Nature Environment & Pollution Technology

This is a peer-reviewed prepublished version of the paper

Estimation of Flood Hazard Zones of Noa River Basin Using Maximum Entropy Model in GIS

Nilotpal Kalita[1†], Niranjan Bhattacharjee[2], Nirmali Sarmah[2] and Manash Jyoti Nath[3]
[1]Department of Geography, Nowgong Girls' College, Assam, India
[2]Department of Geography, Pandu College, Guwahati, Assam, India
[3]Department of Geography, Gauhati University, Guwahati, Assam, India
†Corresponding author: Nilotpal Kalita; nilotpalkalita4@gmail.com
Nilotpal Kalita - Orcid id 0000-0001-7744-349X
Niranjan Bhattacharjee- 0000-0003-4205-0682
Nirmali Sarmah -Orcid id 0009-0005-0598-201X
Manash Jyoti Nath-Orcid id 0000-0002-7532-5075

Key Words	Maximum entropy, Flood hazard, MaxEnt, Environmental parameters
DOI	https://doi.org/10.46488/NEPT.2024.v24i01.B4216 (DOI will be active only after the final publication of the paper)
Citation of the Paper	Nilotpal Kalita, Niranjan Bhattacharjee, Nirmali Sarmah and Manash Jyoti Nath, 2025. Estimation of flood hazard zones of Noa river basin using maximum entropy model in GIS. *Nature Environment and Pollution Technology*, Vol. 24, No. 1, B4216.

ABSTRACT

This study aims to develop a comprehensive flood hazard map for effective hazard management in the Noa river basin, located in Assam, India, through the integration of Geographic Information System (GIS) tools and a Maximum Entropy (MaxEnt) model. The MaxEnt machine learning algorithm was employed, utilizing eight selected geographic and environmental parameters as predictors to generate the flood hazard map. The accuracy of the generated map was evaluated using the Area Under the Curve (AUC) metric. Key findings of the study identified elevation and slope as critical parameters in the assessment of flood risk. Results revealed that the flood hazard map produced by the MaxEnt model achieved an AUC value of 0.85, indicating high predictive accuracy. The research underscores the significance of flood hazard maps as essential tools for policymakers, enabling the identification of areas vulnerable to severe environmental and economic damage. By providing a reliable and precise assessment of flood-prone zones, this study contributes valuable insights for the formulation of effective flood management strategies and mitigation measures. The implementation of such hazard maps is crucial in enhancing preparedness and resilience against flooding events, ultimately safeguarding lives, property, and infrastructure in the Noa River basin.

INTRODUCTION

Rivers are dynamic entities having hydrologic, geomorphic, ecologic, environmental and economic significance. It causes hazards and sometimes disasters, particularly in riverine areas with high to medium rainfall and dense human habitation. However, one of the most significant developments of rivers is the flooding. Floods in some riverine areas, primarily in moist climatic regions, have been rapidly increasing, impacting various geomorphological forms and patterns. The Mangaldai Sub-division, a part of a riverine built-up plain composed of fine alluvial sediments and washed by streams like Bega-*nadi*, Nanoi, Bar-*nadi*, Mangaldai-*nadi,* and Noa-*nadi,* etc., has of late, witnessed increasing floods. The *nadi* is the local term for a river. The river Noa is causing sheet flooding, river bank erosion, and channel shifting. The alarming situation with rivers and floods as geomorphological agents in the sub-division has caused serious damage to geomorphic features, land, the habitat of floodplain dwellers, standing crops, and the environment.

Floods are a natural hazard as well as disaster that has affected human civilizations since time immemorial. They are often caused by heavy rainfall, melting of snow, storm surges, and the collapse of dams and embankments. Understanding floods is crucial for creating strategies to reduce their danger and harm (Garg & Babu 2023). Flood hazard is often referred to as the potential risk and adverse effects associated with the loss of life and property. The effects of flood hazards can not be controlled fully, but they can be reduced or mitigated by proper planning strategies (Harshasimha & Bhatt 2023).

