Efficiency Exploration of Frequency Ratio, Entropy, and Weights of Evidence-Information Value Models in Flood Susceptibility Assessment: A Study of Raiganj Subdivision, Eastern India

Sunil Saha
Raiganj University
Debabrata Sarkar
Raiganj University
Prolay Mondal (mon.prolay@gmail.com)
Raiganj University https://orcid.org/0000-0003-3263-4377

Research Article

Keywords: Flood susceptibility mapping, Frequency ratio, Shannon's Entropy, Weights-of-evidence, GIS, Raiganj Subdivision

Posted Date: July 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-729511/v1

Efficiency Exploration of Frequency ratio, Entropy, and Weights of evidence-Information Value models in flood susceptibility assessment: A study of Raiganj Subdivision, Eastern India

Sunil Saha[a] , Debabrata Sarkar[b] , Prolay Mondal[c]*

[a]Research Scholar, Department of Geography, Raiganj University, Raiganj, West Bengal, 733134, India. E-mail: sahasunil2507@gmail.com

[b]Research Scholar, Department of Geography, Raiganj University, Raiganj, West Bengal, 733134, India. E-mail: debabratas077@gmail.com

[c]Assistant Professor, Department of Geography, Raiganj University, Raiganj, West Bengal, 733134, India. E-mail: mon.prolay@gmail.com

*Corresponding Author

Abstract

The primary objective of this research was to assess the efficiency of RS and GIS for predicting the flood risk using the Frequency ratio, Entropy index, and Weight of evidence-information Value models. Consequently, for spatial analyses fourteen flood conditioning variables are constructed. The research region is experiencing floods consequently in the past with moderate to high intensities. The assessment demonstrates that parameters such as elevation, LULC, rainfall, distance from rivers, and drainage density contributed significantly to the occurrence of floods. The validation of the results indicates that the success rate of the presently constructed models. Around 33 per cent to 47 per cent of the total area of each block in the subdivision is projected to be in danger of floods. The report's findings will aid planners in designing flood prevention measures as part of regional flood risk management programs and will serve as a platform for future research in the study area.

Keywords

Flood susceptibility mapping, Frequency ratio, Shannon's Entropy, Weights-of-evidence, GIS, Raiganj Subdivision.

1. Introduction

Flood is the most common and destructive weather phenomenon in India. Furthermore, second to Bangladesh, the most flood-affected country in the world is India with around one-eighth of its total land (40 million hectares) susceptible to flooding. The Central water commission (2010) estimated about 7.6 million hectares of geographical area are severely affected due to flooding each year. Between 13 years (1965-1978) the flood-affected area has been significantly increased from 7.6 million to 17.5 million hectares. Monsoon rainfall in India creates most of the flooding which happens in conjunction with the shifting nature of monsoonal low-pressure centers and cyclones throughout the monsoon period between June to September. Indian rivers namely Ganga, Brahmaputra, Mahanadi, Krishna, Kaveri are known for devastating flooding, massive flooding occurs on these rivers as a consequence of monsoon outbreaks.

Gupta et al. (2003), Singh and Kumar (2013) in their paper investigated that more than 30 million individuals are impacted by flooding in India and more than 1500 people are killed each year, accounting for nearly one-fifth of the world deaths caused by this catastrophe. According to a government estimate (CWC, 2010), floods killed over 92 thousand people and cost $200 billion in economic damage over a 56-year period (1953-2009). Kalsi and Srivastava (2006) reported that cyclone-induced floods in Orissa coastal area (Paradeep) in 1999 caused devastating damages to properties and about 10 thousand people are died within two days (29 and 30 October). In Andhra Pradesh and Orissa, extremely destructive and sudden floods wreak havoc on property and kill a considerable number of people (Dhar et al., 1981a; WMO, 1994; Dhar and Nandargi, 2003). Due to the dam breaking as a consequence of the unexpectedly severe flooding, flash floods might occur downstream. Near 1979, the Machchhu dam collapsed in Morbi town in Saurashtra, destroying unimaginable amounts of land, crops, animals, and almost 10000 human lives (Dhar et al., 1981b; Dhar and Nandargi, 2003). A total of 34 million hectares of land is under threat of flooding as per the national flood commission (NFC) in 1980 report and in the present day this number significantly increases due to climate change and unpredictability. Though the NFC study claimed that the government has secured 10 million hectares of land against flooding, effective protection may only be available for 6 million hectares.

According to Cloke and Pappenberger (2009), it is somewhat impossible to completely avoid or bypass the devastation of floods altogether, however, the severity of this catastrophe

may be somewhat reduced by adequate risk and susceptibility assessment measures. Prevention is better than quire, due to this relevance flood susceptibility evaluation and risk assessment are regarded as an extremely beneficial techniques to minimize flood severity (Sarkar and Mondal, 2020). In the various geospatial environments, the influencing capacities of individual flood conditioning factors are different. It is thus extremely important to evaluate numerous conditions that cause flood before any flood risk assessments. If we consider several flood conditioning factors and performed flood susceptibility mapping and risk assessment the resulted outcomes will more precise and accurate rather, we use one flood-induced variable.

Now a day, in the field of flood susceptibility analysis, geographical information system (GIS) and remotely sensed (RS) information serve as the main ingredient and achieved widespread acceptance (Kafira et al., 2014; Tehrany et al., 2014; Sachdeva et al., 2017; Al-Jauidi et al., 2018; Lee et al., 2018; Dano et al., 2019; Kanani-Sadat et al., 2019; Lappas and Kallioras, 2019; Castache et al., 2020; Pham et al., 2020; Khoirunisa et al., 2021). The RS and GIS techniques are used to gather reliable information about earth surface characteristics, topographic, morphological character, LULC, meteorological information, and a wide range of other associated information related to a specific environment. Concurrently, the GIS platform provides a fascinating environment for flood susceptibility mapping using RS datasets. The RS and GIS methodologies have effectively been used worldwide in the field of flood susceptibility mapping.

Around the globe, the RS and GIS based different numerical, statistical and machine learning methods have been used to assess flood risk and vulnerability such as, flood susceptibility assessment using SVM in Kuala Terengganu (Tehrany et al., 2015), bivariate and multivariate statistical approach based flood susceptibility mapping in Jeddah city (Youssef et al., 2016), FR and WoE model based flood susceptibility analysis in Iran (Khosravi et al., 2016a), SEI, SI and WoE model based flood susceptibility analysis in Iran (Khosravi et al., 2016b), flood susceptibility analysis using FR and WoE in Iran (Golastan province) (Rahmati et al., 2016), flood susceptibility mapping using hybrid AI in Haraz watershed (Chapi et al., 2017), Flood risk assessment using entropy index model in Madarsoo watershed (Haghizadeh et al., 2017), flood susceptibility analysis in Hengfeng using neuro-fuzzy in Hengfeng county (Hong et al., 2018), flood susceptibility analysis using FR model in Markham basin, Papua New

Guinea (Samanta et al., 2018a), flood risk assessment in Subarnarekha river basin using FR technique (Samanta et al., 2018b), flood risk evaluation using probabilistic techniques in Gucheng county (Tang et al., 2018), flood susceptibility assessment using MCDM and fuzzy in Haraz watershed (Azareh et al., 2019), fuzzy logic based flood risk assessment in Koshi river basin (Sahana and Patel, 2019), FR and LR based flood susceptibility assessment in Brisbane catchment, Australia (Tehrany et al., 2019), flood vulnerability mapping in Kulik river basin using frequency ratio model (Sarkar and Mondal, 2020), frequency ratio (FR) and SVM based flood susceptibility assessment in Sundarban BR (Sahana et al., 2020), novel ensemble and SVM based flood risk assessment (Saha et al., 2021),

In this present study, the bivariate techniques such as frequency ratio (FR), Shannon's entropy index (SEI), and weights of evidence (WoE) were utilized along with geospatial information of the study region. These bivariate methods are very effective and widely accepted techniques for susceptibility analysis. The main purpose of the present study is to construct and deploy a quantitative approach using bivariate techniques in conjunction with RS and GIS for flood susceptibility mapping and risk assessment of the study region. For this purpose, various flood conditioning parameters are taken into consideration.

2. Description of the Study Area

Raiganj subdivision is part of Uttar Dinajpur district (Fig. 1). The Uttar Dinajpur district was established on April 1, 1992, with the division of West Dinajpur into Dakshin and Uttar Dinajpur, and after the formation of the Raiganj Sub-division. The research area is spread out over a large geographical area with an extension of 25°14’34’’ N latitudes to 25°50’2.99’’ N latitudes and 88°01’16.49’’ E longitudes to 88°27’33.70’’ E longitudes. Raiganj Sub-division is divided into four Community Development Blocks and four Panchayat Samitis. The study area is around 1328.25 sq.km., or over 45 percent of the district's total area. Bihar in the east, Malda in the south and southwest, Dakshin Dinajpur district in the east, and Bangladesh in the north surround the research region. The older alluvial sediments are found to be from the Pleistocene epoch. Uttar Dinajpur has been blessed with extremely fertile land. The usual average high temperature is 39 °C (102 °F) in July and 26 °C (79 °F) in January, owing to the sedimentary deposition that aids in the growth of rice, Jute, Mesta, and Sugarcane, among other crops. The average annual temperature is around 25 ° (79

°). The study area is situated in the Mahananda-Nagar river basin region where the Kulik, Gamari, Chhiramati (Srimati), and Tangon, Sooin (Saha et al., 2021) are major rivers flowing as slope direction. In the year 2017, a major flash flood wreaked havoc in the Raiganj sub-division, North Dinajpur, W/B. The destruction of the 2017 flood was caused by unplanned construction, incorrect growth of the Raiganj sub-division region, and excessive rainfall in adjacent states and river basins. North Dinajpur and Raiganj have previously experienced floods in 1982, 1987, 1992, 1995, 1998, 2000, 2002, and 2005 and among the mentioned years, the floods of 1987, 1992, 1998, and 2000 had large magnitudes and wreaked havoc on the neighborhood (Saha and Mondal 2020). The flow diagram of the whole study is given in Fig. 2.

Fig. 1 Geographical Location of the research area

Fig.2 Flow diagram of the research

3. Database and Method

3.1. Data Sources

3.1.1. Flood Inventory Map

The probability of a future flood event in a certain area can be predicted by looking at previous flood records (Manandhar 2010). The flood inventory map (Fig. 3) is therefore regarded as a significant component in predicting future floods. In this study, a flood inventory map containing 70 training points was prepared using various sources of the district reports. The flood inventory map was randomly partitioned into (Pradhan 2010; Pourtaghi & Pourghasemi 2014) two datasets such as 40 flood locations and 30 non-flood locations for training and 30 validation points, respectively.

Fig. 3 Flood Inventory mapping of the research area

3.1.2. Flood predisposing Variables

It is believed that the precision of susceptibility simulation improves when the analysis procedure integrates all event control factors. The acquisition of data and the construction of databases are essential parts of any research study. There were fourteen variables selected for the current simulation procedure on the expert's recommendation. The simulated flood susceptibility regions of the study area were afterward determined by FR, SEI, and WoE methodologies.

3.1.3. Physical Variable

The process of flood susceptibility mapping entails determining the flood-conditioning factors (Kia et al. 2012). This study anticipates a total of 10 independent physical factors.

The information about rainfall was collected from the CHRS data portal and thereafter it is verified with the Indian Meteorological Department's data.

