method is one of the probabilistic methods of a bivariate statistical approach, which is used to envisage the correlation between floods and flood factor classes

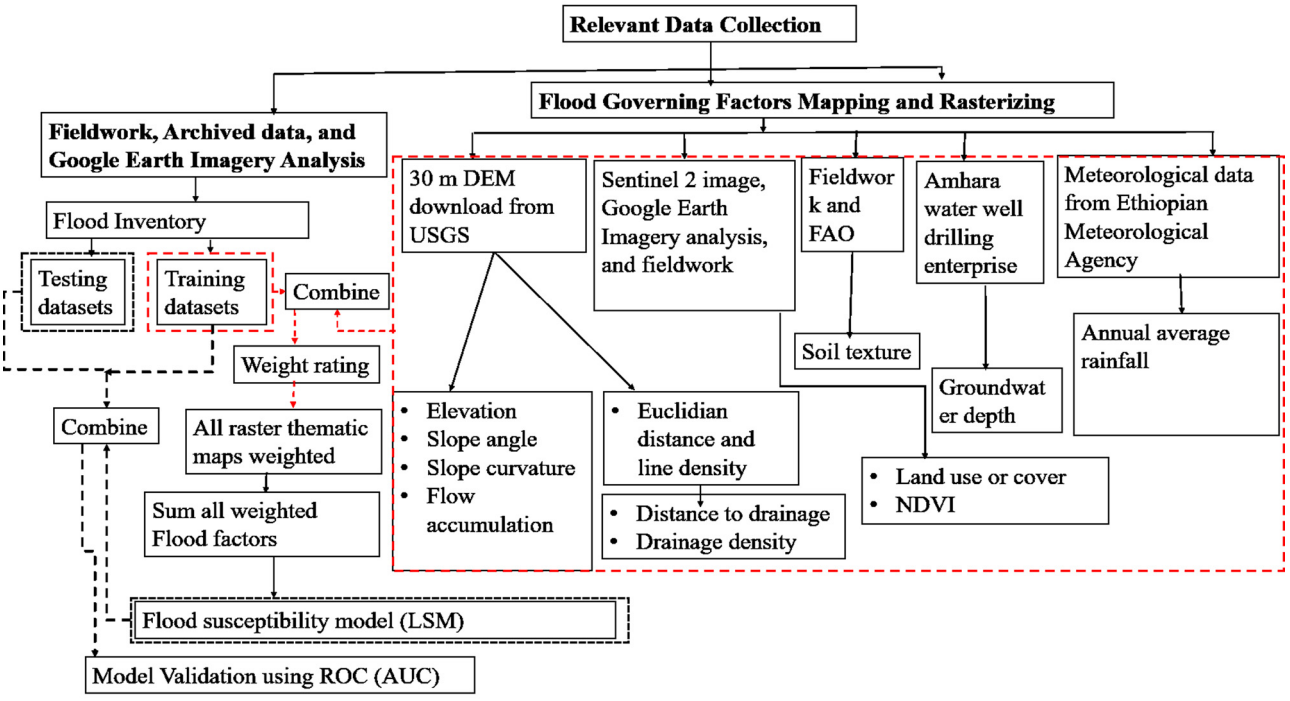

Figure 5: General methodological flow chart of the study.

[64]. It was used by earlier studies [11,64], and it provided reliable results. The IV method requires preexisting flood data to analyze the statistical relationship between flood occurrences and triggering factors, which is the limitation of the IV methods. IV is not effective to determine the effect of factors for flood occurrence rather than factor classes. This is another limitation of the IV method. As shown in equation (1), the IVs are often calculated within the natural logarithm of the ratio of preexisting flood in each flood-driving factor class to the entire flood density within the study area [11,63,64] (equation (1)).

$$\mathrm{IV} = \ln\left(\frac{\text{Conditional probability (CP)}}{\text{Prior probability (PP)}}\right) = \frac{\frac{\text{Nfopix}}{\text{Ncpix}}}{\frac{\text{Ntfopix}}{\text{Ntcpix}}}. \quad (1)$$

Conditional probability is the ratio of the pixel of a flood in class to the pixel of a class [11,63,64]. Prior probability is the ratio of the total number of pixels of the flood to the total number of pixels of the study area [11,63,64]. Nfopix is a flood pixel/area in a flood factor class. Ntfopix is the total area of a flood in the entire study area. Ncpix is the class area in the study area, and Ntcpix is the total pixel area in the whole study area. When the IV > 0.1, the flood occurrence within the factor classes has a high correlation, which means it will have a high probability of flood occurrence. However, when the IV < 0.1 or IV < 0, it is a low correlation between flood factors and flood occurrence. The FSI of the study area was calculated using equation (2).

$$\mathrm{FSI} = \sum_{i=1}^{n} \mathrm{IV}_i X_i, \quad (2)$$

where FSI is the flood susceptibility index and IV_i is the information value of each factor class. The higher value of FSI has indicated a higher probability of flood occurrence.

4.2.2 LR model

LR is one of the favored multivariate statistical methods, which may be used to establish a multivariate regression relationship between floods and sets of independent variables [65]. LR is a commonly used method in natural hazard studies [66]. Among other statistical methods, the LR model has proven one among the foremost reliable approaches for flood susceptibility mapping to work out the foremost flood influencing factors [39,45,67–69]. It was used by earlier studies [11,66] and provided reliable results. The LR method requires preexisting flood data to analyze the statistical relationship between flood occurrences and triggering factors, which is the limitation of the LR method. LR is not effective to determine the effect of factor classes for flood occurrence. This is another limitation of the LR method. The problem of using the LR model lies within the sample size selection of dependent and independent variables. There are three ways of sampling flood and nonflood points [66]. The primary way is using all data from all the study areas. However, this results in an uneven proportion of nonflood and flood pixels, incorporating an outsized volume of data within the analysis. Using all flood pixels with equal nonflood pixels is that the second method, leading to a less reliable output, but it can reduce sample size and sampling bias. The third method uses an unequal or equal proportion of flood and nonflood pixels by classifying floods into training and testing datasets. In this study, the dependent data are binary variables and made from 0 and 1, representing the absence and presence of floods, respectively. Consequently, an equal number of nonflood sample points, whose variable value is 0, were randomly selected from flood-free areas to represent the absence of floods using GIS. The equal number of flood points and nonflood points was merged. Moreover, all the values of independent variables contained flood and nonflood points were extracted from the maps of every flood-governing factor. Before LR analysis applies, collinearity (interrelatedness of independent variable) of independent variables was checked using tolerance (TOL) and variance inflation factor (VIF) index in the SPSS software. When the TOL < 0.2 and VIF > 5, the given independent variables have multicollinearity. Based on the test, the variables did not show a multicollinearity problem (Table 2). Then, the LR analysis was conducted, and coefficients were calculated within the SPSS program. It is expressed mathematically [68,70] as follows:

$$P = \frac{1}{1 + \mathrm{e}^{-z}}, \quad (3)$$

where P is the probability of flood occurrence varies from zero to one. z is the linear combination of the predictors and varies from $-1 < z < 0$ for higher odds of nonflood occurrence to $0 < z < 1$ for higher flood occurrence odds. z can be defined as follows:

$$z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \ldots \beta_n X_n, \quad (4)$$

where x_1, x_2, x_3 ... x_n are independent variables, β_0 is the intercept of the slope of LR analysis, and β_1, β_2, β_3 ... β_n are the coefficients of the LR analysis.