Anthropogenic activities such as deforestation, unsustainable agricultural practices, and construction projects can lead to the drastic modification of ecosystems (Cerdà 2007, Kavian 2017, Javidan 2021). Physical phenomena like gully erosion, landslides, and floods occur over geological timescales and exhibit variability in both time and space (Achour & Pourghasemi 2020). These are categorized as hazard events with the potential to be influenced by human activities or occur naturally (Kelarestaghi & Ahmadi 2009).

MATERIALS AND METHODS

The study employs eight spatial variables to estimate flood hazard zones in the Noa-*Nadi* river basin. Table 1 provides a list of the data sets used and their sources. All the factors were processed into a raster grid of 30-arc-second grid cells. These raster layers were then converted into the Universal Transverse Mercator spheroid with datum WGS1984. Additional related maps were generated from SRTM DEM using appropriate spatial analysis tools in ArcGIS (Fig. 1). The spatial variables used are primarily continuous, and some of them were classified into different categories based on expert knowledge and a literature review.

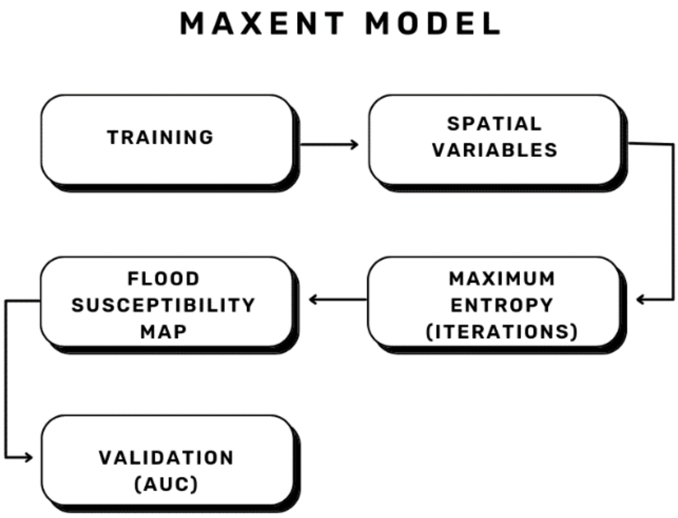


Fig. 1: Flowchart showing MAXENT modeling.

Table 1: List of the data sets used and their sources.

Spatial variables	Source
Elevation	worldclim.org
Slope	SRTM (Shuttle Radar Topography Mission)
Topographic Werness Index (TWI)	Derived from SRTM DEM
Soil Texture	FAO soil portal
Distance from River Channel	https://www.diva-gis.org/gdata
Distance from road	https://www.diva-gis.org/gdata
Precipitation	https://power.larc.nasa.gov/data-access-viewer/
Population density	NASA SEDAC (Socioeconomic Data and Applications Center) databases

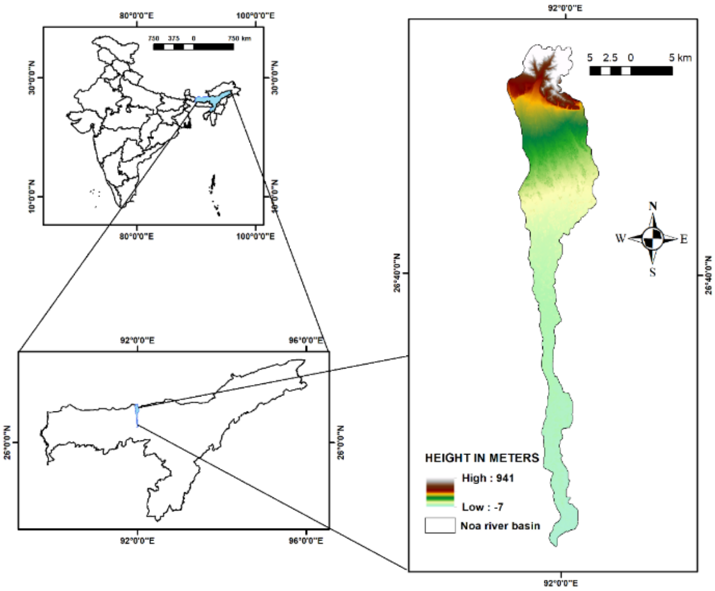


Fig. 2: Study area.