The Digital Elevation Models (DEMs) were downloaded from the Alaska Satellite Facility with 12.5 m cell size and thereafter the Elevation of the study area was estimated (Fernandez & Lutz 2010).

Similarly, the Slope of the study area was generated using the above DEMs to access the topographic inclination. The slope orientation regulates surface runoff infiltration and stream flow rate of any region (Adiat et al., 2012).

The erosive force of surface runoff is measured using SPI. Stream Power Index (SPI) was generated from the DEMs using Eq. 1.

$$SPI = C_{area} \times \tan\theta \quad (1)$$

Where, SPI: Stream Power Index; C_{area} : Catchment Area (m^2/m); $\tan\theta$: Angle of the Slope.

Plan Curvature (PC), Flow Direction (FD), and Stream Density (SD) were also generated in the GIS environment using Alaska DEMs. The **predisposing** variable that significantly contributes to flooding prevalence is Stream Density (Eq. 2) (Gül 2013).

$$S_{denssity} = {}^{s_{length}}\!\Big/\!_{C_{area}} \quad (2)$$

Where, $S_{denssity}$: Stream Density (km^2); s_{length} : Length of the streams in the specific catchment; C_{area} : Catchment Area.

The data about geomorphology was collected from the data portal of the Geological Survey of India.

The Normalized Difference Vegetation Index (NDVI) was estimated (Eq. 3) using the pixel values of Near-Infrared Band (IR) and Red Band (RB) of Landsat-8 OLI imageries. This data was collected from the USGS Earth Explorer data portal with a 30 m cell size.

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)} = \frac{(Band5 - Band4)}{(Band5 + Band4)} \quad (3)$$

Where, NDVI: Normalized Difference Vegetation Index; NIR: Near-Infrared Band.

3.1.4. Socio-economic Variable

In order to evaluate the probable flood sensitivity region of the study area, some of the four variables were studied from the socio-economic group. From the Census of India, the statistics on population-related characteristics such as total population, total area, the total number of houses, total number of farmers, were collected to estimate the Population Density, Household Density, and Cultivator Density of the study area.

The River network of the study area was manually digitized from the Google Earth Pro software to measure the distance from the river (Tehrany et al. 2015).

The LULC data was collected from the Bhuvan data portal and thereafter the LULC layer of the study area was generated. A detailed account of variables and their sources is mentioned in table 1.

Table 1 Data and their sources for the research work

3.1.5. The rationale for the Selecting of the variables

Rainfall is one of the prominent flood **predisposing** variables. Massive rainfall leads to a little probability that rainwater will be absorbed by the ground (infiltration), so it will wash off into the river. The quicker the water reaches the river, the greater the chance of flooding (Saha et al. 2020). Excessive rainfall causes flooding when existing watercourses are unable to carry the excess water (Ouma et al. 2014). The elevation influences flood vulnerability (Fernandez and Lutz 2010) by regulating stream direction and producing changes in vegetation and soil properties, which influence the runoff of the research area (Aniya 1985). Steep slopes restrict water penetration into the ground, allowing water to flow quickly down to rivers as surface runoff, causing flooding. SPI was used to identify and analyze areas where soil management actions could reduce surface runoff-induced degradation (Jebur et al., 2014). Plan curvature is the frequency of slope changes in a specific direction (Wilson and Gallant 2000) and it is also determining the flow pattern (Oh and Pradhan 2011). Plan curvature measurements can be used to extract useful geomorphologic data from any area. The alignment and number of cells in the catchment area that contribute flow into a given cell are incorporated into the flow direction study, which reflects not only the drainage network but also the route a potential flood may take (Kourgialas and Karatzas 2011; Sahana et al. 2019). Geomorphology employs numerous research projects, modeling, and simulation to provide a better understanding of landform

histories and dynamics, as well as to identify its future shifts. The variation in reflectance of plant cover between visible and near-infrared wavelengths can be used to calculate the density of vegetation on a plot of land (Weier and Herring, 2000).

The density of the population as intended is a critical measure of vulnerability because the vulnerability is negligible if there is no population. Higher farmers' densities provide greater assurance that crops will be destroyed during flooding. A higher density of households creates a higher certainty of the destruction of houses during a flooding event. Distance from the river is the most significant contributing variable due to its significant effect on flood extent and intensity (Glenn et al. 2012). Improper land use and land cover pattern can entertain every type of susceptibilities to occur. The most common land-use types within the study area are cropland, Vegetation, bare land, built-up land, and water body. Surface runoff and floods are increased in built-up areas, which are primarily made up of impermeable surfaces (Tehrany et al. 2013). Due to the positive link between infiltration capability and vegetation density, however, vegetated areas are less prone to flooding.

3.2. Methods of the study

3.2.1. Frequency Ratio Method (FR)

The predictive relationship between dependent and independent variables can be easily quantified using the FR model. The FR method is a Basic Statistical Analysis method that assesses the impact of each type of conditioning factor on projected flooding (Lee et al., 2012; Sarkar and Mondal 2020). According to Bonham-Carter (1994), FR is the possibility of the emergence of a certain event and it reveals the relationship between the flood sites and their relevant variables.

This approach has been employed in several natural catastrophe circumstances including the estimation of groundwater potentiality (Manap et al., 2014), to estimate the blast-induced air-blast (Keshtegar et al. 2019), landslide susceptibility mapping (Lee and Pradhan 2007; Mondal and Maiti 2013), and flood susceptibility mapping (Lee et al., 2012; Rahamti et al. 2015; Khosravi et al. 2016). The advantage of this model is that it is easy to apply and gives totally understood outcomes (Ozdemir and Altural 2013). Among numerous bivariate statistical approaches, the FR model was used in the current study to assess the flood susceptibility tendency of the study area (Althuwaynee et al. 2014). This method can be expressed by equation 4.

241 $$FR = \frac{Pix_i^v \big/ Pix_i^c}{Pix_t^v \big/ Pix_t^c} \quad (4)$$

242 Where, FR: Frequency Ration; Pix_i^v : Total vegetable pixel of i_{th} alternative; Pix_i^c : Total class
243 pixel of i_{th} alternative; Pix_t^v : Total vegetable pixel of the criteria; Pix_t^c : Total pixel of the
244 criteria.

245 In order to assess the association between the flood site and predictor classes, the Relative
246 Frequency Index (RF) was used. It is the normalized values of the previous frequency ratio
247 (FR) values of the variables. Eq. 5 demonstrates the formulation of FRF for a given class
248 factor field.

249 $$RF = FR_i \Big/ \sum FR^c \quad (5)$$

250 Where, RF: Relative Frequency; FR: Frequency Ratio; FR_i : FR value i_{th} alternative; $\sum FR^c$:
251 Summation of FR values of C criteria.

252 The prediction rate was determined using equation 3 to recognize the relational
253 interrelationships among the independents. The entire procedure can be based on a
254 frequency ratio (FR) model and the individual rating of variables was expressed as the
255 prediction rate (Eq. 6) (Sabatakakis et al., 2013).

256 $$PR = \frac{\left(Max_{RF} - Min_{RF}\right)}{Min_{\left(Max_{RF} - Min_{RF}\right)}} \quad (7)$$

257 Where, PR-Prediction Rate; Max_{RF} -Highest value from the RF values; Min_{RF} -Least value
258 from the RF values; $Min_{\left(Max_{RF} - Min_{RF}\right)}$ - is the minimum value from all selected variables
259 $\left(Max_{RF} - Min_{RF}\right)$.

260 The flood susceptibility was calculated using a factor analysis algorithm that projected the
261 probable influence and predictor interrelationships as a function of factor analysis. The
262 following (Eq. 7) is the proposed algorithm, which combines the amount of the products of
263 fourteen separate factor variables.

$$FSZ = V^{1}_{\otimes PR} \oplus V^{2}_{\otimes PR} \oplus .....V^{n}_{\otimes PR}$$

$$Where, V^{1}_{\otimes PR} = V^{Slope}_{a \otimes RF} .....V^{n}_{\otimes PR} \quad (7)$$

Where, FSZ: Flood Susceptibility Zone; V: Variable; PR: Prediction Rate; a: Alternatives of specific variables; RF: Relative Frequencies of the alternatives of that specific variable.

In each area of pixels, the Flood Susceptibility Zone (FSZ) was generated independently from relative frequency values of the selected fourteen conditioning variables. Following that, the pixel values were classified using a natural break classification scheme (Moghaddam et al., 2013) in the ArcMap-10.5 environment.

3.3.2. Shannon Entropy Weight Method

Entropy is the measure of unpredictability, volatility, imbalanced behaviors, energy distribution, and instability in a system (Pourghasemi et al. 2012; Khosravi et al. 2016). Stephan Boltzmann was the first to propose this theory, while Shannon was the first to present it statistically in 1948. Based on the Boltzmann theorem, Shannon expanded the entropy model for information theory. The entropy model has been widely used in hazard identification and risk management studies to calculate the leverage ratio of natural hazards and to represent the priority of some variables in influencing a particular hazard (Al-Hinai et al. 2021). The core aim of this research work is to measure flood susceptibility of the study area using the entropy technique. The following Eq. 8 was used to normalize the arrays of performance indices to obtain the probability density.

$$Pd_{ij} = \frac{FR_{ij}}{\sum_{i=1}^{m_j} FR_{ij}} \quad (8)$$

Where, Pd_{ij} : Probability Density or Project outcomes; FR_{ij} : Frequency Ratio value of the specific cell.

After calculating the probability density, the entropy measure of these probability densities is estimated using equation 9.

287 $$Ev_j = -\sum_{i=1}^{m_j} Pd_{ij} \log_2 Pd_{ij}, j = 1,...,n \quad (9)$$
$$Ev_{j\max} = \log_2 m_j$$

288 Where, Ev_j & $Ev_{j\max}$: Entropy Value; m_j : Number of classes.

289 Furthermore, the weights of the criteria were estimated using Equations 10 and 11.

290 $$Ic_j = \left(\frac{Ev_{j\max} - Ev_j}{Ev_{j\max}} \right), I = (0,1)1\, j = 1,....n \quad (10)$$

291 $$Cw_j = I_j FR \quad (11)$$

292 Where, Ic_j : information coefficient; Cw_j : Weight of the criteria.

293 The criteria weights of all criteria were estimated using Shannon's entropy technique on MS
294 Excel. After calculating the entropy values, the final flood vulnerability map was generated
295 using the raster calculator tool in the ArcMap environment.

296 **3.3.3. Weight of Evidence and Information Value (WoE-IV)**

297 WoE is a quantitative data-driven strategy for predicting the recurrence of events based on
298 the Bayes rule (Rahmati et al., 2016). This approach, which is based on the Bayesian
299 probability model, has more facilitation when compared to other statistical methods
300 (Tehrany et al., 2014). The WoE model is built around the concept of determining positive
301 and negative weights. Based on the presence or absence of cultivations inside a certain
302 area, the approach estimates the weight of each cultivation conditioning variable. In this
303 study the following Eq. 12 was used to predict the WoE values.

304 $$WoE = \ln\left(\%Pix_t \Big/ \%Pix_v \right) \quad (12)$$

305 Where, WoE : Weight of Evidence; $\ln$: Natural Log; $\%Pix_t$: Percentage of the total pixel of
306 the alternative; $\%Pix_v$: Percentage of the vegetable pixel of the same alternative.

307 A negative weight (W-) reflects the magnitude of negative association and implies the
308 absence of the conditioning variable and a positive weight (W+) denotes the presence of

the conditioning factor at the cultivation sites, and its magnitude reflects the positive association between the presence of the conditioning component and vegetable cultivation (Pradhan et al. 2010).