4.2.3 FR model

FR is a commonly used method in natural hazard mapping due to its simple and rapid calculation procedures, also as it provides reliable results [58]. FR was used by earlier studies [11,62–64], and it provided reliable results. The FR method requires preexisting flood data to analyze the statistical relationship between flood occurrences and triggering factors, which is the limitation of the FR method. FR is not effective to determine the effect of factors for flood occurrence rather than factor classes. This is another limitation of the FR method. The FR is that the ratio of areas where the flood occurred in areas to areas where the flood did not occur. When the ratio value is more significant than one, it indicates a robust correlation between factor classes and flood occurrence in a given terrain. However, the ratio value less than one showed a weak correlation between flood occurrence and flood factors [71]. It can calculate using equation (5).

$$\mathrm{FR} = \frac{a}{b} = \frac{\frac{\mathrm{Nfopix}}{\mathrm{Ntfopix}}}{\frac{\mathrm{Ncpix}}{\mathrm{Ntcpix}}}, \tag{5}$$

where FR is frequency ratio, Nfopix is a flood pixel/area in a flood factor class, Ntfopix is that the total area of a flood in the whole study area (a), Ncpix is an area of the category within the study area, and Ntcpix is that the total pixel area within the whole study area (b). In the present research work, the FR for every causative factor class was calculated using equation (5), and the results are summarized in Table 1. To generate the FSI, each factor type or class's FR is summed as in equation (6).

$$\mathrm{FSI} = \sum_{i=1}^{n} \mathrm{FR}_i X_i. \tag{6}$$

FSI is the flood susceptibility index, n is the number of flood factors, X_i is the flood factor, and FR_i is the frequency ratio of each flood factor type or class. FR is the frequency ratio value of each factor class.

4.2.4 AHP

The AHP method is a structured tool used to analyze difficult decisions based on mathematics and psychology [72–75]. The AHP method has three basic steps: the problem must first be broken down and organized into a hierarchy of subproblems, second, data must be collected and quantified using pairwise attribute comparisons, and third, priority weights of factors or items in each level must be calculated [72–75]. The AHP method is very important to predict flood susceptibility models based on expert experience. Subjectivity during parameter weight rating is the limitation of the AHP. AHP is a very helpful method to evaluate the effects of both individual factors and classes. This makes the AHP unique from statistical methods. In this study, an excel pairwise comparison matrix was used to provide weights for each parameter or criteria that were considered in Saaty's ranking scale [76,77]. Scale is critically crucial before performing a pairwise comparison matrix (Table 1). Saaty devised a scale that goes from 1 to 9 for this purpose. One implies that the parameters are of equal relevance, whereas nine indicates that the factor or parameters are of tremendous importance. Consistency of weight for flood factor class examined by the consistency ratio (CR), calculated using equation (7) [76,78]. When the CR is less than 0.1, the weight of the factor class calculated using the comparison matrix is consistent. If it is greater than 0.1, the comparison matrix is inconsistent and will be revised. In this study, the CR is less than one. After each factor class's weight is determined, the flood susceptibility map is produced as shown in equation (9) [16].

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}}, \tag{7}$$

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n}, \tag{8}$$

$$\mathrm{FSI} = \sum_{i=1}^{n} W_i \star X_n, \tag{9}$$

Table 1: Fundamentals of Saaty scale [71]

Scales	Preferential level	Explanation
1	Equally important	The parent decision element is influenced equally by two-decision components
3	Moderately more important	One decision factor has a moderately higher influence than the other does
5	Strongly important	One decision factor has a greater impact than the other does
7	Very strongly important	One decision factor has a lot greater influence than the other does
9	Extremely important	The contrast in influences between the two choice elements is enormous
2, 4, 6, 8	Intermediate judgment values	Weights 1, 3, 5, 7, and 9 are used to illustrate compromises between preferences

where CR is consistency ratio, CI is consistency index, RI is the average random consistency index of the judgment matrix, and λ_{max} is the largest Eigenvalue derived from the pairwise comparison matrix, and n is the number of flood factor, W_i is the weight of the flood factor, X_n is the flood factors, and FSI is flood susceptibility index. CI is calculated using equation (8).

4.2.5 Model validation

Flood susceptibility modeling without prediction and model performance evaluation is nonsense to apply for disaster reduction programs. Although researchers used many techniques to validate the flood susceptibility model, the receiver operating characteristics (ROC) method is routinely used [37,79,80] because it is simple and produces reliable results [1,27,53,56]. The prediction and model performance of the flood susceptibility map of the study area were validated by comparing the flood model with existing flood data [78,81]. The prediction accuracy and model performance of the flood susceptibility maps were evaluated quantitatively using the ROC curve based on true-positive and false-positive rates [58,82]. The predictive rate curves for the four models were obtained by comparing testing flood datasets with the FSI, while the success rate curves were also obtained by comparing training flood datasets. The AUC value ranges from 0.5 to 1 [83], and the closer the value to one indicates higher the accuracy of the model.

5 Result and discussion

5.1 Result

5.1.1 FR

The flood occurrence ratio was divided by the area ratio to calculate the FR for each class of flood-conditioning factor. Table 2 shows the derived FR result for each factor class. A higher FR weight indicates a stronger link between that class and flood occurrence. For example, the data revealed that the first classes of slope and elevation had the greatest FR values. This supports the theory that flooding occurs mostly in low-lying or gentle slope areas. As indicated in Table 2, as the slope angle increases, the degree of flood occurrence decreases. The details for all parameters' weight ratings are summarized in Table 2.

5.1.2 IV

Table 2 shows the IV weights for each flood-conditioning component derived using Microsoft Excel and ArcGIS. The higher the IV weight for each class of every conditioning element, the more likely a flood will occur within that class. Furthermore, the negative IV weights show a negative relationship between the class and the occurrence of floods. The first classes of the slope and elevation, for example, received positive IV weights, while the next classes received negative IV weights. This means that as the slope degree and elevation increase, the likelihood of flooding decreases. The distance to the stream of the first four classes indicates a positive correlation with flood occurrence but the rest factor class of the distance to stream shows negative correlations for flood occurrence.

5.1.3 LR

Before LR analysis, a multicollinearity test was performed using TOL and VIF in SPSS software. As indicated in Table 3, no independent variables showed any multicollinearity problem. SPSS software was used to conduct the LR analysis, and LR coefficients for each flood-conditioning component were calculated. Negative LR weights, like IV, show that flood occurrence is negatively related to the conditioning factor. Stream density, NDVI, rainfall, and curvature give negative weights in this study, while other flood-conditioning factors give positive weights.

5.1.4 AHP pairwise comparison matrix

The first row in Table 4 illustrates the importance of the first slope angle compared to the other slope angle classes. For instance, the first slope angle class (0–5°) is more important than the other slope classes, placed in the column and assigned 9. However, the last classes of the slope angle at the row have less significance for flood probability. Therefore, the last class of the slope assigned the reciprocal value (e.g. 1/9 for the last class of the slope, 29–77°). The consistency of the factor class weight was evaluated using the CR. When the CR value is less than 0.1, the weights' consistency is correct. The CR value for all factor classes ranges from 0.014 to 0.09, which is less than 0.1. This showed no weights inconsistency.

Table 2: Statistical analysis results of flood occurrence and flood factors using FR and IV methods