Study Area

The Noa-*nadi* (Fig. 1), extending from 91°55’30” E to 92°2’30” E and 26°55’10” N to 26°21’ 55” N, is a north bank tributary of the mighty Brahmaputra River in Assam, India. The basin covers an area of 241.85 sq.km, with the river stretching for about 104.275 kilometers in length. The Noa River is located in the Udalguri district (to the north) and Darrang district (to the south) in the Brahmaputra valley, Assam. It starts at 520 meters above sea level near Bhutiachang on the southern slope of the Bhutan Himalayas and has three tributaries: Bhola, Lakshmi, and Batiamari. The middle part of the river is called Kuyapani, and in the Rik Veda, it is referred to as the Anjashi River (Noa). The river separates from Khampajuli in Bhutan and bends eastward after passing Bhutiachang in Silputa Mouza. It flows southward, crossing National Highway 52, and joins the Mangaldai River in Darrang district. The Noa River, like other rivers, cannot form meanders if the banks are too hard or too unstable. The elevation drops 350 meters from the Assam-Bhutan border to its mouth, giving an average gradient of 5.69 meters per kilometer. The gradient varies across different parts of the river: 10.86 meters/km in the upper part, 4.41 meters/km in the middle, and 0.76 meters/km in the lower part. The river has a mean annual maximum discharge of 73.13 cubic meters per second, a mean maximum water level of 58.33 meters, and carries 166 tons of sediment per square kilometer each year (Bhattacharjee 2008).

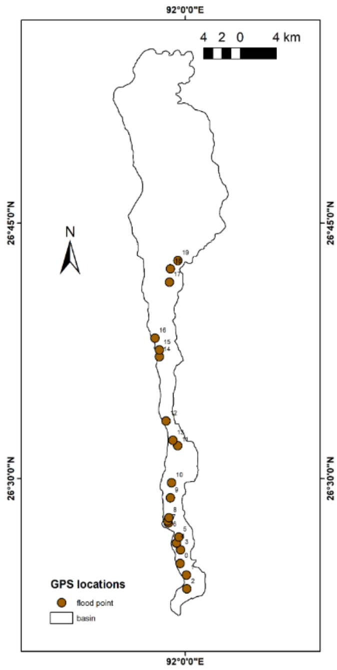


Fig. 3: Flood point of the study area.

Spatial Variables

In this study, eight spatial variables (Fig. 4) are utilized: Elevation, Slope, Soil Texture, Topographical Wetness Index, Distance from river, Distance from road, Precipitation, and Population density.

Elevation

Elevation plays a crucial role in identifying flood hazards and mapping flood zones, especially during the monsoon season when downstream areas are at high risk due to sedimentation and increased river flow. Accurate flood mapping requires analyzing factors influencing natural and human-made hazards. Understanding elevation variation is essential for predicting flood propagation in river basins. The Elevation data is downloaded from worldclim.org, and other datasets are clipped and resampled to 1 km spatial resolution using the nearest neighbor resampling technique. Furthermore, to enable model performance, all data layers have been converted to ASCII grid format.

Slope

The topography of a region, particularly its steepness and length, significantly influences discharge and flooding. Steep or high slopes lead to rapid precipitation runoff, while low or flat slopes are susceptible to waterlogging and high infiltration. A slope map was generated using SRTM data to depict these variations in terrain.