One of the most useful techniques for selecting important variables in a predictive model is information value (IV). It facilitates the classification of variables according to their importance. The IV is determined using the formula below (Eq. 13).

$$IV = \left(\%Pix_t - \%Pix_v\right) \times \ln\left(\frac{\%Pix_t}{\%Pix_v}\right) \quad (13)$$

Where, $\ln$: Natural Log; $\%Pix_t$: Percentage of the total pixel of the alternative; $\%Pix_v$: Percentage of the vegetable pixel of the same alternative.

3.3.4. Weighted Linear Combination Method (WLCM)

Several scholars have used AHP to investigate the 'Potentiality' of its effective mathematics qualities (Triantaphyllou and Mann, 1995; Srdevic et al. 2011). 'WLC' or 'The 'Weighted Linear Combination Model' (Fig. 8) is intended to rank the criteria in order to get a probable appropriateness index. The priority values of the parameters were evaluated using AHP, and then the Potential Suitability Zoning map was produced utilizing equation 14 in a GIS environment.

$$FSZ_{WLC} = \left(Variable \times Rating_1\right) + ...\left(Variable \times Rating_n\right) \quad (14)$$

Where, FSZ: Flood Susceptibility Zone; WLC: Weighted Linear Combination; n: no. of the variable

4. Discussion of the results

4.1. Execution Frequency ratio model

The FR approach was used to determine the degree of correlation between flood sites and conditioning factors i.e. the frequency ratio model can be used to evaluate the correlation between each flood factor and the distribution of prior floods, which is based on probabilistic (statistical) methodologies (Manandhar 2010; Sarkar and Mondal 2018). Table 2 shows the results of the geographic connection between the flood location and the conditioning factors using the FR model. In general, an FR value of 1 indicates that flood sites and affective factors have an average correlation (Pradhan 2010). There is a high

337 correlation if the FR value is greater than 1 and a lesser correlation if the FR value is less
338 than 1.

339 The relative frequencies of all variables were estimated using equation 5 which is helpful in
340 the calculation of priority values of the individual variable. A higher value of RF depicts a
341 higher susceptibility to flooding and vice-versa. The estimation of FR (RF) for the relationship
342 between flood sites and rainfall, it is found that the rainfall zone (Fig. 4a) of <1750 mm has
343 the highest FR (RF) value (FR: 2.457; RF: 0.374) followed by the zone of 1850-1900 mm
344 rainfall (FR: 1.067; RF: 0.162) which depict a higher probability of flooding. In the case of
345 Elevation, the elevation zone (Fig. 4b) of 10 to 30 m has the larger FR value (FR: 1.639; RF:
346 0.336) which indicates the higher chances of the flood occurring. The slope zone (Fig. 4c) of
347 <1.0° has the higher FR value (FR: 1.008; RF: 0.211) followed by the zone of 1.0-2.0° which
348 depicts the higher probabilities of the flood. The NDVI zones (Fig. 4d) of -0.12 – 0.05 (FR:
349 1.119; RF: 0.228) and 0.05-0.12 (FR: 1.059; RF: 0.216) are most sensitive to flood with the
350 higher values of FR. From the estimation of correlation between the plan curvature and
351 flood sites, it is observed that flat curvature (Fig. 4e) areas are most susceptible to flood
352 having a higher FR value (FR: 1.012; RF: 0.344). Flood locations are centered in the
353 areas where the stream density (Fig. 4f) is 6.5 - 8.1 (FR: 1.866; RF: 0.303), SPI is 0.24-4.05
354 (FR: 1.010; RF: 0.204) (Fig. 4h). It is observed that the higher stream density indicates higher
355 FR values which is an indicator of higher flooding probabilities. It is because the infiltration
356 of surface water is very low and the water carrying capacity of lands has become very low in
357 the places where the density is higher in the rainy season. The flow direction zone (Fig. 4g)
358 of South (FR: 1.045; RF: 0.131) and West (FR: 1.032; RF: 0.130) is most vulnerable to flood
359 with a higher FR rating. The estimation of FR (RF) for the depiction of the relationship
360 between flood sites and geomorphology is also carried out. It is observed that an old flood
361 plain-water body is more susceptible to flood with FR 2.402 (RF: 0.442) followed by the old
362 alluvial plain and the Young alluvial plain (Fig. 4i). The distance from the river (Fig. 4j) zone
363 <1.2 km is more prone to flood with FR 1.255 (RF: 0.270). From the result, it is observed that
364 the frequency of flood increased with the decreasing distance from the river. The results
365 demonstrated that the flooding event mostly occurs where the Household Density (Fig. 4k)
366 is 160 to 180 (FR: 1.427; RF: 0.280), Cultivator Density (Fig. 4l) is 88-98 (FR: 1.060; RF: 0.234)
367 Population Density (Fig. 4m) is 1160 to 1270 (FR: 1.234; RF: 0.245) respectively. The layer of

Rainfall, Household Density, Cultivator Density, and Population Density were generated in the GIS environment using the IDW tool (Fig. 4p). The correlation between flood sites and the land use pattern (Fig. 4n) indicates that water bodies and cropland have a higher chance to flood followed by the Built-up area.

The variables-wise weighting was estimated using relative frequency values where Geomorphology is the most influencing variable with a priority value of 21.40 per cent followed by Rainfall, Elevation, and LULC with priorities values of 16.84 per cent, 13.84 per cent, and 10.10 per cent. SPI (Fig. 4h; Fig. 4o), Flow Direction, and Plan Curvature are the least influencing variables to flood with 0.53 per cent, 0.65 per cent, and 0.97 per cent variable priorities values.

Finally, the Flood Susceptibility Zoning map (Fig. 5a; Table 3) was generated using equation 14 in the GIS environment. From the result, it is found that 11.02 per cent of the total area is very highly susceptible to flood whereas only 8.80 per cent area is least susceptible. 32.06 per cent of the total area is moderately susceptible to flood.

4.2. Execution Shannon Entropy Index

The relative frequency values were used to calculate the Shannon's Entropy Index to estimate the priority weights of the variables. A higher relative frequency value indicates a higher probability of the flood occurring and a lower value indicates a lesser probability to flood occurring. From the estimation of the relationship between the flood sites and variables, it is found that the rainfall zone of <1750 mm (RF: 0.374), the elevation zone of 10-30 m (RF: 0.336), the slope zone of <1.0 ° (RF: 0.211), the NDVI zone of -0.12-0.05 (RF: 0.228), the flat curvature zone (RF: 0.344), the stream density zone of 6.5-8.1 (RF: 0.303), the flow direction zone of South (RF: 0.131), the SPI zone of 0.24-4.05 (RF: 0.204), the old flood plain/water body geomorphic zone (RF: 0.442), the distance from the river zone of <1.2 km(RF: 0.270), the household density zone of 160-180 (RF: 0.280), the cultivator zone of 108-118 (RF: 0.252), the population density zone of 1160-1270 (RF: 0.252), and crop land of LULC zones have higher probabilities of flooding.

In Shannon's Entropy Index estimation, it found that Plan Curvature is the most influencing variable with 10.09 per cent criteria weight followed by Geomorphology with 9.65 per cent, LULC with 8.45 per cent of criteria weights. Flow direction is the least influencing variable

with a 3.34 per cent criteria weight. The rainfall and Elevation influence the estimation with 7.62 per cent and 7.25 per cent of criteria weights respectively.

Finally, the Flood Susceptibility Zoning map (Fig. 5b; Table 3) was generated using equation 14 in the GIS environment. From the result, it is found that 13.90 per cent of the total area is very highly susceptible to flood whereas only 7.20 per cent area is least susceptible. 31.44 per cent of the total area is moderately susceptible to flood.

4.3. Execution Weight of Evidence model

According to Armaş (2012), WofE is one of the most extensively used statistical data integration methods since it can be utilized with only a few predictive variables. According to Rahmati et al. (2015), the WoE method is an appropriate technique for flood hazard modeling, because its unpredictability is related to hazard events and their connections with the complex landscape. From the calculation, it was found that the rainfall zone of 1750-1800 mm (WoE: 0.023243), and 1800-1850 mm (WoE: 0.994778) are positively correlated with flood occurrence in the study area. In the case of elevation, the zone of 40-55 m is positively co-related with the flood with WoE 1.01696. The slope zone of <1.0 represents a WoE value of -0.00832 which indicates the negative influence of the alternative zone on the flood. The NDVI zones of 0.17-0.25 and 0.25-0.46 indicate a positive influence on the flood. The flat curvature zone is negatively co-related with the flood with a WoE of -0.01242. The stream density zone of 3.4-5.0 (WoE: 0.215628), Flow direction zones of North, North-East (WoE: 0.053785), East, and South-East, the SPI zone of -3.1- -0.38 are representing a positive relation with the flood. The geomorphic zone of the Active Alluvial Plain positively (WoE: 1.598498) influences the flood. Excluding <1.2 km (WoE: -0.22745) distance from the river zones, all other zones positively influence the flood. The household density zone of 160-180 is negative and 180-200 positively influences the flood. The cultivator density zone of 70-80, and 98-108 positively influences the flood with WoE of 0.308081, and 0.540594 respectively. The population Density zone of 740-840 (WoE: -0.21018), and 1160-1270 (WoE: -0.23727) negatively influence the flood. The LULC zone of Built-up represents a positive value of WoE 0.386191 which indicates the positive influence of the alternative on the flood.

Table 3 Areal distribution of flood susceptibility zones of the study area

The WoE values are used to estimate the Information Values (IV) of the variables as well as of all the alternatives of the variables. Using Information Values, the final flood susceptibility map was generated. The IV of rainfall is 31.42 per cent followed by geomorphology with 27.78 per cent, elevation with 12.93 per cent whereas SPI represents a lower IV of 0.01 per cent followed by slope with 0.03 per cent respectively.

Finally, the Flood Susceptibility Zoning map (Fig. 5c; Table 3) was generated using equation 14 in the GIS environment. From the result, it is found that 11.50 per cent of the total area is very highly susceptible to flood whereas only 8.08 per cent area is least susceptible. 31.46 per cent of the total area is moderately susceptible to flood.

Fig. 4 Performing Variables of the research; (a) Rainfall; (b)Elevation; (c) Slope; (d)NDVI; (e)Plan Curvature; (f) Stream Density; (g) Flow Direction; (h) SPI; (i) Geomorphology; (j) Dist. from the river; (k) Household Density; (l) Cultivator Density; (m) Population Density; (n) LULC; (o) Stream Network; (p) Spots of the IDW.

Table 2 Areal distribution of the variables with their relative weight of applied statistical methods

4.4. Comparison of FR, SEI, and WoE-IV results

The final results of all the models that were applied in the evaluation nearby represent a valid result. From the result changing matrix of FR and SEI, it is found that in the case of SEI estimation the very high susceptibility zone is larger than 2.88 per cent from FR estimation. The very low, Low, and Moderate susceptibility regions decreased in SEI estimation by -1.60 per cent, -3.32 per cent, and -0.62 per cent respectively. The result of WoE-IV is also valid with FR estimation, where the higher flood susceptibility region is just 0.48 per cent more than the FR result. The result of SEI and WOE-IV are also comparing using changing matrix (Table 4; Fig. 7). The very high flood susceptibility region of the WoE-IV estimation is - 2.40 per cent lower than SEI estimation (Table 5).