Factors	Class	Class pixel	% Class pixel (*b*)	Flooded area pixel	% Flooded area (*a*)	FR = *a*/*b*	Con_P	Prio_P	Con_P/ Prior_P	IV = ln (Con_P/ Prio_P)
Slope	<5°	11,426,722	45.18	507,460	95.99	2.12	0.044	0.02	2.12	0.75
	5–11°	6,780,625	26.81	18,543	3.51	0.13	0.003	0.02	0.13	−2.03
	11–19°	4,173,258	16.50	2,457	0.46	0.03	0.001	0.02	0.03	−3.57
	19–29°	2,159,660	8.54	207	0.04	0.00	0.000	0.02	0.00	−5.38
	29–77°	750,796	2.97	8	0.00	0.00	0.000	0.02	0.00	−7.58
Elevation (m)	1,774–1,972	10,237,743	40.48	523,039	98.93	2.44	0.051	0.02	2.44	0.89
	1,972–2,220	6,383,841	25.24	5,292	1.00	0.04	0.001	0.02	0.04	−3.23
	2,220–2,513	5,148,369	20.36	344	0.07	0.00	0.000	0.02	0.00	−5.75
	2,513–2,979	3,038,070	12.01	0	0.00	0.00	0.000	0.02	0.00	
	2,979–4,037	483,038	1.91	0	0.00	0.00	0.000	0.02	0.00	
Flow accumulation	Very low	25,250,270	99.84	524,941	99.29	0.99	0.021	0.02	0.99	−0.01
	Low	25,502	0.10	1,653	0.31	3.10	0.065	0.02	3.10	1.13
	Moderate	8,076	0.03	1,037	0.20	6.14	0.128	0.02	6.14	1.82
	High	3,257	0.01	532	0.10	7.81	0.163	0.02	7.81	2.06
	Very high	3,956	0.02	512	0.10	6.19	0.129	0.02	6.19	1.82
Distance to stream (m)	0–100	1,310,596	5.18	75,517	14.28	2.76	0.058	0.02	2.76	1.01
	100–300	2,399,920	9.49	99,494	18.82	1.98	0.041	0.02	1.98	0.68
	300–500	2,288,224	9.05	73,168	13.84	1.53	0.032	0.02	1.53	0.43
	500–700	2,153,831	8.52	53,693	10.16	1.19	0.025	0.02	1.19	0.18
	700–6116.5	17,138,490	67.77	226,803	42.90	0.63	0.013	0.02	0.63	−0.46
Stream density (km^2)	0–0.8	6,882,039	27.92	46,291	8.76	0.31	0.007	0.02	0.31	−1.16
	0.8–2.1	4,983,095	20.21	58,174	11.00	0.54	0.012	0.02	0.54	−0.61
	2.1–3.5	6,317,902	25.63	99,289	18.78	0.73	0.016	0.02	0.73	−0.31
	3.5–5.1	4,350,662	17.65	174,722	33.05	1.87	0.040	0.02	1.87	0.63
	5.1–8.8	2,118,013	8.59	150,199	28.41	3.31	0.071	0.02	3.31	1.20
Slope curvature	Concave	4,388,463	17.35	71,032	13.44	0.77	0.016	0.02	0.77	−0.26
	Flat slope	11,840,022	46.82	296,510	56.09	1.20	0.025	0.02	1.20	0.18
	Convex slope	9,062,576	35.83	161,133	30.48	0.85	0.018	0.02	0.85	−0.16
LULC	Waterbody	170,378	0.67	53,875	10.19	15.13	0.316	0.02	15.13	2.72
	Forest	1,584,350	6.27	25,202	4.77	0.76	0.016	0.02	0.76	−0.27
	Bush	185,078	0.73	8,200	1.55	2.12	0.044	0.02	2.12	0.75
	Settlements	291,928	1.15	9,189	1.74	1.51	0.031	0.02	1.51	0.41
	Wetland	72,446	0.29	812	0.15	0.54	0.011	0.02	0.54	−0.62
	Bare land	2,288,296	9.05	52,119	9.86	1.09	0.023	0.02	1.09	0.09
	Grazing land	1,017,397	4.02	71,962	13.61	3.38	0.071	0.02	3.38	1.22
	Crop land	19,678,779	77.82	307,316	58.13	0.75	0.016	0.02	0.75	−0.29
NDVI	Very low	101,398	0.26	3,352	0.63	2.47	0.033	0.01	2.47	0.90
	Low	18,762,295	47.48	209,315	39.59	0.83	0.011	0.01	0.83	−0.18
	Moderate	11,571,866	29.28	165,084	31.23	1.07	0.014	0.01	1.07	0.06
	High	6,389,220	16.17	100,817	19.07	1.18	0.016	0.01	1.18	0.17
	Very High	2,692,708	6.81	50,107	9.48	1.39	0.019	0.01	1.39	0.33
Soil texture	Loam	17,424,445	68.91	61,780	11.69	0.17	0.004	0.02	0.17	−1.77
	Silty sand	33,442	0.13	2,502	0.47	3.58	0.075	0.02	3.58	1.27
	Clay	4,514,511	17.85	462,543	87.50	4.90	0.102	0.02	4.90	1.59
	Sandy loam	3,313,428	13.10	1,809	0.34	0.03	0.001	0.02	0.03	−3.65
Groundwater depth (m)	1,750–1,951	11,643,964	46.04	514,021	97.23	2.11	0.044	0.02	2.11	0.75
	1,951–2,202	6,130,330	24.24	9,137	1.73	0.07	0.001	0.02	0.07	−2.64
	2,202–2,467	3,367,218	13.31	3,792	0.72	0.05	0.001	0.02	0.05	−2.92
	2,467–2,664	2,064,256	8.16	1,725	0.33	0.04	0.001	0.02	0.04	−3.22
	2,664–2,902	2,085,293	8.25		0.00	0.00	0.000	0.02	0.00	
Rainfall (mm)	83–96	1,207,382	4.77	1,763	0.33	0.07	0.001	0.02	0.07	−2.66
	96–106	1,641,787	6.49	6,480	1.23	0.19	0.004	0.02	0.19	−1.67
	106–113	8,856,030	35.02	362,318	68.53	1.96	0.041	0.02	1.96	0.67
	113–118	8,706,231	34.42	139,032	26.30	0.76	0.016	0.02	0.76	−0.27
	118–125	4,879,631	19.29	19,082	3.61	0.19	0.004	0.02	0.19	−1.68

Note: IV is information value, FR is frequency ratio, *a* is flooded area in a factor class, *b* is an area of factor class, Con_P is conditional probability, and Prio_P is the prior probability.

5.1.5 Flood susceptibility model

The four flood susceptibility indexes (Figure 5) were reclassified into five susceptibility classes of very low, low, moderate, high, and very high using the natural break method (Table 5). As shown in Table 5, high and very high susceptibility classes of the FR, IV, LR, and AHP models covered about 19.8 and 20.7%, 20.3 and 20.2%, 13.2 and 9.3%, 19.8 and 10.2% of the study area, respectively. However, the remaining, 14.1, 23.6, and 21.7%; 13.1, 23.9, and 22.5%; 54.3, 11.2, and 12.1%; and 19.7, 24.8, and 25.6% of the study area, were covered by very low, low, and moderate flood susceptibility areas, respectively (Table 5).

5.2 Discussion

5.2.1 Correlation of flood factors and flood occurrence

The IV, AHP, LR, and FR methodologies were used to assess the relationship between flood incidence and flood governing factors (Table 2). All of the methods were successful in determining the effects of each factor and factor class on the occurrence of floods. Based on LR coefficients, stream density, NDVI, rainfall, and curvature showed a negative relationship with flood occurrence while other flood factors such as slope, aspect, distance to stream, groundwater depth, soil texture, LULC, and elevation are the most flood-driving factors in the study