Topographic Wetness Index

The slope map, generated from DEM data, illustrates the topographic characteristics of the area. TWI (Topographic Wetness Index) indicates the movement of water down the slope based on gravitational forces and wetness distribution. This index, derived from DEM, plays a crucial role in regulating surface runoff, with wetter areas experiencing higher runoff rates. Additionally, Land Use Land Cover (LULC) data is essential for understanding surface runoff regulation. The general formula of TWI is $TWI = \ln(a/\tan\beta)$. Here, a is the upslope contributing area per unit contour length, and $\tan\beta$ represents the local slope gradient.

Step for TWI in ArcGIS

DEM>Fill>FDR>Flow Accumulation>Slope in degree>Radians of slope=(slope in degree*1.570796)/90>Tan slope= con (slope>0, tan(slope),0.001)> flow accumulation scaled=(flow accumulation+1)*cell size> TWI=ln (flow accumulation scaled/tan slope).

Soil Texture

Soil texture is recognized as a significant factor influencing both infiltration and runoff generation, as well as hazard occurrence. The data for this layer was obtained from the FAO soil portal, revealing that the soil texture in the study area consists mainly of loam and clay.

- Ao- Orthic Acrisols: The Orthic Acrisols soil type is characterized by a very deep, well-drained, dark brown loamy soil.
- Rd- Dystric Regosols: These are soils made up of loose mineral material that is not too coarse and doesn't have fluvic properties.

Distance from River Channel

The distance from the river basin region to the natural drainage, including all streams and rivers in the study area, was calculated using the Euclidean Distance tool in ArcGIS. This tool provides an estimate of the distance from the river basin to the natural drainage features within the study region.

Distance from Road

In ArcGIS mapping, considering the distance from roads can be crucial for various spatial analyses and decision-making processes. The distance from roads can provide valuable insights into accessibility, infrastructure planning, environmental impact assessment, and emergency response planning. In this study, the Euclidean Distance tool is utilized as it calculates the straight-line distance from each cell in a raster to the nearest road. While it is useful for basic distance analysis, but does not consider terrain or obstacles.

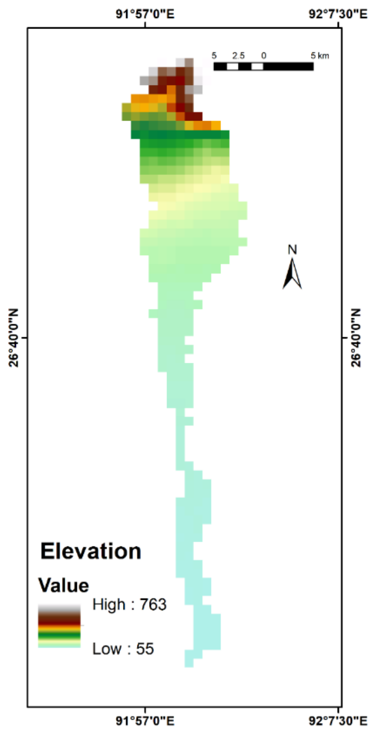
Precipitation

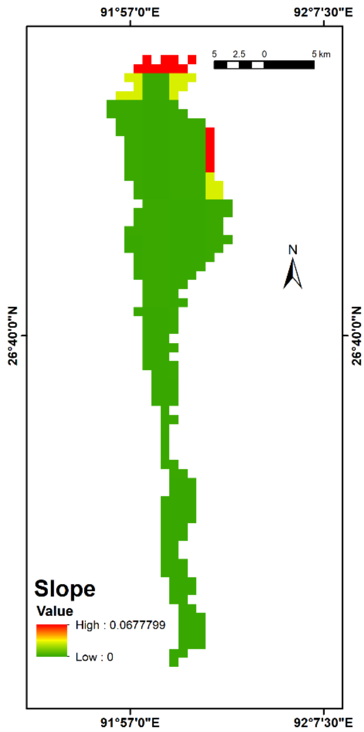
Rainfall is a crucial factor in this study as floods primarily occur during the monsoon season, often termed "rain-induced floods." The rainfall map for the study area was created using the Inverse Distance Weighted (IDW) approach, utilizing NASA’s POWER data access viewer. https://power.larc.nasa.gov/data-access-viewer/