Table 4 Changing comparison matrix of different alternatives of the FSZ (Flood Susceptibility Zone) for applied techniques

Table 5 Areal changes of the result of using the methods

Fig. 5 Potential Flood Susceptibility Outputs; (a) Frequency Ratio; (b) Shannon's Entropy Index; (c) WoE-IV

4.5. District-wise Flood Susceptibility Analysis

From the estimation of district-wise flood susceptibility regions, it is found that more than 33.39 per cent areas of the entire blocks are vulnerable to flood (Fig. 6). The middle, western and south-western portions of the Raiganj block are highly susceptible to flood. The southern and western portion of the Hemtabad block has very high susceptibility to flooding during the opposite time. The Kaliyaganj block is the least flood-affected region whereas only 8.16 per cent of the total area falls under the very high flood susceptibility zone. Most probably, itahar is the most vulnerable block to flood where 52.34 per cent of the total area falls under the very high flood susceptibility region. The southern and middle itahar is highly susceptible to flood during the rainy season.

Fig. 6 Block-wise flood susceptibility mapping; (a)Raiganj; (b)Hemtabad; (c)Kaliyaganj; (d) Itahar

Fig.7 Efficiency Comparison of applied methods

Fig. 8 Weighted Linear Combination model for potentiality estimation

5. Validation of the flood susceptibility maps

The principal objective is to find areas that may be affected by future flooding in the analysis of susceptibility to floods. So it is very crucial to check the resulting flood susceptibility maps in respect of any further unknown floods regardless of whether integration process is employed (Chung & Fabbri 2003). The examination of the accuracy of the flood susceptibility maps obtained using FR, SEI, and WoE-IV models was conducted using the Receiver Operating Characteristics (ROC) analysis (Rahmati et al., 2014; Sarkar and Mondal, 2020). A receiver operating characteristic curve (ROC curve) is a graph that depicts a classification model's output across all classification thresholds. The ROC (Receiver operating features) curve, derived from the fields of signal detection, visual representation of the hit, and false-alarm rate exchange. TPR vs. FPR at different prediction thresholds is plotted on a ROC curve. As the rating threshold is reduced, more objects are labeled as positive, resulting in a rise in both False Positives and True Positives. It is obvious that the

AUC (Area Under Curve) is around 0.911 in flood susceptibility maps, which corresponds with the reliability of the forecast of 91.10 per cent by methods of the FR model (Fig. 9) whereas it is 86.70 per cent and 90.70 per cent for SEI and WoE-IV estimation. Therefore, the FR, SEI, and WoE-IV demonstrated almost the same and fair results based on the computed AUC.

Fig. 9 Validation of the results using ROC curve

6. Conclusion

Floods have been labeled the world's worst and most devastating catastrophe in recent years, amid several geo-environmental risks and catastrophes. Flood susceptibility mapping can be used to detect flood-prone locations or regions. Flood susceptibility mapping is essential for flood-prone areas to prepare for proper field management. The major goal of creating a flood susceptibility map was to promote community awareness of the flood threat among residents, local government, and other organizations. The current research aims to define flood-prone locations in the Raiganj Sub-division by developing flood susceptibility maps. The geospatial technology and BSA-based statistical methodologies were used to design and assess the results of this study. This semi-quantitative and BSA-based statistical technique is useful for determining appropriate flood management strategies. Based on 70 training points, The FR approach was used to investigate the relationship between past flood events and the likelihood of forecasted future flood episodes. From the final outputs, it is found that about 11 to 14 per cent of the total area belongs to the high flood susceptibility region. The results of the FR Shannon's Entropy value and WoE-IV estimations demonstrate that variables like rainfall, elevation, LULC, Geomorphology, distance from the river, and drainage density play a crucial role in resulting in floods in the study region. The success rate of the frequency ratio model is 0.911, while the success rate of Shannon's entropy is 0.867. The success rate of the WoE-IV model is 0.907 indicating that the used models for analyzing flood susceptibility of the Raiganj sub-division might be considered genuine. Itahar and Raiganj blocks of the district are demarcated as the area having very high flood risk potentiality. The main barrier to flood susceptibility mapping is the scarcity of trustworthy flood occurrence data sources, information, as well as a suitable modeling approach for replicating the findings. This study demonstrates that bivariate approaches such as FR, SEI, and WoE-IV, when

combined with a Geographic Information System, are always effective measures (GIS) to flood susceptibility mapping. A flood management strategy includes conducting a risk assessment and developing a flood susceptibility map. Developing and implementing an effective flood vulnerability map can help in identifying the flood hazard zones, putting adequate management systems in place and early warning systems, preparation for a swift response during a flood, flood shelters, and flood rescue planning are all instances of resilience methods.

Acknowledgments

The authors are grateful to acknowledge all the agencies especially, the Indian Meteorological Department (IMD), Geological Survey of India, Survey of India (SOI), USGS, and Alaska Sattelite Facility as sources of data required for the study. We'd like to thank Dr. Gopal Chandra Debnath (Retired Professor of Visva Bharati University, W/B) for his assistance with the data analysis and model validation parts.

Compliance with ethical standards

The manuscript of the research was prepared as per the Journal's ethical standards and all the authors were equally contributed to the manuscript preparation. The authors hereby declared that there is no conflict of interest. During the research work, no humans or animals are wounded or harmed in any way.

Reference

Adiat KAN, Nawawi MNM, & Abdullah K (2012) Integration of geographic information system and 2D imaging to investigate the effects of subsurface conditions on flood occurrence. *Mod App Sci 6*(3):11-21

Al-Hinai H, Abdalla R (2021) Mapping Coastal Flood Susceptible Areas Using Shannon's Entropy Model: The Case of Muscat Governorate, Oman ISPRS Int J Geoinf 10(4):252

Al-Juaidi AE, Nassar AM, & Al-Juaidi OE (2018) Evaluation of flood susceptibility mapping using logistic regression and GIS conditioning factors. *Ara J Geo 11*(24): 1-10

542 Althuwaynee OF, Pradhan B, Park HJ, & Lee JH, (2014) A novel ensemble bivariate statistical
543 evidential belief function with knowledge-based analytical hierarchy process and
544 multivariate statistical logistic regression for landslide susceptibility mapping. *Cat 114* :21-36

545 Aniya M, (1985) Landslide-susceptibility mapping in the Amahata river basin, Japan. *Ann Ass*
546 *Amer Geo 75*(1):102-114

547 Azareh A, Rafiei SE, Choubin B, Barkhori S, Shahdadi A, Adamowski J, & Shamshirband S,
548 (2019) Incorporating multi-criteria decision-making and fuzzy-value functions for flood
549 susceptibility assessment. *Geoca Inter* 1-21

550 Bonham-Carter GF, (1994) Geographic information systems for geoscientists-modeling with
551 GIS. *Comp meth geosci 13*:398

552 Central Water Commission (CWC), (2010) Water and related statistics Water Resource
553 Information System Directorate. New Delhi 198–247

554 Chapi K, Singh VP, Shirzadi A, Shahabi H, Bui DT, Pham BT, & Khosravi K, (2017) A novel
555 hybrid artificial intelligence approach for flood susceptibility assessment. *Envir mod soft. 95*:
556 229-245

557 Chung CF, Fabbri AG, (2003) Validation of spatial prediction models for landslide hazard
558 mapping. Nat Haz 30:451–472

559 Cloke HL, & Pappenberger F, (2009) Ensemble flood forecasting: A review. *J hydr 375*(3-4):
560 613-626

561 Costache R, Pham QB, Sharifi E, Linh NTT, Abba SI, Vojtek M, & Khoi DN (2020) Flash-flood
562 susceptibility assessment using multi-criteria decision making and machine learning
563 supported by remote sensing and gis techniques. *Rem Sens 12*(1):106

564 Dano UL, Balogun AL, Matori AN, Wan YK, Abubakar IR, Said Mohamed MA, & Pradhan B
565 (2019) Flood susceptibility mapping using GIS-based analytic network process: A case study
566 of Perlis, Malaysia. *Wat 11*(3): 615

567 Dhar ON, Mandal BN, and Ghose GC (1981a) *Vamsadhara flash flood of September 1980 - a*
568 *brief appraisal.* Va Man 11: 7-11

569 Dhar ON, Nandargi S (2003) Hydrometeorological aspects of floods in India. Nat Haz 28(1):1-
570 33

571 Dhar ON, Rakhecha PR, Mandal BN, Sangam RB (1981b) The rainstorm which caused the
572 Morvi dam disaster in August 1979/L'orage qui a provoqué la catastrophe du barrage Morvi
573 août 1979 *Hydr* Sci J 26(1):71-81

574 Fernández DS, Lutz MA (2010) Urban flood hazard zoning in Tucumán Province, Argentina,
575 using GIS and multicriteria decision analysis. *Eng Geo* 111:90–98

576 Glenn E, Morino K, Nagler P, Murray R, Pearlstein S, Hultine K (2012) Roles of saltcedar
577 (Tamarix spp) and capillary rise in salinizing a non-flooding terrace on a flow-regulated
578 desert river. *J Ari Envi* 79:56–65

579 Gül GO, (2013) Estimating flood exposure potentials in Turkish catchments through index-
580 based flood mapping Nat Haz. 69:403–423

581 Gupta S, Javed A, Dutt D (2003) Economics of flood protection in India. Nat Haz 28:199–210

582 Haghizadeh A, Siahkamari S, Haghiabi AH, & Rahmati O (2017) Forecasting flood-prone areas
583 using Shannon's entropy model. *J Ear Sys Sci 126*(3): 39

584 Hong H, Panahi M, Shirzadi A, Ma T, Liu J, Zhu AX, & Kazakis N (2018) Flood susceptibility
585 assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with genetic
586 algorithm and differential evolution. *Sci tot Envi 621*:1124-1141

587 Jebur MN, Pradhan B, Tehrany MS, (2014) Using ALOS PALSAR derived high-resolution
588 DInSAR to detect slow-moving landslides in tropical forest, Cameron Highlands, Malaysia. J
589 *Geo Nat Haz Ri* 1–19. https:// doi:101080/194757052013860407

590 Kafira V, Albanakis K, & Oikonomidis D (2014) Flood Susceptibility Assessment using GIS An
591 example from Kassandra Peninsula, Halkidiki, Greece. *Proc 10th Inte Congress Hel Geo Soci*
592 *Thessaloniki, Greece* 287-308

593 Kalsi SR, & Srivastava KB (2006) Characteristic features of Orissa super cyclone of 29th
594 October, 1999 as observed through CDR Paradip. *Maus 57*(1):21

595 Kanani-Sadat Y, Arabsheibani R, Karimipour F, & Nasseri M (2019) A new approach to flood
596 susceptibility assessment in data-scarce and ungauged regions based on GIS-based hybrid
597 multi criteria decision-making method. *J Hydr 572*:17-31

598 Keshtegar B, Hasanipanah M, Bakhshayeshi I, Sarafraz ME, (2019) A novel nonlinear
599 modeling for the prediction of blast-induced airblast using a modified conjugate FR
600 method. *Mea* 131:35-41

601 Khoirunisa N, Ku CY, & Liu CY (2021) A GIS-Based Artificial Neural Network Model for Flood
602 Susceptibility Assessment. *Inter J Envir Res Pub Hea 18*(3):1072