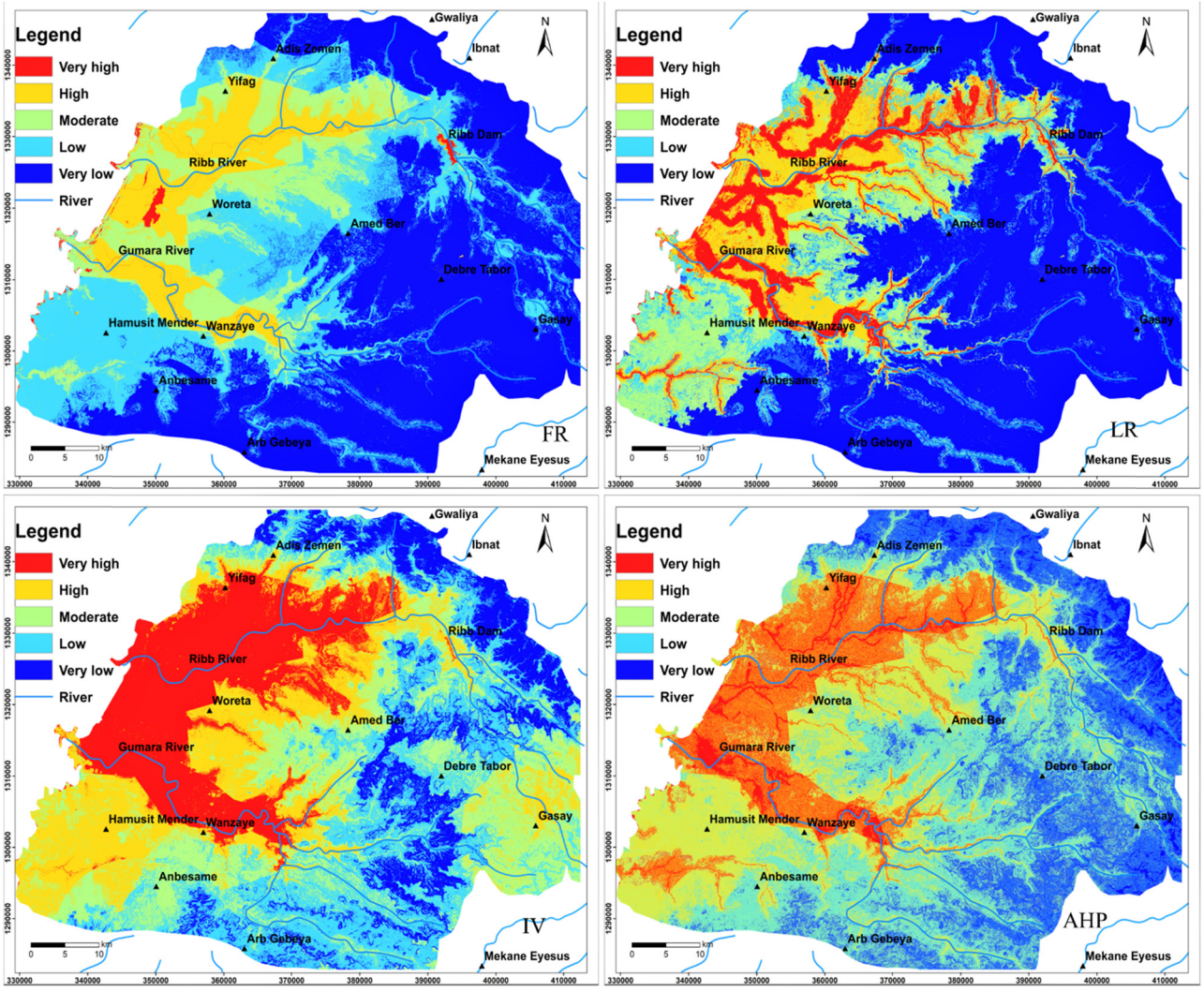

Figure 6: Flood susceptibility maps of FR, IV, LR, and AHP methods (projection: WGS 1984 UTM Zone 37N). The red color represented a very high flood susceptibility area; the yellow color represented high flood susceptibility area, but green, light blue, and blue color represent moderate, low, and very low flood susceptibility areas, respectively.

area. Statistical analysis was also performed to determine the effects of factor class for flood occurrence. For example, the first four classes of distance to a stream showed positive correlations, while the rest factor class of distance to stream exhibited negative correlations. The slope gradient larger than 5° has a weaker correlation with flood occurrence. This finding supported the theory that as the slope angle increases, the likelihood of flooding in a particular train decreases [84,85]. The flat class of curvature showed very good correlation with flood occurrence (Table 1), and 56.1% of the flooded area belongs into this category. A flat class exhibited a positive correlation due to its higher capacity for rainwater concentration and low infiltration of precipitation due to its impermeable soil behavior in class. Other investigations [61,86,87] confirmed this findings. When the elevation of the region increases, the probability of flood occurrence decreases [86]. This suggested that the first class of 1,774–1,972 m had a great flood probability association. This area accounts for 99% of the flooded land. This study revealed that the likelihood of flooding in elevated lands is lower than in low-lying lands [86]. This conclusion is consistent with the results of an earlier research [86]. Distance to the river is crucial for flood occurrence [87]. The first four classes of distance to streams/rivers (0–100, 100–300, 300–500, and 500–700 m) showed a substantial relationship with flood occurrence. Furthermore, these classes account for 57.1% of flooded land. As the distance to the riverbed increases, the probability of flood occurrence decreases (Table 2). This result confirmed the theory that the closer you get to the riverbank, the more likely you are to get flooded [86,87]. One of the most important criteria in flood susceptibility mapping is flow accumulation [87]. As the flow accumulation increase, the FR, IV, and AHP weight (W) value increase. In this study, the River zone or water body, barren area, grazing land, settlement, and bush of all land use land cover classes showed a higher relationship with flooding. Furthermore, these land-use classifications account for 37% of flooded land. Because of their impermeable characteristics, urban and pastureland, indicating a higher flood incidence association. This conclusion is consistent with the findings of an earlier study [86]. Higher NDVI indicated dense vegetation that can reduce and slow water flow [87]. This gives the water time to infiltrate into the ground and decrease surface water volume and less probability of flood occurrence. However, it depends on the hydraulic behavior of soil and the depth of groundwater. As the vegetation density increase, the flood susceptibility of a region decreases depending on the depth of groundwater and vegetation type. In this study, the first, third, fourth, and fifth classes of the NDVI have shown a higher flood occurrence correlation due to shallow groundwater depth. The study result showed that as a stream density increases, flood occurrence increases [86,87]. The stream density classes (3.5–5.1 and 5.1–8.8) have indicated a strong correlation with flood occurrence, and 61.5% of flooded area falls in these classes.

The hydraulic behavior of the region's soil determines the quantity of surface water concentration and rainwater infiltration rate. The water penetration rate into the ground will be higher when the soil mass in a region is extremely pervious, but the amount of surface water concentration will be lower. This will increase the likelihood of a nonflood occurrence in a given area. The depth of groundwater, however, will have a significant impact. The result of this study showed that silty sand and clay soil types have a strong correlation with flood occurrence than loam and sandy loam soil masses. The fraction of pore space between soil grains increases as the grain size of the soil mass increases, but the pore space diameter decreases as grain size decreases. This causes a restriction in the passage of water within the soil. These soil types will be able to hold a lot of water. This boosted the water flow overland once more. Furthermore, the silty sand and clay soil types account for 88% of the inundated area. The shallow groundwater class indicates a significant correlation with the likelihood of flooding. 97.2% of the flooded area is covered by groundwater with a shallow depth.

Even though rainfall is one of the most prominent flood drivers, the nature of the land and the depth of the river channel heavily influence its impact. The annual mean rainfall of class (106–113 mm) indicates a strong relationship with the likelihood of flood incidence. This is due to the impermeable hydraulic behavior of the soil mass, as well as the low slope angle and shallow groundwater depth. The class (106–113 mm) covers 68.5% of the flooded area.

5.2.2 Model validation and comparison

The flood susceptibility class assessment values were calculated for each of the four models using a testing dataset to validate and a training dataset to compare performance. As indicated in Table 5 and Figure 7, in evaluating the model's performance, the FR model (AUC = 97.9%) indicated superior in classifying flooded pixels correctly followed by AHP (AUC = 82.5%), LR (AUC = 75.6%), and then the IV (AUC = 71%). In the case of model prediction accuracy, the FR model again indicated superior performance (AUC = 99.1%), followed by the AHP model (AUC = 86.9%), LR model (AUC = 81.4%), and then the IV model (AUC = 78.2%). Therefore, in this study, the FR model

indicated the highest model accuracy and prediction capacity from the AUC results while the IV model demonstrated relatively less model accuracy and predictive capability. Moreover, the four models (FR, AHP, LR, and IV) resulted in AUC >75%, which is good, very good, and excellent model performance [76], respectively. This finding is similar to the results of earlier research [1,87]. Besides the ROC curve, flood-testing datasets that are not used for model development were overlaid on the four flood susceptible maps. The number of flood points that fall in the very high susceptibility class was measured. As shown in Table 4, 85.2, 55.3, 85.1, and 93.92% of flood points fall in the very high susceptibility class of FR, LR, IV, and AHP models. Here also, the FR and AHP models confirm again its excellent performance followed by the IV model.

The results of this study were also compared with existing studies considering AUC values. In literature, Turoğlu and Dölek [87] stated that the FR model (AUC = 83.69%) is better than the prediction rate of 81.22% of the IV method. This finding is analogous to the present study. In the present study, the FR method showed better performance for both success rates (AUC = 97.9%) and predictive rate curve (AUC = 99.1%) than the IV method of success rate curve (AUC = 71.0%) and predictive rate curve (AUC = 78.2%). As shown from an earlier study [87], the LR model shows high predictive accuracy of AUC value of 79.45% in comparison to the FR and IV models with prediction rate curve value (AUC = 67.33% for FR, AUC = 78.18% for IV). Nevertheless, within the present study model, the FR model showed a comparatively large difference in prediction rate value (AUC = 99.1%) than the IV and LR models with prediction rate value (AUC = 78.2% for IV, AUC = 81.4% for LR). From the results of an earlier study [50], the FR and AHP models showed a bit different in predictive capacity, which is 96.57% for the FR model and 94.92% for the AHP model. This result is in line with the present work. The prediction rate of 99.1% using the FR model is the best performance than the prediction rate curve of 86.9% for the AHP model.