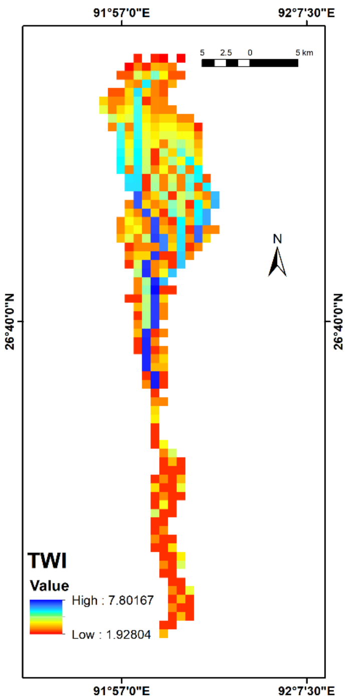
The map reflects the annual total rainfall of the year 2020, which was selected as it was considered a particularly flood-prone year.

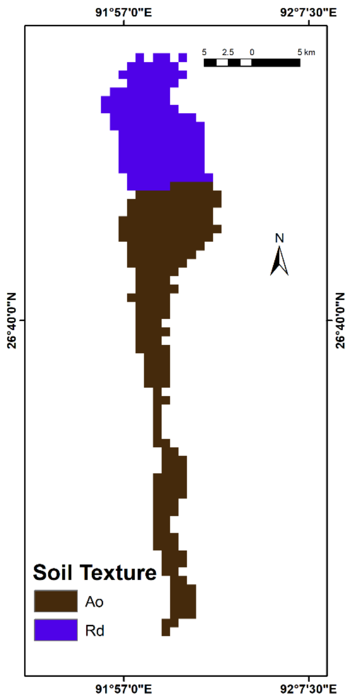
Population Density

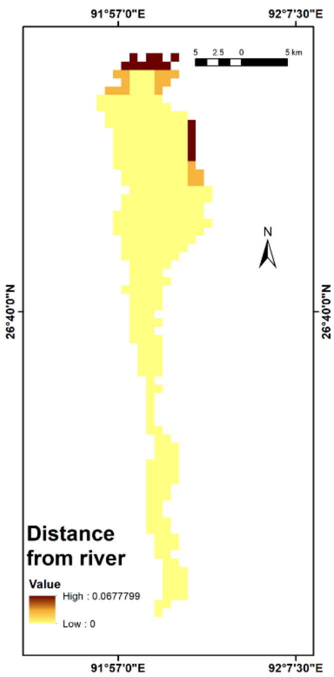
Population density is a critical factor in flood vulnerability research, as it helps analyze the social loss and damage suffered by communities in flood-prone areas. The data for this study was downloaded from SEDAC databases, which are estimates for the year 2020.

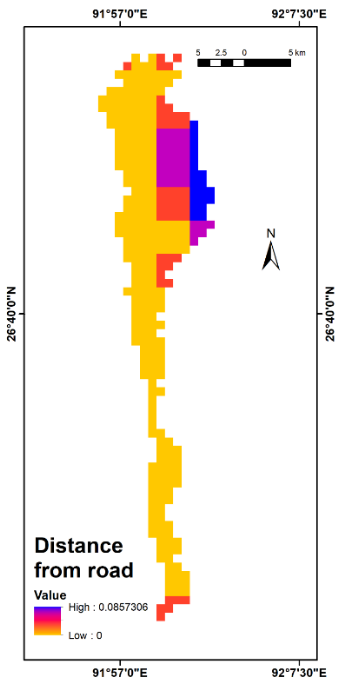





91°57'0"E
92°7'30"E
5 2.5 0 5 km
N
26°40'0"N
TWI
Value
High : 7.80167
Low : 1.92804