603 Khosravi K, Nohani E, Maroufinia E, & Pourghasemi HR (2016a) A GIS-based flood
604 susceptibility assessment and its mapping in Iran: a comparison between frequency ratio
605 and weights-of-evidence bivariate statistical models with multi-criteria decision-making
606 technique. *Nat Haz 83*(2):947-987

607 Khosravi K, Nohani E, Maroufinia E, Pourghasemi HR (2016b) A GIS-based flood susceptibility
608 assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-
609 evidence bivariate statistical models with multi-criteria decision-making technique. Nat
610 Haz 83(2):947-987

611 Khosravi K, Pourghasemi HR, Chapi K, & Bahri M (2016c) Flash flood susceptibility analysis
612 and its mapping using different bivariate models in Iran: a comparison between Shannon's
613 entropy, statistical index, and weighting factor models. *Envi moni asse 188*(12):1-21

614 Kia MB, Pirasteh S, Pradhan B, Mahmud AR, Sulaiman WNA, & Moradi A (2012) An artificial
615 neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Envi Ear*
616 *Sci 67*(1):251-264

617 Kourgialas NN, Karatzas GP (2011) Flood management and a GIS modelling method to assess
618 flood-hazard areas: a case study. Hydr Sci J 56:212–225

619 Lappas I, & Kallioras A (2019) Flood susceptibility assessment through GIS-based multi-
620 criteria approach and analytical hierarchy process (AHP) in a river basin in Central
621 Greece. *para (Malczewski, 1999)*:*6*(03)

622 Lee MJ, Kang JE, & Jeon S (2012) Application of frequency ratio model and validation for
623 predictive flooded area susceptibility mapping using GIS. *Int geosci rem sens sym* 895-898

624 Lee S, and Pradhan B (2007) Landslide hazard mapping at Selangor, Malaysia using
625 frequency ratio and logistic regression models Landslides. 4(1). 3341doi:101007/s10346-
626 006-0047-y

627 Lee S, Lee S, Lee MJ, & Jung HS (2018) Spatial assessment of urban flood susceptibility using
628 data mining and geographic information System (GIS) tools. *Sust 10*(3):648

629 Manandhar B (2010) Flood Plain Analysis and Risk Assessment of Lothar Khola, MSc Thesis,
630 Tribhuvan University, Phokara, Nepal, 64

631 Manap AM, Nampak H, Pradhan B, Lee S, Sulaiman WNA, & Ramli MF (2014) Application of
632 probabilistic-based frequency ratio model in groundwater potential mapping using remote
633 sensing data and GIS. *Ara J Geosci 7*(2):711-724

634 Moghaddam DD, Rezaei M, Pourghasemi HR, Pourtaghie ZS, Pradhan B (2015) Groundwater
635 spring potential mapping using bivariate statistical model and GIS in the Taleghan
636 watershed, Iran. *Ara J Geosci* 8(2):913-929

637 Mondal S, Maiti R, (2013) Integrating the analytical hierarchy process (AHP) and the
638 frequency ratio (FR) model in landslide susceptibility mapping of Shiv-khola watershed,
639 Darjeeling Himalaya. *Inter J* Dis *Ris Sci* 4(4):200-212

640 Oh HJ, Pradhan B, (2011) Application of a neuro-fuzzy model to landslide susceptibility
641 mapping for shallow landslides in a tropical hilly area. *Comp Geosci* 37:1264–1276

642 Ouma YO, & Tateishi R (2014) Urban flood vulnerability and risk mapping using integrated
643 multi-parametric AHP and GIS: methodological overview and case study
644 assessment. *Wat 6*(6):1515-1545

645 Ozdemir A, Altural T (2013) A comparative study of frequency ratio, weights of evidence and
646 logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW
647 Turkey. J Asian Earth Sci64:180-197

648 Pham BT, Avand M, Janizadeh S, Phong TV, Al-Ansari N, Ho, LS, PrakashI (2020) GIS based
649 hybrid computational approaches for flash flood susceptibility assessment. Water 12(3):683

650 Pourghasemi HR, Mohammadi M, Pradhan B (2012) Landslide susceptibility mapping using
651 index of entropy and conditional probability models at Safarood Basin, Iran. Catena 97:71-
652 84

653 Pourtaghi, ZS, Pourghasemi, HR, (2014) GIS-based groundwater spring potential assessment
654 and mapping in the Birjand Township, southern Khorasan Province, Iran.Hydrogeol Jdoi:
655 101007/s10040-013-1089-6

656 Pradhan B (2010) Flood susceptible mapping and risk area estimation using logistic
657 regression, GIS and remote sensing.J Spat Hydrol 9(2):1–18

658 Rahmati O, Nazari SA, Mahdavi M, Pourghasemi HR, Zeinivand H (2014) Groundwater
659 potential mapping at Kurdistan region of Iran using analytic hierarchy process and
660 GIS.Arab J Geoscidoi: 101007/s12517-014-1668-4

661 Rahmati O, Samani AN, Mahdavi M, PourghasemiHR, Zeinivand H (2015) Groundwater
662 potential mapping at Kurdistan region of Iran using analytic hierarchy process and
663 GIS. Arab J Geosci8(9):7059-7071

664 RahmatiO, Pourghasemi HR, Zeinivand H (2016) Flood susceptibility mapping using
665 frequency ratio and weights-of-evidence models in the Golastan Province, Iran. Geocarto
666 Int31(1):42-70

667 Sabatakakis N, Koukis G, Vassiliades E, Lainas S (2013) Landslide susceptibility zonation in
668 Greece. Nat Hazards65(1):523-543

669 Sachdeva S, Bhatia T, Verma AK (2017) Flood susceptibility mapping using GIS-based support
670 vector machine and particle swarm optimization: A case study in Uttarakhand (India)
671 In 2017 8th International conference on computing, communication and networking
672 technologies (ICCCNT) (pp 1-7) IEEE

673 Saha A, Pal SC, Arabameri A, Blaschke T, Panahi S, ChowdhuriI,Arora A (2021) Flood
674 susceptibility assessment using novel ensemble of hyperpipes and support vector regression
675 algorithms. Water, 13(2):241

Saha S, Mondal P (2020) A Catastrophic Flooding Event in North Bengal, 2017 and its Impact Assessment: A Case Study of Raiganj CD Block Uttar Dinajpur, West Bengal. AppliGeospat Tech Geomorpho Environ IGI Conf ISBN 978-81-925799-3-1

Saha S, Sarkar D, MondalP, Goswami S (2021) GIS and multi-criteria decision-making assessment of sites suitability for agriculture in an anabranching site of sooin river, India. Model Earth Syst Environ7(1):571-588

Sahana M, Patel PP (2019) A comparison of frequency ratio and fuzzy logic models for flood susceptibility assessment of the lower Kosi River Basin in India. Environ Earth Sci78(10):1-27

Sahana M, Rehman S, Sajjad H, Hong H, (2020) Exploring effectiveness of frequency ratio and support vector machine models in storm surge flood susceptibility assessment: A study of Sundarban Biosphere Reserve, India. Catena 189:104450

Samanta S, Pal DK, Palsamanta B (2018a) Flood susceptibility analysis through remote sensing, GIS and frequency ratio model. Appl Water Sci8(2):1-14

SamantaRK, Bhunia GS, Shit PK,Pourghasemi HR (2018b) Flood susceptibility mapping using geospatial frequency ratio technique: a case study of Subarnarekha River Basin, India. Model Earth Syst Environ 4(1):395-408

Sarkar D, Mondal P (2020) Flood vulnerability mapping using frequency ratio (FR) model: a case study on Kulik river basin, Indo-Bangladesh Barind region. Appl Water Sci 10(1):1-13

Singh O, Kumar M(2013) Flood events, fatalities and damages in India from 1978 to 2006. Nat hazards, 69(3):1815-1834

Srdevic Z, Blagojevic B, Srdevic B (2011) AHP based group decision making in ranking loan applicants for purchasing irrigation equipment: a case study Bulgarian. J Agric Sci17(4):531-543

Tang Z, Yi S, Wang C, Xiao Y (2018) Incorporating probabilistic approach into local multi-criteria decision analysis for flood susceptibility assessment. Stoch Environ Res Risk Assess 32(3):701-714

703 Tehrany M, Kumar L, NeamahJeburM, Shabani F (2019) Evaluating the application of the
704 statistical index method in flood susceptibility mapping and its comparison with frequency
705 ratio and logistic regression methods. Geomat Nat Haz Risk 10(1):79-101

706 Tehrany MS, Pradhan B, Mansor S, Ahmad N (2015) Flood susceptibility assessment using
707 GIS-based support vector machine model with different kernel types. Catena 125:91–101

708 TehranyMS, Pradhan B,Jebur MN (2014) Flood susceptibility mapping using a novel
709 ensemble weights-of-evidence and support vector machine models in GIS. J hydrol 512:332-
710 343

711 Triantaphyllou E, Mann SH (1995) Using the analytic hierarchy process for decision making
712 in engineering applications: some challenges. Int J Ind Eng: Theory Appl Pract2(1):35-44

713 Weier J,Herring D, (2000) Measuring Vegetation (NDVIEVI) NASA Earth
714 Observatory Washington, DC, USA

715 WHO (2003) World Health Organization Disaster data-key trends and statistics in world
716 disasters report WHO, Geneva, Switzerland

717 Wilson JP, Gallant JC (2000) Terrain Analysis: Principles and Applications New York. Wiley, p
718 479

719 Youssef AM, Pradhan B, Sefry SA (2016) Flash flood susceptibility assessment in Jeddah city
720 (Kingdom of Saudi Arabia) using bivariate and multivariate statistical models. Environ Earth
721 Sci 75(1):12