Rahman et al. [88] found that the LR (AUC = 86.8%) gave a more realistic flood susceptibility map than the FR (AUC = 85.6%) and AHP (AUC = 64%) model. However, this result is not in line with the present work, which is that the FR is better than AHP and the LR model. This difference happens mostly due to the amount of and kinds of input parameters for model construction. Generally, the AHP, IV, FR, and LR statistical methods in the literature and this study show a closer prediction capacity with AUC >64% and AUC >75%, respectively. Their AUC values fall within the range of good and very good/excellent performance [87].

The four models' flood validation results (FR, LR, IV, and AHP) are closer to each other. Therefore, from these results, the research work finds out that in flood susceptibility mapping, the four models have equal potential to get flood-prone areas, but factor selection should be playing a more important role than the methods.

Although individual flood susceptibility models are valuable for regional land use planning, they will bring confusing decision-makers and planners if they use all maps as a suggestion [87]. For this, within the present study, we recommended one flood susceptibility map to assist planners to spot suitable areas for a different purpose by evaluating the relative importance among the four models via the area under the curve method (AUC).

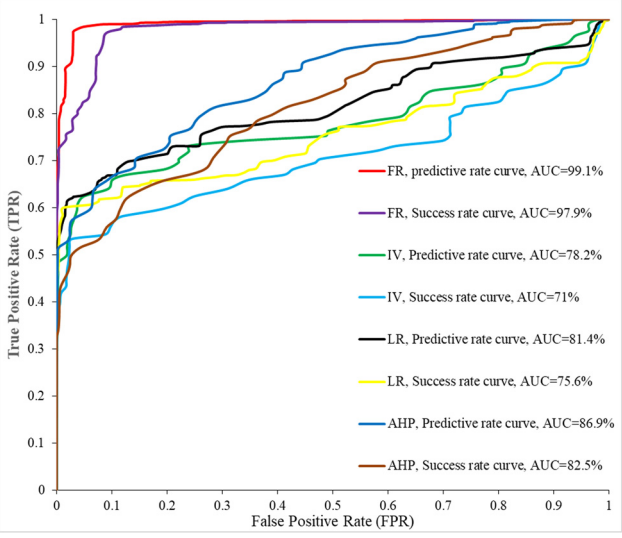

Figure 7: ROC curve of predictive and success rate curves for IV, LR, FR, and AHP methods in which *Y*-axis is true-positive rate value and *X*-axis is false-positive rate value.

Table 3: Logistic coefficients of flood factors and multicollinearity statistics

Factors	LR coefficients (β)	Collinearity statistics	
		TOL	VIF index
Curvature	−0.04	0.983	1.017
Elevation	0.804	0.441	2.267
Flow accumulation	0.222	0.957	1.045
Groundwater depth	0.006	0.485	2.062
LULC	0.159	0.947	1.056
NDVI	−1.198	0.925	1.081
Rainfall	−0.148	0.652	1.534
Slope	0.769	0.608	1.644
Soil texture	0.106	0.58	1.724
Distance to stream	1.73	0.61	1.641
Stream density	−0.095	0.65	1.538
Constant	−4.383		

Table 4: Pairwise comparison matrix and weight of flood factor and classes

Factors	Subfactor class	Subfactor class					
Slope (W_i = 22%)	Class	<5°	5°–11°	11°–19°	19°–29°	>29°	W_i
	<5°	**1**	2	5	7	9	**0.509**
	5–11°	0.5	**1**	2	3	4	**0.229**
	11–19°	0.2	0.5	**1**	2	5	**0.143**
	19–29°	0.14	0.33	0.5	**1**	2	**0.075**
	>29°	0.11	0.25	0.2	0.5	**1**	**0.044**
Elevation (W_i = 8%)	Class (m)	1,774–1972	1,972–2,220	2,220–2,513	2,513–2,979	>2,979	W_i
	1,774–1,972	**1**	3	6	7	9	**0.543**
	1,972–2,220	0.33	**1**	3	3	4	**0.222**
	2,220–2,513	0.17	0.33	**1**	2	5	**0.123**
	2,513–2,979	0.14	0.33	0.5	**1**	2	**0.071**
	>2,979	0.11	0.25	0.2	0.5	**1**	**0.042**
Flow (W_i = 3%)	Class	Very low	Low	Moderate	High	Very high	W_i
	Very low	**1**	0.33	0.11	0.11	0.11	**0.031**
	Low	3	**1**	0.33	0.33	0.33	**0.092**
	Moderate	9	3	**1**	0.33	0.33	**0.186**
	High	9	3	3	**1**	1	**0.346**
	Very high	9	3	3	1	**1**	**0.346**
Distance stream (W_i = 21%)	Class (m)	<100	100–300	300–500	500–700	>700	W_i
	<100	**1**	3	5	9	9	**0.529**
	100–300	0.33	**1**	3	3	5	**0.229**
	300–500	0.2	0.33	**1**	3	7	**0.147**
	500–700	0.11	0.33	0.33	**1**	1	**0.053**
	>700	0.11	0.2	0.14	1	**1**	**0.042**
Stream density (W_i = 4%)	Class (km^2)	<0.8	0.8–2.1	2.1–3.5	3.5–5.1	>5.1	W_i
	<0.8	**1**	0.33	0.2	0.14	0.11	**0.033**
	0.8–2.1	3	**1**	0.33	0.2	0.14	**0.064**
	2.1–3.5	5	3	**1**	0.33	0.14	**0.124**
	3.5–5.1	7	5	3	**1**	1	**0.324**
	>5.1	9	7	7	1	**1**	**0.455**
NDVI (W_i = 5%)	Class	Very low	Low	Moderate	High	Very high	W_i
	Very low	**1**	2	5	9	9	**0.489**
	Low	0.5	**1**	3	5	7	**0.282**
	Moderate	0.2	0.33	**1**	3	7	**0.144**
	High	0.11	0.2	0.33	**1**	1	**0.047**
	Very high	0.11	0.14	0.14	1	**1**	**0.039**
Rainfall (W_i = 4%)	Class	83–96	96–106	106–113	113–118	>118	W_i
	83–96	**1**	0.5	0.2	0.11	0.11	**0.033**
	96–106	2	**1**	0.33	0.2	0.14	**0.057**
	106–113	5	3	**1**	0.33	0.14	**0.123**

(*Continued*)

Table 4: *Continued*

Factors	Subfactor class	Subfactor class								
	113–118	9	5	3	**1**	1	**0.335**			
	>118	9	7	7	1	**1**	**0.452**			
GW (W_i = 2%)	Class	1,750–1,951	1,951–2,202	2,202–2,467	2,467–2,664	>2,664	W_i			
	1,750–1,951	**1**	3	5	9	9	**0.522**			
	1,951–2,202	0.33	**1**	3	5	7	**0.256**			
	2,202–2,467	0.2	0.33	**1**	3	7	**0.139**			
	2,467–2,664	0.11	0.2	0.33	**1**	1	**0.046**			
	>2,664	0.11	0.14	0.14	1	**1**	**0.038**			
LULC (W_i = 16%)	Class	Water body	Dense forest	Moderate forest	Settlements	Wetland	Bare land	Grazing land	Agricultural land	W_i
	Water body	**1**	4	4	0.33	2	0.5	0.5	0.33	**0.091**
	Dense forest	0.25	**1**	0.5	0.2	0.14	0.11	0.11	0.11	**0.021**
	Moderate forest	0.25	2	**1**	0.2	0.33	0.11	0.11	0.11	**0.027**
	Settlements	3	5	5	**1**	3	1	2	2	**0.225**
	Wetland	0.5	7	3	0.33	**1**	0.5	0.5	0.33	**0.08**
	Bare land	2	9	9	1	2	**1**	2	1	**0.204**
	Grazing land	2	9	9	0.5	2	0.5	**1**	1	**0.16**
	Agricultural land	3	9	9	0.5	3	1	1	**1**	**0.192**
Soil (W = 16%)	Class	Loam	Silty sand	Clay	Sandy loam	W_i				
	Loam	**1**	3	0.11	4	**0.148**				
	Silty sand	0.33	**1**	0.11	2	**0.07**				
	Clay	9	9	**1**	9	**0.735**				
	Sandy loam	0.25	0.5	0.11	**1**	**0.047**				
Curvature (W_i = 1%)	Class	Concave	Flat slope	Convex slope	W_i					
	Concave	**1**	0.5	0.5	**0.196**					
	Flat slope	2	**1**	2	**0.493**					
	Convex slope	2	0.5	**1**	**0.311**					