91°57'0"E
92°7'30"E
5 2.5 0 5 km
N
26°40'0"N
Soil Texture
Ao
Rd


91°57'0"E
92°7'30"E
5 2.5 0 5 km
N
26°40'0"N
Distance from river
Value
High : 0.0677799
Low : 0


91°57'0"E
92°7'30"E
5 2.5 0 5 km
N
26°40'0"N
Distance from road
Value
High : 0.0857306
Low : 0

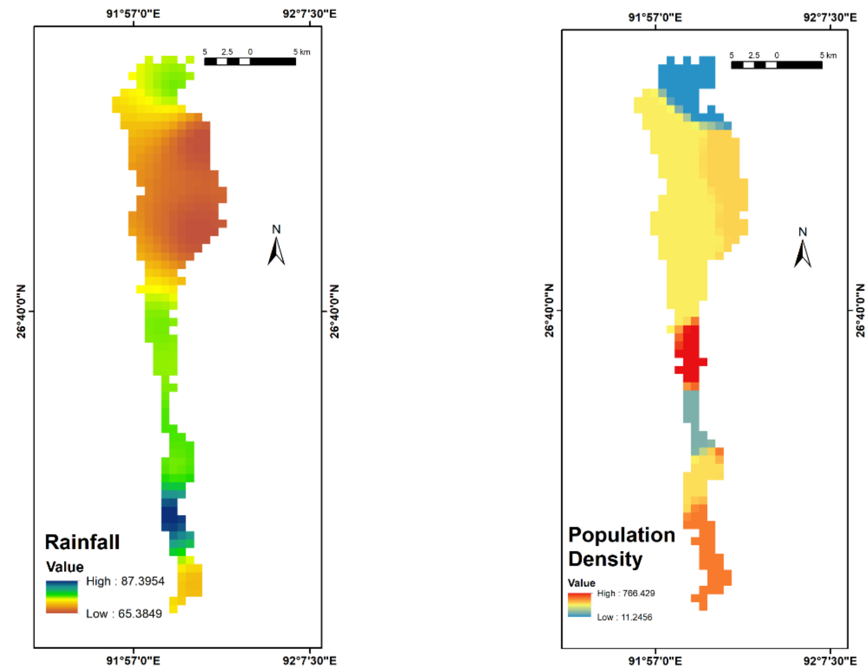


Fig. 4: Spatial variables used in MaxEnt modeling.

MaxEnt Modeling

A field survey was conducted frequently to cover different seasons (peak monsoon months and non-monsoon months), and samples were collected at various stages from the Noa River basin (Fig. 3). Field observations were recorded in a field notebook, and photographs were taken with an SLR camera (Canon SX 530 HS). For GPS locations, a"Garmin eTrex" model was used. In the present study, 70% of each hazard dataset was considered for model construction (training), while the remaining 30% of each hazard was used for validation in Maxent. The formula of maximum entropy used in maxent is

$$P^*(z(x_i)) = \exp(z(x_i)l)/\sum_i \exp(z(x_i)l)$$

Here, z is a vector of environmental variables at location x_i, and l is a vector of regression coefficients.

The Maxent model (Phillips & Dudik 2008), initially conceived at the American Museum of Natural History in partnership with AT&T-Research, continues to be developed and maintained by Steven J. Phillips and his team. Introduced in 2004 by Steven J. Phillips, Miroslav Dudík, and Robert E. Schapire, Maxent employs maximum entropy-based machine learning to estimate species distribution probabilities based on environmental parameters, gaining popularity among scientists for species distribution modeling.

Model Calibration and Validation

The model was run with 500 iterations with 10 training samples and 4 test samples. The test AUC and training AUC of the maxent model are both 0.85. The model utilized 10000 background points, and a total of 100 runs were conducted for the model. The area under the receiving operator curve (AUC) was used as the criterion for the goodness of fit, with the model showing the best AUC being selected for further analysis. The jackknife test (Fig. 5) was conducted to assess the importance of the variables. The final map of flood suitability was prepared with four different classes: high potential, moderate potential, low potential, and very low potential.