Figures

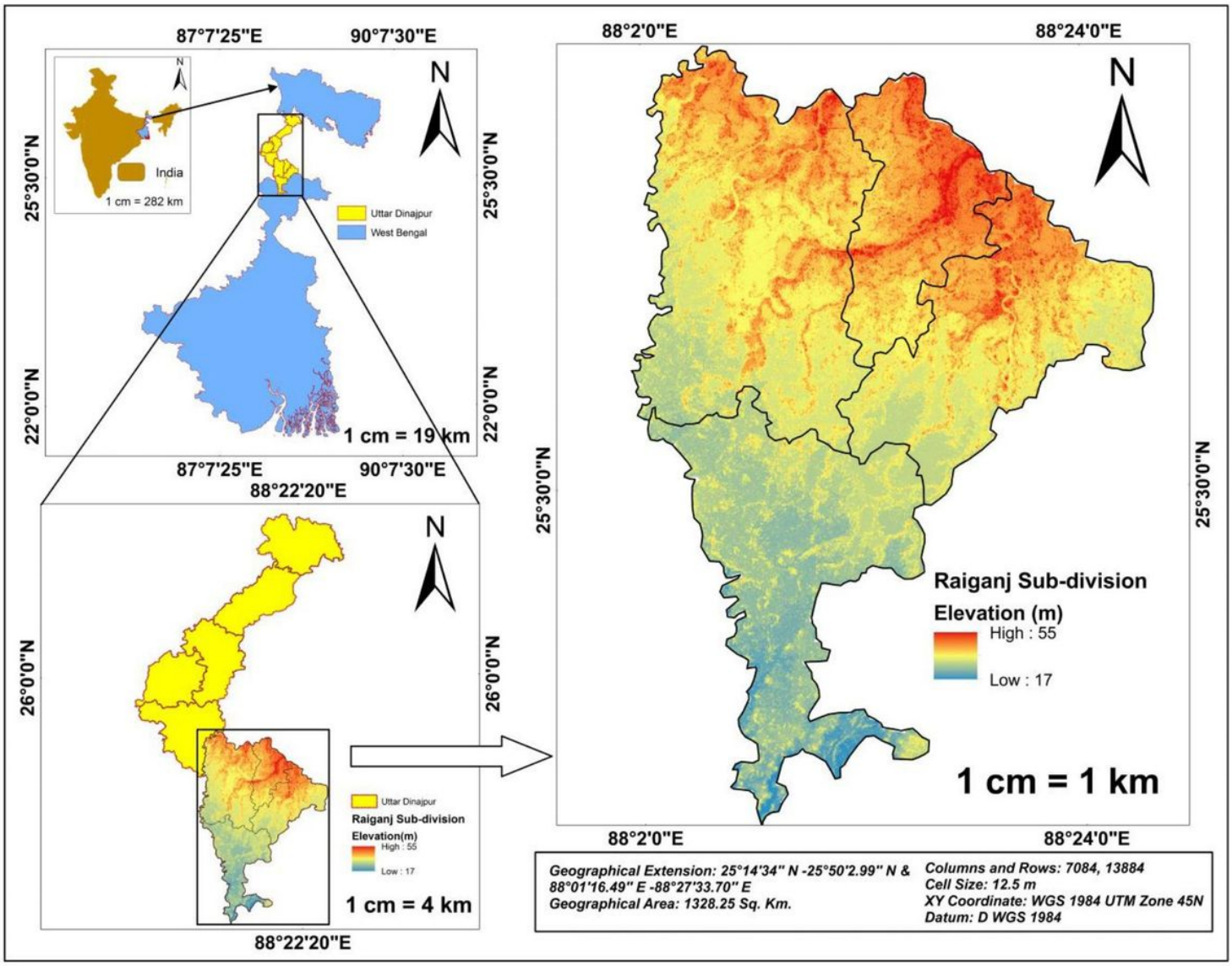


Figure 1

Geographical Location of the research area

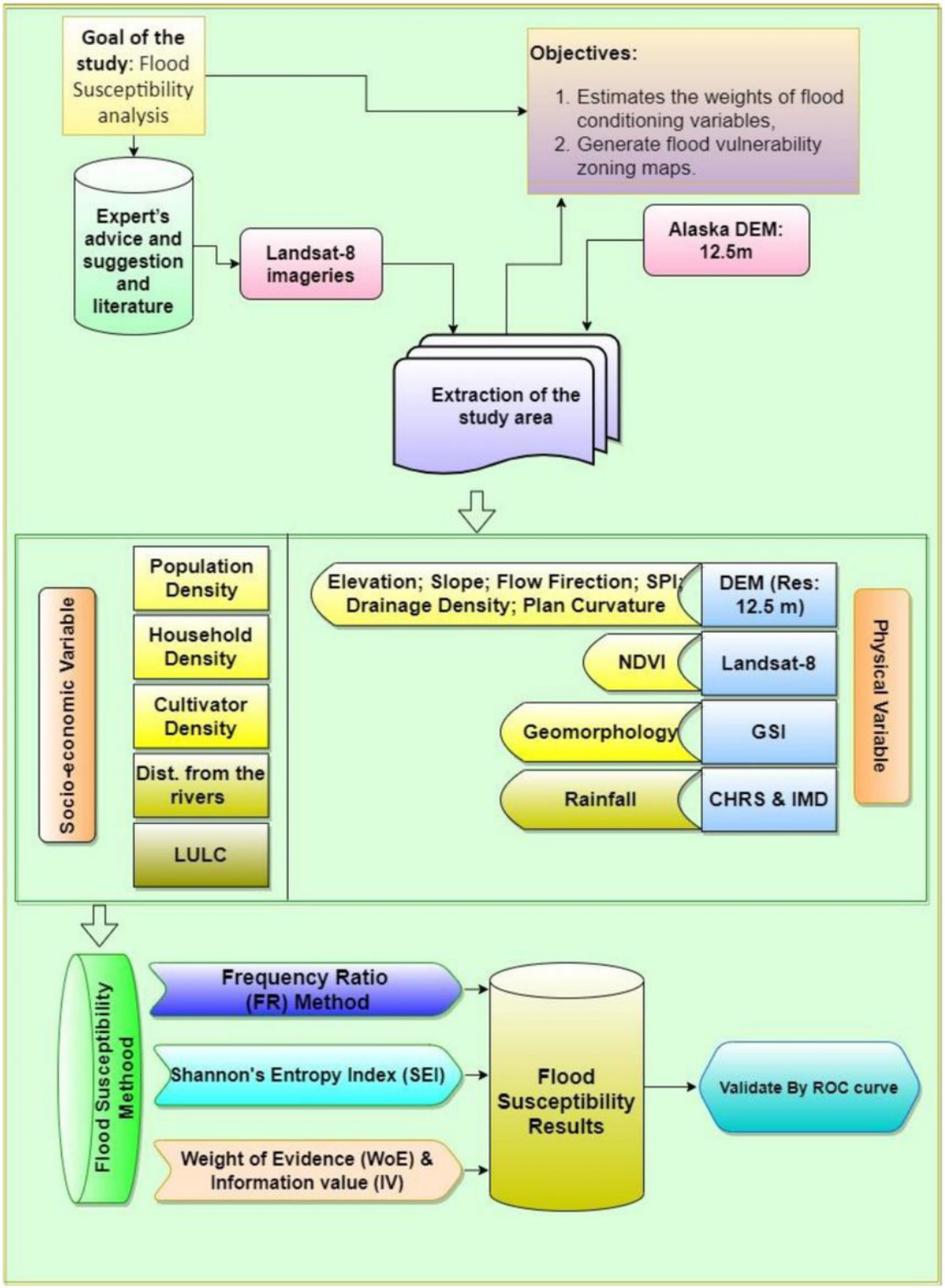


Figure 2

Flow diagram of the research

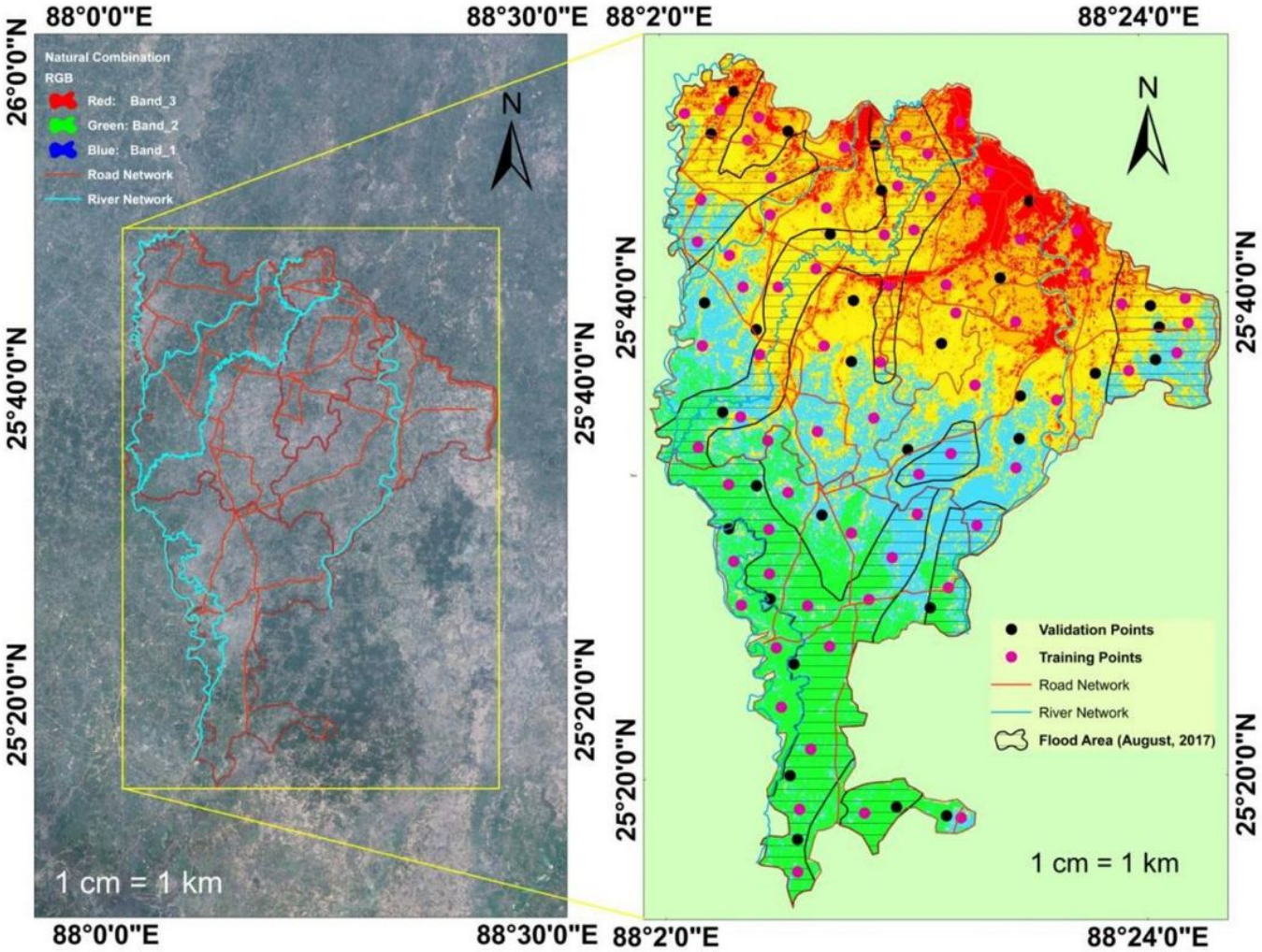


Figure 3

Flood Inventory mapping of the research area

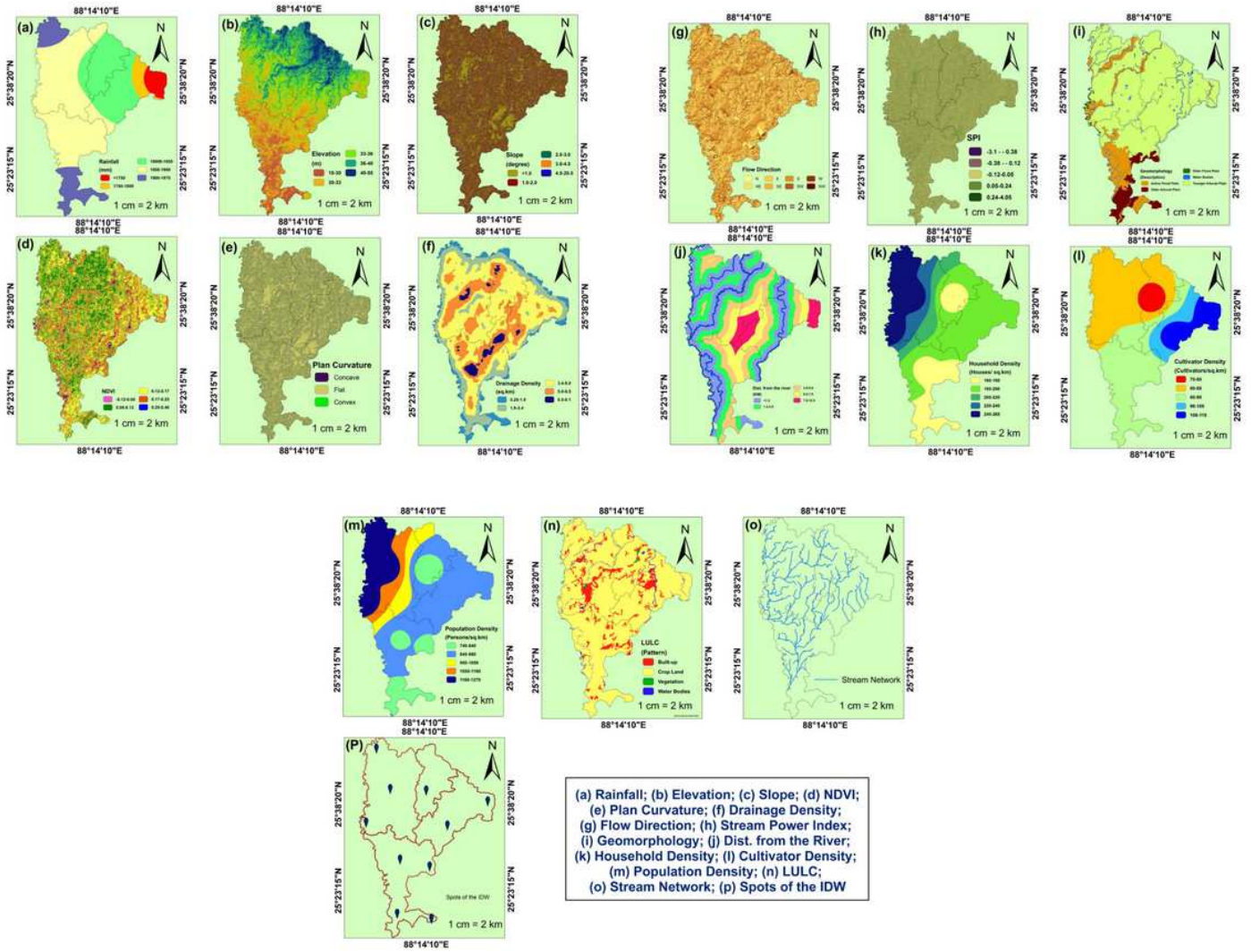


Figure 4

Performing Variables of the research; (a) Rainfall; (b)Elevation; (c) Slope; (d)NDVI; (e)Plan Curvature; (f) Stream