Table 5: Statistical model summary of FR, LR, IV, and AHP methods

IVFSI	Class	IVFSP	% FSM	VFP	% VF	LRFSI	LRFSP	% FSM	VFP	% VF
−25 to −15.1	Very low	3226367	13.1	13	0.01	0–0.1	13381271	54.3	627	0.34
−15.1 to −10	Low	5901361	23.9	472	0.26	0.1–0.3	2756345	11.2	4470	2.46
−10 to −5	Moderate	5535540	22.5	4816	2.65	0.3–0.5	2972834	12.1	15071	8.28
−5 to 1.2	High	4996851	20.3	21844	12.00	0.5–0.7	3243717	13.2	61228	33.64
1.2–13	Very high	4982782	20.2	154863	85.09	0.7–1	2288737	9.3	100612	55.28
FRFSI	Class	FRFSP	% FSM	VFP	% VF	Methods	Success rate curve, AUC%	Prediction rate curve AUC%		
4–9	Very low	3480969	14.1	15	0.01	LR	75.6	81.4		
9–14	Low	5825312	23.6	511	0.28	FR	97.9	99.1		
14–19	Moderate	5356987	21.7	4775	2.62	IV	71	78.2		
19–27	High	4874089	19.8	21635	11.89					
27–46	Very high	5105544	20.7	155072	85.20					
AHPFSI	Class	AHPFSP	% FSM	VFP	% VF	AHP	82.5	86.9		
0.5–1.7	Very low	4849344	19.7	0	0.00					
1.7–2.3	Low	6122024	24.8	12	0.01					
2.3–2.9	Moderate	6298368	25.6	558	0.31					
2.9–3.6	High	4887029	19.8	10491	5.76					
3,6–5.3	Very high	2486139	10.1	170947	93.92					

Note: AHPFSI is analytical hierarchy process flood susceptibility index, IVFSI is information value flood susceptibility index, IVFSP is information value flood susceptibility pixel, FSM is flooded susceptibility map, VFP is validation flood pixel, VF is validation flood, LRFSI is logistic regression flood susceptibility index, LRFSP is logistic regression flood susceptibility pixel, FRFSI is frequency ratio flood susceptibility index, and FRFSP is frequency ratio flood susceptibility pixel.

Therefore, we recommend a flood susceptibility model produced using the frequency ratio method (FR) for regional land use planning.

6 Conclusion

The effects of eleven factors on flood occurrence were evaluated using FR, LR (using SPSS), IV, and AHP methods under the GIS environment. Based on the LR coefficients elevation, flow accumulation, groundwater depth, land use land cover, slope, distance to a river, and soil texture have shown a great statistical correlation with flood probability. The research finds out that the gentle slope, intensive rainfall, low-lying land, near the riverbeds, and unplanned land use practice are the major factors for flood occurrence besides stream density and flow accumulation. The final flood susceptibility maps of the study area were produced using a raster calculator and classified into five zones of very low, low, moderate, high, and very high flood susceptibility. As the result shows, the high and very high flood susceptibility areas exposed in the low lying area of the Ribb and Gumara rivers (western and southwestern part of the study area), however, north, northeast, central, south, and southeastern parts of the study area exposed to moderate, low and very low susceptibility areas. A ROC curve was applied to evaluate the accuracy and performance of the LR, FR, AHP, and IV models. The result shows that the applied methods for flood susceptibility mapping in the study area achieved reliable results. The application of FR, IV, LR, and AHP models was evaluated in flood susceptibility mapping, and their results were compared to each other using AUC values. The results showed that the FR method is relatively better than the AHP, LR, and IV methods. However, the prediction accuracy value for all four methods indicates that the FR (AUC = 99.1%), AHP (AUC = 86.9%), LR (AUC = 81.4%), and IV (AUC = 78.2%) methods are capable of producing acceptable flood susceptibility models. The models generated using the bivariate, multivariate statistical, and AHP models can help to understand the flood susceptibility problems in the study area. On the other hand, this advancement study helps us to understand the relationship between flood occurrence and flood-driving factors to reduce damages from flood hazards.

Although the resulting maps cannot forecast the time and frequency of a flood can occur, it has provided the spatial distribution of flood probability. These models can also provide important information to the researchers, local

people, government, and planners to reduce the flood hazard problems in the study area. Therefore, the concerned bodies may at the Wereda/District, Zone, Region, and Federal levels take tangible activities to mitigate the flood problem. This may be watershed management on upper catchment, construction of a flood barrier along Riversides, construction of check dams on tributaries, and avoiding permanent activities at the high and very high flood susceptibility areas.

Despite the fact that this study achieved a high level of effectiveness and accuracy in mapping flood-prone areas within the Lake Tana sub-basin, it had a number of limitations. As a result, future research should address the following constraints in order to provide a more complete picture of flooding in the region. This will make flood control decisions in the area much easier. These are the following: For hydrological research and flow rate analyses, there is a lack of information on stream discharge rates. Adequate flood data is missing from archived old newspapers, unpublished reports, and/or peer-reviewed works. Although geology is one of the important factors in flood susceptibility mapping, it is excluded in this study due to the lack of detailed geological maps. In the future, hazard and risk analysis should be included in flood potentiality assessments in places with high flood-susceptibility zones.

Acknowledgements: We would like to thank the University of Gondar for financial and equipment supports. We also would like to thank all contributing project partners. We also want to give our special thanks to the West Amhara Meteorological Agency, AWWDE, and Ethiopian Mapping Agency.

Author contributions: The first author has conceptualized the statistical analysis; done the completed modeling process, and wrote the original drafts, reviewed and edited by all coauthors. Although all authors have actively participated in data collection and data organization, the second author has dedicated a lot to data collection and data organization.

Conflict of interest: We declare that we do not any conflict of interest.

References

[1] Samanta RK, Bhunia GS, Shit PK, Pourghasemi HR. Flood susceptibility mapping using geospatial frequency ratio technique: a case study of Subarnarekha River Basin, India. Model Earth Syst Env. 2018;4:395–408. doi: 10.1007/s40808-018-0427-z.

[2] Gain AK, Mojtahed V, Biscaro C, Balbi S, Giupponi C. An integrated approach of flood risk assessment in the eastern part of Dhaka City. Nat Hazards. 2015;79:1499–530. doi: 10.1007/s11069-015-1911-7.

[3] Nageswara Rao G. Occurrence of heavy rainfall around the confluence line in monsoon disturbances and its importance in causing floods. Proc Indian Acad Sci Earth Planet Sci. 2001;110:87–94. doi: 10.1007/bf02702232.

[4] Scheuer S, Haase D, Volk M. Integrative assessment of climate change for fast-growing urban areas: Measurement and recommendations for future research. PLoS One. 2017;12:1–27. doi: 10.1371/journal.pone.0189451. PMID: 29232695.

[5] Khosravi K, Pourghasemi HR, Chapi K, Bahri M. Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: a comparison between Shannon's entropy, statistical index, and weighting factor models. Env Monit Assess. 2016;188:656. doi: 10.1007/s10661-016-5665-9. PMID: 27826821.

[6] Charlton R, Fealy R, Moore S, Sweeney J, Murphy C. Assessing the impact of climate change on water supply and flood hazard in Ireland using statistical downscaling and hydrological modelling techniques. Clim Change. 2006;74:475–91. doi: 10.1007/s10584-006-0472-x.

[7] Tehrany MS, Pradhan B, Jebur MN. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. J Hydrol. 2013;504:69–79. doi: 10.1016/j.jhydrol.2013.09.034.