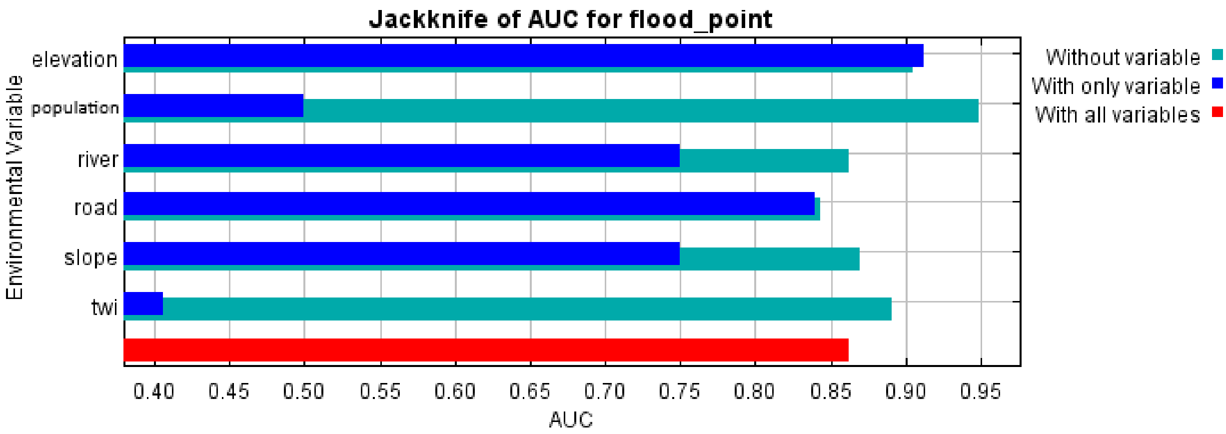


Fig. 5: Jackcknife test for evaluating the relative importance of environmental variables.

RESULTS AND DISCUSSION

The high flood hazard zones (Fig. 6) are estimated from the model to be within the low-lying area of the basin. The model also suggests that the soil (60 % contribution) and rainfall (32 % contribution) play significant roles in the estimation of spatial distribution of the flood hazard zones. Out of the total area of 241.85 sq. Km; 48 Sq. The km (20%) area is demarcated as having good and high flood potential. The total population in the river basin was 1,28,142 persons in 2001, out of which 57,159 persons were affected by flood. By the year 2011, the population increased to 1,41,146, with 63,044 persons directly and indirectly affected by floods. The population data are extracted from the district census handbook of Assam, India.

Relevant findings were also observed from different river systems within India, including Yamuna River (Sneha et al. 2018), Subarnarekha Basin (Das & Gupta 2021), Jia Bharali (Debnath et al. 2023), Rapti River (Khan et al. 2023), Ganga River (Yaseen 2024), and many others. Most of these studies depict flood hazards in river basins with varied topography and consistent rainfall, utilizing Analytical Hierarch Process (AHP) techniques for flood hazard mapping. However, the present study focuses on Maximum Entropy (MaxEnt) based modeling to identify probable flood hazard zones and yields similar results for the Noa basin, which is also characterized by mostly flat topography and ample rainfall (average of 210 cm).

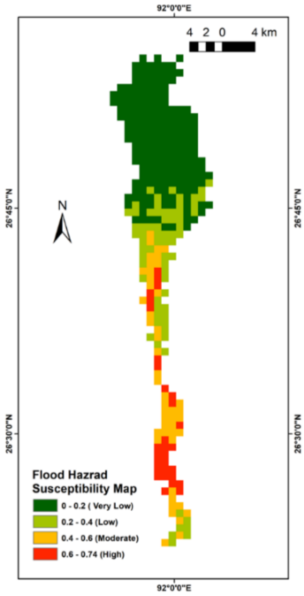


Fig. 6: Flood hazard susceptibility map.