Density; (g) Flow Direction; (h) SPI; (i) Geomorphology; (j) Dist. from the river; (k) Household Density; (l) Cultivator Density; (m) Population Density; (n) LULC; (o) Stream Network; (p) Spots of the IDW.

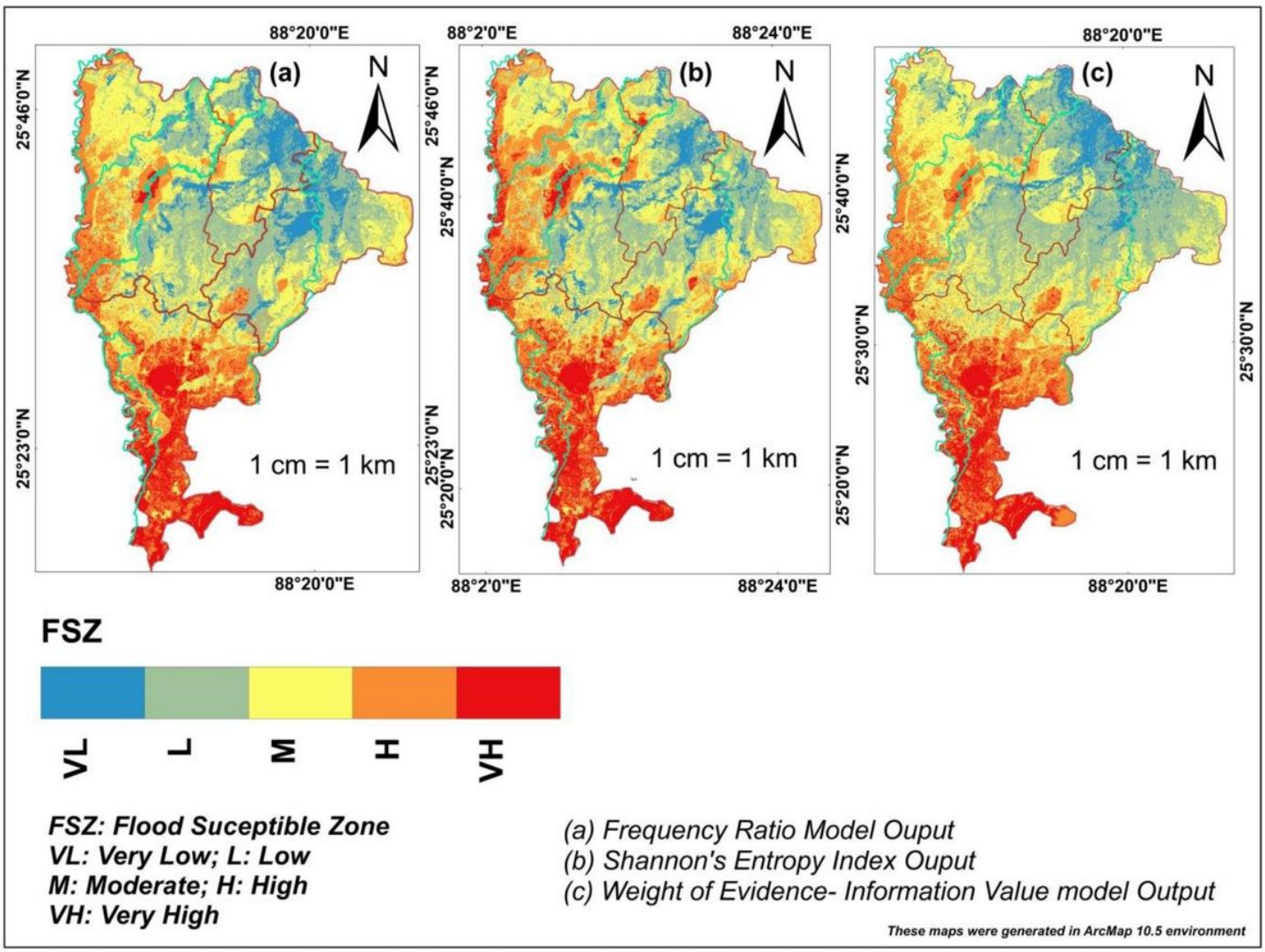


Figure 5

Potential Flood Susceptibility Outputs; (a) Frequency Ratio; (b) Shannon's Entropy Index; (c) WoE-IV

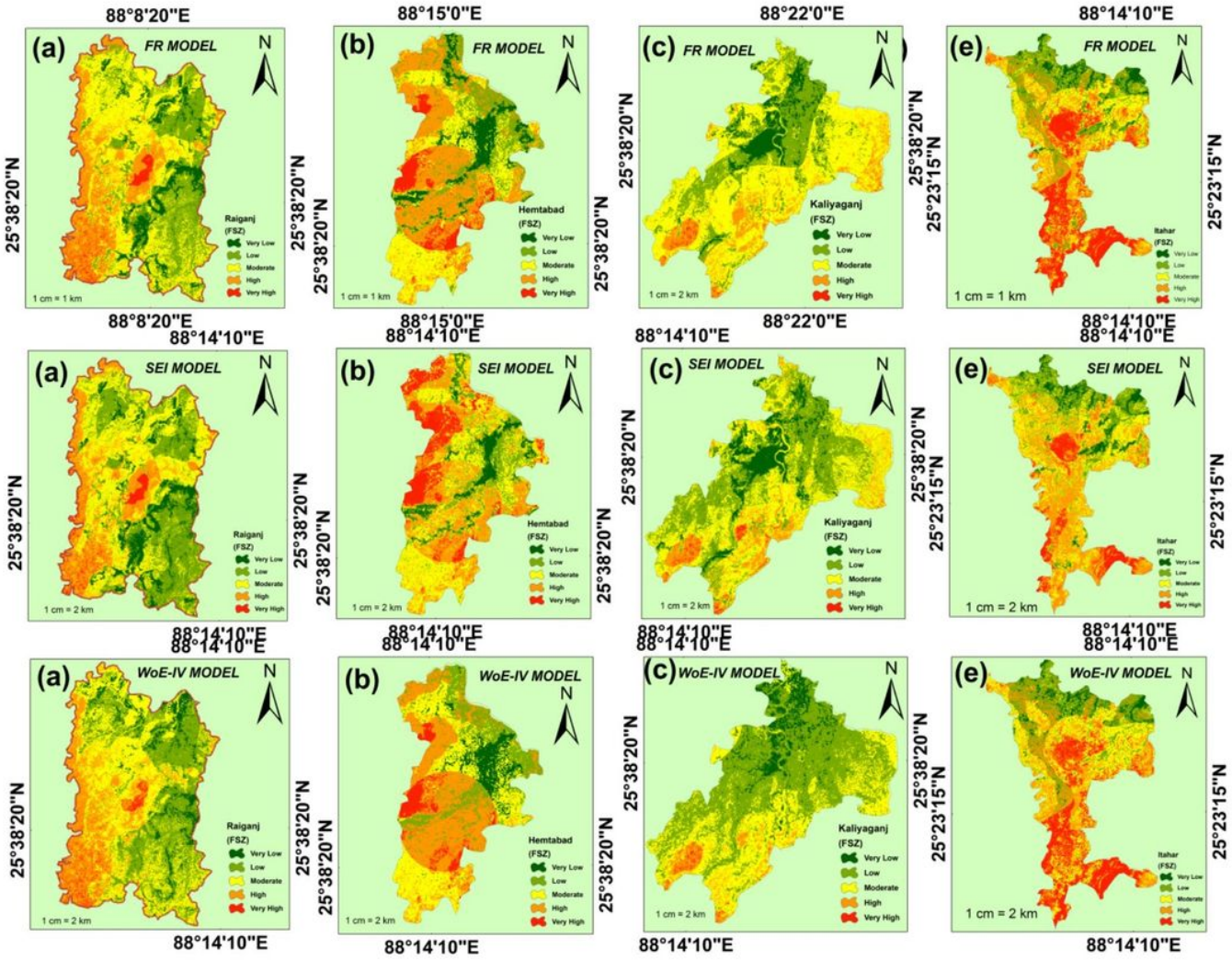
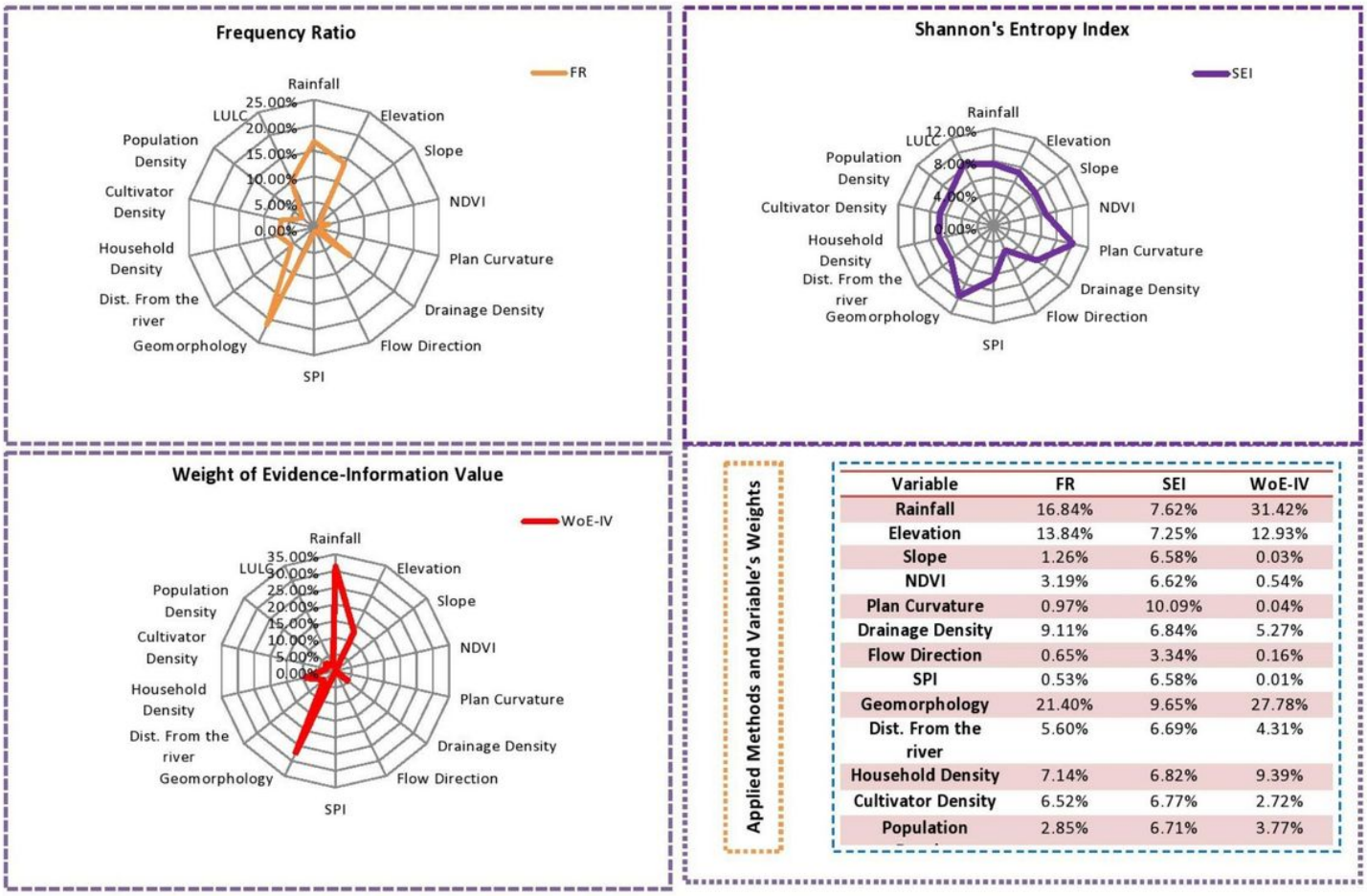