[8] Tehrany MS, Pradhan B, Mansor S, Ahmad N. Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. Catena. 2015;125:91–101. doi: 10.1016/j.catena.2014.10.017.

[9] Calil J, Beck MW, Gleason M, Merrifield M, Klausmeyer K, Newkirk S. Aligning natural resource conservation and flood hazard mitigation in California. PLoS One. 2015;10:1–14. doi: 10.1371/journal.pone.0132651. PMID: 26200353.

[10] Zou Q, Zhou J, Zhou C, Song L, Guo J. Comprehensive flood risk assessment based on set pair analysis-variable fuzzy sets model and fuzzy AHP. Stoch Env Res Risk Assess. 2013;27:525–46.

[11] Adger NW. Vulnerability. Glob Env Chang. 2006;16:268–81.

[12] Dandapat K, Panda GK. Flood vulnerability analysis and risk assessment using analytical hierarchy process. Model Earth Syst Env. 2017;3:1627–46. doi: 10.1007/s40808-017-0388-7.

[13] Das S. Geographic information system and AHP-based flood hazard zonation of Vaitarna basin, Maharashtra, India. Arab J Geosci. 2018;4:11. doi: 10.1007/s12517-018-3933-4.

[14] Dou X, Song J, Wang L, Tang B, Xu S, Kong F, et al. Flood risk assessment and mapping based on a modified multi-parameter flood hazard index model in the Guanzhong Urban Area, China. Stoch Env Res Risk Assess. 2018;32:1131–46. doi: 10.1007/s00477-017-1429-5.

[15] Guo E, Zhang J, Ren X, Zhang Q, Sun Z. Integrated risk assessment of flood disaster based on improved set pair analysis and the variable fuzzy set theory in central Liaoning Province, China. Nat Hazards. 2014;74:947–65. doi: 10.1007/s11069-014-1238-9.

[16] Tehrany MS, Pradhan B, Jebur MN. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. J Hydrol. 2014;512:332–43.
[17] Bui DT, Ngo PTT, Pham TD, Jaafari A, Minh NQ, Hoa PV, et al. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. Catena. 2019;179:184–96.
[18] Rahmati O, Pourghasemi HR, Zeinivand H. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. Geocarto Int. 2016;31:42–70.
[19] Naulin JP, Payrastre O, Gaume E. Spatially distributed flood forecasting in flash flood prone areas: Application to road network supervision in Southern France. J Hydrol. 2013;486:88–99. doi: 10.1016/j.jhydrol.2013.01.044.
[20] Jacinto R, Grosso N, Reis E, Dias L, Santos FD, Garrett P. Continental Portuguese Territory Flood Susceptibility Index – Contribution to a vulnerability index. Nat Hazards Earth Syst Sci. 2015;15:1907–19.
[21] Dawson CW, Abrahart RJ, Shamseldin AY, Wilby RL. Flood estimation at ungauged sites using artificial neural networks. J Hydrol. 2006;319:391–409.
[22] Chapi K, Singh VP, Shirzadi A, Shahabi H, Bui DT, Pham BT, et al. A novel hybrid artificial intelligence approach for flood susceptibility assessment. Environ Model Softw. 2017;95:229–45.
[23] Shafizadeh-Moghadam H, Valavi R, Shahabi H, Chapi K, Shirzadi A. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. J Env Manag. 2018;217:1–11.
[24] Amade N, Painho M, Oliveira T. Geographic information technology usage in developing countries – A case study in Mozambique. Geo Spat Inf Sci. 2018;21:331–45.
[25] Chen W, Hong H, Li S, Shahabi H, Wang Y, Wang X, et al. Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. J Hydrol. 2019;575:864–73.
[26] Rahmati O, Pourghasemi HR, Zeinivand H. Ls in the Golastan Province, Iran. Geocarto Int. 2016;31:42–70. doi: 10.1080/10106049.2015.104.
[27] Saaty TL. The analytic hierarchy process. New York: McGraw-Hill New York; 1980. p. 287.
[28] Rao M, Sastry S, Yadar P, Kharod K, Pathan S, Dhinwa P, et al. A weighted index model for urban suitability assessment – a GIS approach. Bombay: Bombay Metropolitan Regional Development Authority; 1991.
[29] Rozos D, Bathrellos G, Skillodimou H. Comparison of the implementation of rock engineering system and analytic hierarchy process methods, upon landslide susceptibility mapping, using GIS: a case study from the Eastern Achaia County of Peloponnesus, Greece. J Env Earth Sci. 2011;63(1):49–63.
[30] Bathrellos GD, Gaki-Papanastassiou K, Skilodimou HD, Skianis GA, Chousianitis KG. Assessment of rural community and agricultural development using geomorphological–geological factors and GIS in the Trikala prefecture (Central Greece). J Stoch Environ Res Risk Assess. 2013;27(2):573–88.
[31] Subramanian N, Ramanathan R. A review of applications of Analytic Hierarchy Process in operations management. Int J Prod Econ. 2012;138(2):215–41.
[32] Kayastha P, Dhital MR, De Smedt F. Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: a case study from the Tinau watershed, west Nepal. J Comput Geosci. 2013;52:398–408.
[33] Rahmati O, Zeinivand H, Besharat M, Risk. Flood hazard zoning in Yasooj region, Iran, using GIS and multi-criteria decision analysis. J Geomatics, Nat Hazards. 2016b;7(3):1000–17.
[34] Vogel R. Methodology and software solutions for multicriteria evaluation of floodplain retention suitability. J Cartography Geographic Inf Sci. 2016;43(4):301–20.
[35] Das S, Pardeshi SD. Comparative analysis of lineaments extracted from Cartosat, SRTM and ASTER DEM: a study based on four watersheds in Konkan region, India. J Spat Inf Res. 2018;26(1):47–57.
[36] Rahman M, Ningsheng C, Islam MM, Dewan A, Iqbal J, Washakh RMA, et al. Flood Susceptibility Assessment in Bangladesh Using Machine Learning and Multi-criteria Decision Analysis. J Earth Systems. Environment. 2019;3(3):585–601.
[37] Bednarik M, Yilmaz I, Marschalko M. Landslide hazard and risk assessment: a case study from the Hlohovec–Sered' landslide area in south-west Slovakia. Nat Hazards. 2012;64:547–75. doi: 10.1007/s11069-012-0257-7.
[38] Chen Z, Wang J. Landslide hazard mapping using a logistic regression model in Mackenzie Valley. Can Nat Hazards. 2007;42:75–89.
[39] Pradhan B, Mansor S, Pirasteh S, Buchroithner M. Landslide hazard and risk analyses at a landslide-prone catchment area using the statistical-based geospatial model. Int J Remote Sens. 2011;32(14):4075–87. doi: 10.1080/01431161.2010.4844331559.
[40] Regmi AD, Yoshida K, Pourghasemi HR, Dhital MR, Pradhan B. Landslide susceptibility mapping along Bhalubang-Shiwapur area of mid-western Nepal using frequency ratio and conditional probability models. Jour Mt Sci. 2014;11(5):1266–85.
[41] Wang HB, Wu SR, Shi JS, Li B. Qualitative hazard and risk assessment of landslides: a practical framework for a case study in China. Nat Hazards. 2013;69:1281–94. doi: 10.1007/s11069-011-0008-1.
[42] Dai FC, Lee CF. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology. 2002;42:213–28.
[43] Donati L, Turrini MC. An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology applied to an area of the Apennines (Valnerina; Perugia, Italy). Engg Geol. 2002;63:277–89.
[44] Ayalew L, Yamagishi H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains. Cent Jpn Geomorphology. 2005;65:15–31.
[45] Duman TY, Can T, Gokceoglu C, Nefesliogocu HA, Sonmez H. Application of logistic regression for landslide susceptibility zoning of Cekmee area, Istanbul, Turkey. Verlag. 2006;51(2):241–56.
[46] Sarkar S, Rjan Martha T, Roy A. Landslide susceptibility assessment using information value method in parts of the Darjeeling Himalayas. Vol. 82. Geological Society of India; 2013. p. 351–62.
[47] Meten M, Bhandary NP, Yatabe R. GIS-based frequency ratio and logistic regression modeling for landslide susceptibility mapping of Debre Sina area in central Ethiopia. J Mt Sci. 2015;12(6):1355–72.