CONCLUSION

This study is the first attempt to estimate probable flood hazard zones of the Noa river basin using Maximum Entropy modeling in GIS, integrating eight spatial variables: Elevation, Slope, Soil Texture, Topographical Wetness Index, Distance from river, Distance from road, Precipitation, and Population density. The flood hazard susceptibility map depicts flood-prone areas in mostly low-lying areas of the basin. Twenty percent of the area (48 sq. km) is affected by moderate and high floods during the monsoon months. Although this area is relatively small in percentage, the basin is highly populated (141,146 people in 2011), and 63,044 persons are affected by floods. The results generated by the MaxEnt model provide a quick estimation of the probable flood hazard zones. This is of great concern for policymakers, as it not only saves time but also provides valuable information for further investigation.

REFERENCES

Achour, Y. and Pourghasemi, H.R., 2020. How do machine learning techniques help in increasing the accuracy of landslide susceptibility maps? *Geoscience Frontiers*, 11(3), pp.871-883.

Bhattacharjee, N., 2008. Flood and Bank Erosion Problems in Darrang District, Assam: A Fluvio-Geomorphological study. Ph.D. thesis, Gauhati University, India.

Cerdà, A., 2007. Soil water erosion on road embankments in eastern Spain. *Science of the Total Environment*, 378(1-2), pp.151-155.

Das, S. and Gupta, A., 2021. Multi-criteria decision-based geospatial mapping of flood susceptibility and temporal hydro-geomorphic changes in the Subarnarekha basin, India. *Geoscience Frontiers*, 12(5), 101206.

Debnath, J., Sahariah, D., Nath, N., Saikia, A., Lahon, D., Islam, M.N. and Chand, K., 2023. Modelling on assessment of flood risk susceptibility at the Jia Bharali River basin in Eastern Himalayas by integrating multicollinearity tests and geospatial techniques. *Modeling Earth Systems and Environment*, 106, pp.1-27.

Garg, C. and Babu, A., 2023. Extreme flood calibration and simulation using a 2d hydrodynamic model under a multipurpose reservoir. *Nature Environment and Pollution Technology*, 22(2), pp.977-983.

Harshasimha, A.C. and Bhatt, C.M., 2023. Flood vulnerability mapping using maxent machine learning and analytical hierarchy process (AHP) of Kamrup Metropolitan District, Assam. *Environmental Science Proceedings*, 25(1), p.73.

Javidan, N., Kavian, A., Pourghasemi, H.R., Conoscenti, C., Jafarian, Z. and Rodrigo-Comino, J., 2021. Evaluation of multi-hazard map produced using MaxEnt machine learning technique. *Scientific Reports*, 11(1), p.6496.

Kavian, A., Hoseinpoor Sabet, S., Solaimani, K. and Jafari, B., 2017. Simulating the effects of land use changes on soil erosion using RUSLE model. *Geocarto International*, 32(1), pp.97-111.

Kelarestaghi, A. and Ahmadi, H., 2009. Landslide susceptibility analysis with a bivariate approach and GIS in Northern Iran. *Arabian Journal of Geosciences*, 2(1), pp.95-101.

Khan, R., Anwar, J., Said, S., Ansari, S., Pathan, A.I. and Sidek, L.M., 2023. GIS Based Flood Hazard and Risk Assessment Using Multi Criteria Decision Making Approach in Rapti River Watershed, India. In *International Conference on Intelligent Computing & Optimization*, pp.95-103. Cham: Springer Nature Switzerland.

Phillips, S.J. and Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), pp.161-175.

Sneha, S.M., Nagar, V., Moral, S. and Teli, B., 2018. Preparation of flood model and hazard estimation on Yamuna River (using GIS and remote sensing). *International Research Journal of Engineering and Technology*, 5(3), pp.3591-3597.

Yaseen, Z.M., 2024. Flood hazards and susceptibility detection for Ganga River, Bihar state, India: Employment of remote sensing and statistical approaches. *Results in Engineering*, 21, p.101665.