Figure 6

Block-wise flood susceptibility mapping; (a)Raiganj; (b)Hemtabad; (c)Kaliyaganj; (d) Itahar



Variable	FR	SEI	WoE-IV
Rainfall	16.84%	7.62%	31.42%
Elevation	13.84%	7.25%	12.93%
Slope	1.26%	6.58%	0.03%
NDVI	3.19%	6.62%	0.54%
Plan Curvature	0.97%	10.09%	0.04%
Drainage Density	9.11%	6.84%	5.27%
Flow Direction	0.65%	3.34%	0.16%
SPI	0.53%	6.58%	0.01%
Geomorphology	21.40%	9.65%	27.78%
Dist. From the river	5.60%	6.69%	4.31%
Household Density	7.14%	6.82%	9.39%
Cultivator Density	6.52%	6.77%	2.72%
Population	2.85%	6.71%	3.77%

Figure 7

Efficiency Comparison of applied methods

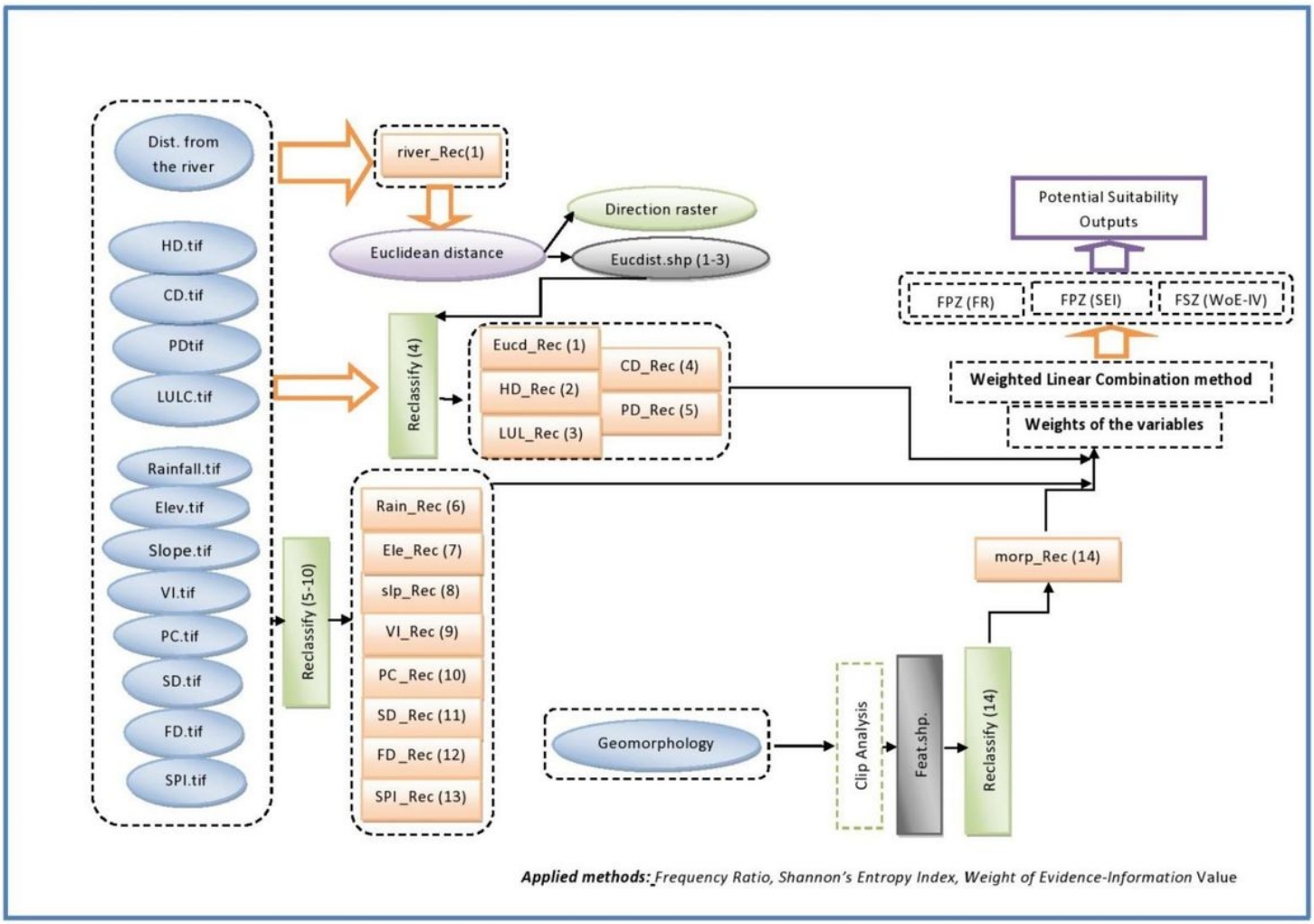


Figure 8

Weighted Linear Combination model for potentiality estimation

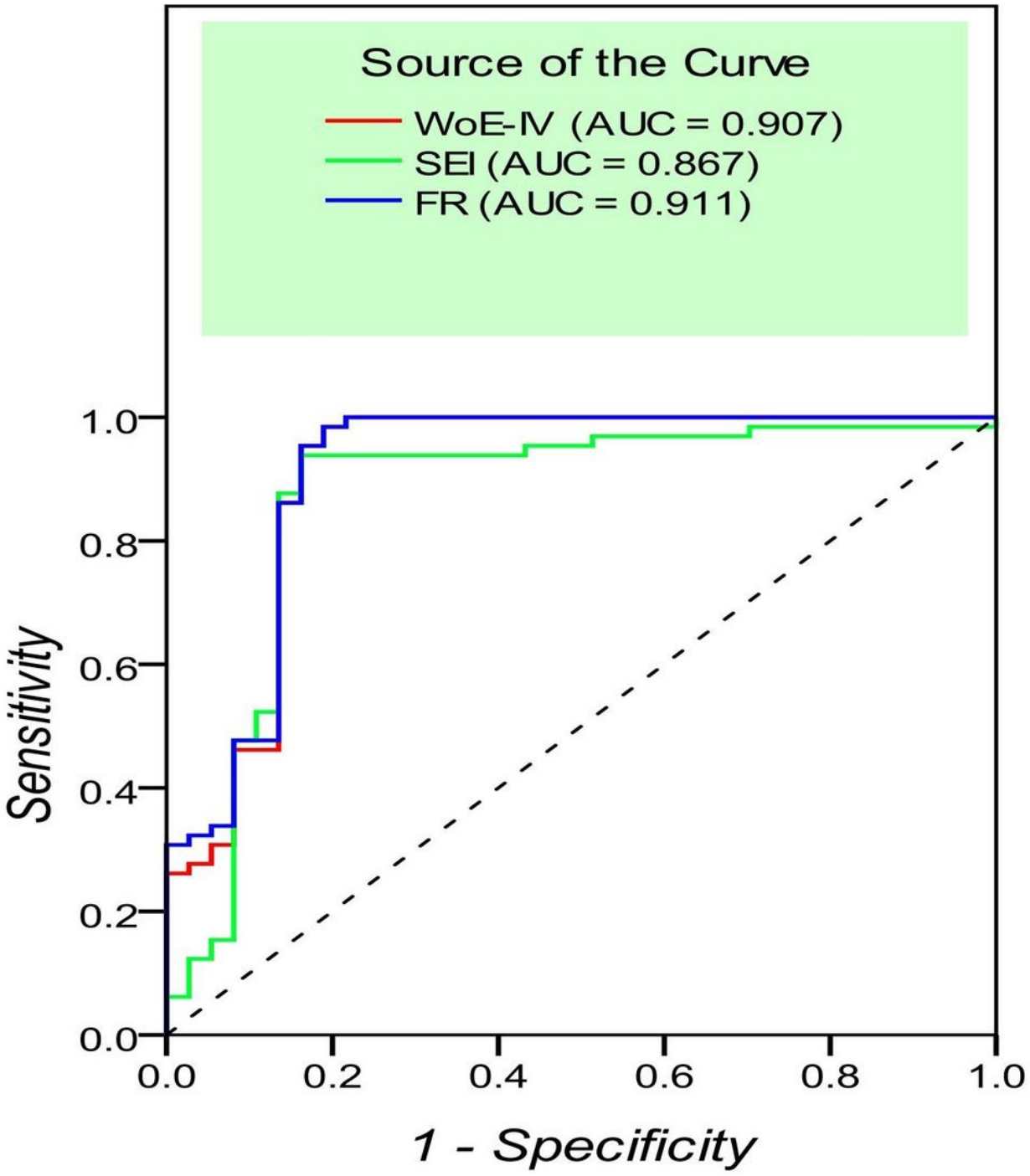


Figure 9

Validation of the results using ROC curve