[48] Chandak PG, Sayyed SS, Kulkarni YU, Devtale MK. Landslide hazard zonation mapping using information value method near Parphi village in Garhwal Himalaya. Ljemas. 2016;4:228–36.
[49] Kouhpeima S, Feizniab H, Ahmadib Moghadamniab AR. Landslide susceptibility mapping using logistic regression analysis in Latyan catchment. Desert. 2017;22(1):85–95.
[50] Wubalem A, Meten M. Landslide susceptibility mapping using information value and logistic regression models in Goncha Siso Eneses area, northwestern Ethiopia. SN Appl Sci, Switz AG. 2020;2:807. doi: 10.1007/s42452-020-2563-0.
[51] Hong H, Junzhi L, A-Xing Z. Modeling landslide susceptibility using logit Boost alternating decision trees and forest by penalizing attributes with the bagging ensemble. Sci Total Environ. 2020;718:3–15. doi: 10.1007/s00477-012-0598-5.
[52] Pham BT, Prakashi I, Singh SK, Shizardi A, Shahabi H, Bui DT. Landslide susceptibility modeling using reduce error pruning trees and different ensemble techniques: hybrid machine learning approach. Catena. 2019b;175:203–18.
[53] Das G, Lepcha K. Application of logistic regression (LR) and frequency ratio (FR) models for landslide susceptibility mapping in Relli Khola river basin of Darjeeling Himalaya India. SN Appl Sci. 2019;1:1453. doi: 10.1007/s42452-019-1499.
[54] Tehrany M, Jones S. evaluating the variations in the flood susceptibility maps accuracies due to the alterations in the type and extent of the flood inventory. ISPRS-Int Arc Photogramm. Remote Sens Spat Inf Sci. 2017;12:209–14.
[55] Shafapour M, Lalit T, Mustafa K, Jebur N, Shabani F. Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods. Geomatics, Nat Hazards Risk. 2019;10(1):79–101. doi: 10.1080/19475705.2018.1506509.
[56] Ohlmacher GC, Davis JC. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Eng Geol. 2003;69:331–43.
[57] Lee MJ, Kang JE, Jeon S. Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS. Geoscience and Remote Sensing Symposium (IGARSS). Munich: IEEE International; 2012. p. 895–8.
[58] Khosravi K, Nohani E, Maroufinia E, Pourghasemi HR. A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of evidence bivariate statistical models with multicriteria decision-making technique. Nat Hazards. 2016;83:1–41.
[59] Kia MB, Pirasteh S, Pradhan B, Mahmud AR, Sulaiman WNA, Moradi A. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Env Earth Sci. 2012;67:251–64. doi: 10.1007/s12665-011-1504-z.
[60] Liuzzo L, Sammartano V, Freni G. Comparison between different distributed methods for flood susceptibility mapping. Water Resour Manag. 2019;33:3155–73. doi: 10.1007/s11269-01902293-w.
[61] Sarkar S, Kanungo D, Ptra A, Kumar P. Disaster mitigation of debris flow, slope failure, and landslides. GIS-based landslide susceptibility case study in Indian Himalaya. Tokyo, Japan: Universal Acadamy Press; 2006. p. 617–24.
[62] Pradhan B, Lee S. Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modeling. Environ Model & Softw. 2010;25:747–59.
[63] Zhang YS, Igbol J, Yae Y. Susceptibility mapping using an integrated model of information value and logistic regression methods in the Bailongjiang watershed, Gansu province, China. J Mt Sci. 2017;14:249–68.
[64] Chau KT, Chan JE. The regional bias of landslide data in generating susceptibility maps using logistic regression: Case of Hong Kong Island. Landslide. 2005;2:280–90.
[65] Lee S, Sambath T. Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. Env Geol. 2006;50:847–55.
[66] Rickli C, Graf F. Effects of forests on shallow landslides – case studies in Switzerland. For Snow Landsc Res. 2009;82:33–44.
[67] Schicker R, Moon V. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. Geomorphology. 2012;161–162:40–57.
[68] Lee S, Talib JA. Probabilistic landslide susceptibility and factor effect analysis. J Env Geol. 2005;47:982–90.
[69] Cho S, Kim J, Heo E. Application of fuzzy analytic hierarchy process to select the optimal heating facility for Korean horticulture and stockbreeding sectors. Renew Sustain Energy Rev. 2015;49:1075–83.
[70] Nguyen AT, Nguyen LD, Le-Hoai L, Dang CN. Quantifying the complexity of transportation projects using the fuzzy analytic hierarchy process. Int J Proj Manage. 2015;33:1364–76.
[71] Saaty TL. Fundamentals of decision-making and priority theory with the analytic hierarchy process. 6, Pittsburgh: Rws Publications; 2000. p. 326.
[72] Zhang W, Lu J, Zhang Y. Comprehensive evaluation index system of low carbon road transport based on fuzzy evaluation method. Procedia Eng. 2016;37:659–68.
[73] Luu C, Von Meding J, Kanjanabootra S. Assessing food hazard using food marks and analytic hierarchy process approach: a case study for the 2013 food event in Quang Nam. Vietnam Nat Hazards. 2018;90:1031–50.
[74] Saaty TL. Decision making with the analytic hierarchy process. Int J serv Sci. 2008;1:83–9.
[75] Saaty TL. The seven pillars of the analytic hierarchy process. In Köksalan M, Zionts S, editors. Multiple criteria decision making in the new millennium. Vol. 3. Berlin, Heidelberg: Springer; 2001. p. 15–37.
[76] Pourghasemi HR, Pradhan B, Gokceoglu C. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed. Iran Nat Hazards. 2012;63:965–96.
[77] Chauhan S, Sharma M, Arora MK. Landslide susceptibility zonation of the Chamoli region, Garhwal Himalayas, using a logistic regression model. Landslides. 2010;7:411–23.
[78] Cao C, Xu P, Wang Y, Chen J, Zheng L, Niu C. Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas. Sustainability. 2016;8(9):948.
[79] Ullah K, Zhang J. GIS-based flood hazard mapping using relative frequency ratio method: A case study of Panjkora River Basin, eastern Hindu Kush, Pakistan. PLoS One. 2020;15(3):e0229153. doi: 10.1371/journal.pone.0229153.
[80] Yesilnacar E, Topal T. Landslide susceptibility mapping: A comparison of logistic regression and neural networks

method in a medium scale study. Turkey: Hendek Region; 2005, Engine.

[81] Rahmati O, Pourghasemi HR. Identification of critical flood-prone areas in data-scarce and ungauged regions: a comparison of three data mining models. Water Resour Manag. 2017;31(5):1473–87.

[82] Skilodimou HD, Bathrellos GD, Chousianitis K, Youssef AM, Pradhan. B. Multi-hazard assessment modeling via multi-criteria analysis and GIS: a case study. Environ Earth Sci. 2019;78:47. doi: 10.1007/s12665-018-8003-4.

[83] Bathrellosa GD, Skilodimoua HD, Chousianitisb K, Youssef AM, Pradhand B. Suitability estimation for urban development using multi-hazard assessment map. Sci Total Environ. 2017;575:119–34. doi: 10.1016/j.scitotenv.2016.10.025.

[84] Falah F, Rahmati O, Rostami M, Ahmadisharaf E, Daliakopoulos IN, Pourghasemi HR. Artificial neural networks for food susceptibility mapping in data-scarce urban areas. In Pourghasemi HR and Gokceoglu C, editors. Spatial modeling in GIS and R for the earth and environmental sciences. Elsevier; 2019. p. 323–36.

[85] Tien Bui D, Pradhan B, Lofman O, Revhaug I, Dick OB. Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. Comput Geosci. 2012;45:199–211.

[86] Mosavi A, Ozturk P, Chau K-W. Flood prediction using machine learning models: literature review. Water. 2018;10:1536.

[87] Turoğlu H, Dölek I. Floods and their likely impacts on the ecological environment in Bolaman River basin (Ordu, Turkey). RJAS. 2011;43(4):167–73.

[88] Matej V, Jana V. Flood susceptibility mapping on a national scale in slovakia using the analytical hierarchy process. Water. 2019;11:364.