Journal of Natural Environment Hazards
University of Sistan and Baluchestan
Print ISSN: 2676 - 4377 Online ISSN: 2676 - 4385
Homepage: https://jneh.usb.ac.ir

Journal of Natural Environmental Hazards, Vol.14, Issue 43, March 2025

Predicting flood-prone areas using generalized linear and maximum entropy machine learning models

Abdulaziz Hanifinia[1], Hirad Abghari[2*]

1. PH.D. Student of Watershed Management, Urmia University, Iran
2. *Corresponding Author,* Associate Professor, Department of Range and Watershed Management, Urmia University, Iran

Article Info

Article type:
Research Article

Article history:
Received: 14 January 2024
Revised: 13 May 2024
Accepted: 08 June 2024

Keywords:
flood, Geographic Information System, SAGA-GIS, ROC curve, Lake Urmia.

ABSTRACT

The purpose of this study is to identify the effective factors, prepare flood risk prediction maps using machine learning models, and finally evaluate the efficiency of these models in the Zive watershed of Urmia. For this purpose, environmental and human factors including morphometric indices; Waterway Power Index (SPI), Slope Length Index (LS), Topographic Wetness Index (TWI), Topographic Position Index (TPI), Land Roughness Index (TRI), Mass Balance Index (MBI), Profile Curvature Index and The surface curvature index (Plan Curvature), rainfall, basin height, slope degree, slope direction, lithology, land use, normalized difference index of vegetation cover (NDVI), distance from waterway, distance from village and distance from fault were used. For this purpose, 96 flood spots were identified in the basin by using field visits and Google Earth images and sources received from the offices. Layers related to morphometric indices from the digital height model (12.5 x 12.5) meters and in the SAGA_GIS environment; And maps of environmental and human factors were prepared and digitized in the ArcGIS geographic information system. The evaluation results of two models using the ROC curve for machine learning (ML) models showed that the maximum entropy model with AUC=0.916 and the generalized linear model with AUC=0.902 have excellent performance in the field The results of the Kappa index for the superior model showed that environmental factors including geology, distance from waterways, height and slope have the greatest impact and the least impact related to profile curvature index factors. , land use, and mass balance index. Identifying high-risk areas and determining factors affecting the occurrence of floods in this basin can be very efficient in reducing possible damages.

Cite this article: Hanifinia, A., & Abghari, H. (2025). Predicting flood prone areas using generalized linear and maximum entropy machine learning models. Journal of Natural Environmental Hazards, 14(43), 7-10. DOI: 10.22111/jneh.2024.47730.2021

 Publisher: University of Sistan and Baluchestan
DOI: 10.22111/jneh.2024.47730.2021

[*] Corresponding Author Email: h.abghari@urmia.ac.ir

INTRODUCTION

Floods are one of the most common natural hazard phenomena worldwide, the occurrence and severity of which are influenced by two factors: human activities and climate change (Esmaili and Taheri, 1401 and Tabari, 2020). Floods cause enormous human and financial losses every year. Forecasts have shown that flood fatalities in recent years have reached more than 5,000 deaths per year in the world (El-Rawi et al., 2023). The factors affecting natural hazards in each basin depend on the specific conditions prevailing in that basin. However, in general, the factors are classified into one of the categories human factors, natural factors, and morphometric indicators. GIS and machine learning (ML)--based algorithms are currently some of the best methods for producing flood hazard maps (Panahi et al., 2021). The nonlinear nature of natural events such as floods can be well covered by these algorithms. Flood occurrence in the Ziveh watershed has been increasing due to predisposing factors such as hydrological, tectonic, topographic, and human factors such as land use change, leading to an increase in tangible and intangible damages in this region. The main objective of this study is to prepare a zoning map of flood-prone areas and predict flood occurrence using powerful machine learning algorithms. Secondary objectives include identifying the most effective factors in this phenomenon using the Kappa index and evaluating the accuracy and reliability of these models.

DATA AND METHODS

The Ziveh watershed, with an area of 21,686 hectares, is located in the southwest of Urmia City and the Margavar district of Silvaneh. The outlet river of the basin is one of the main tributaries of Lake Urmia. The average elevation of the Ziveh basin is about 2265 meters, and the average slope of the Ziveh basin is about 16.5 degrees.

Ninety-six flood points were prepared using field visits and the Natural Resources Department. Information layers of elevation, slope degree, slope direction, watercourse network, and morphometric indicators were prepared from the digital elevation model map. Due to the lack of rain gauge stations, the rainfall layer was extracted from 13 stations around the basin using the simple kriging method. The land use layer was prepared from updated images of 2022. To prepare the lithology layer and distance from faults, the topographic map of 1:100,000 Silvaneh sheet was used. The layer of villages in the region was extracted using Google Earth software and the NDVI layer from Sentinel_2 images of June. The advantages of machine learning models include the ability to learn from data, adaptability to changes, generalizability in different domains, flexibility in parameter setting, and the ability to make accurate and robust predictions. These models are known as powerful tools for solving complex problems and providing intelligent solutions in many fields (Ha et al., 2021; Riazi et al., 2023). In this study, two linear models, the maximum entropy model, and the maximum entropy model, were used to prepare the hazard map, and the results of these two models were evaluated through the ROC curve.

RESULTS AND DISCUSSION

The present study aimed to zone flood sensitivity using the maximum entropy (Max Ent) and generalized linear model (GLM) machine learning models in the Ziveh watershed of Urmia. After preparing a map of the distribution of floods and identifying the factors affecting their occurrence, a map of the region's susceptibility to flood occurrence was prepared using the above two methods. The results of the evaluation of the models using the ROC curve for the above method show that the results of both models have excellent performance (AUC> 0.9) in predicting areas sensitive to flood events. Also, the results of the Kappa index in the superior model (maximum entropy) to determine the factors affecting the occurrence of floods in the area in question show that, respectively, geology, distance from the watercourse, height, and slope, and surface roughness index have the most significant impact, and profile curvature index, land use, and balance index have the least impact on the surface roughness index.

CONCLUSION

The results of the study and modelling showed that, on average, more than 60 percent of the basin is located in high-risk and critical areas. Although in this study, the evaluation of the results of the maximum entropy model (0.916) was better than the generalized linear model (0.902), this difference was minimal, tiny and only about 0.1. The area under the curve for both models shows the excellent performance of the maps created by these models. In general, managing and controlling floods and flood-prone areas where human activities are the leading cause of their occurrence is much more complicated than when environmental or natural factors alone have an effect. In all mountainous areas, floods that occur naturally and inherently have a self-regulating state and nature itself controls and manages these floods in most cases. Controlling floods caused by human intervention in nature is often impossible and will be very costly if done. Based on the results of the Kappa index, it was determined that factors such as geology, distance from the watercourse, altitude, and slope had the most significant impact on the occurrence of floods. Therefore, in this region, the impact of natural factors is more significant than that of human factors, and the slightest intervention by humans can cause a critical situation.

REFERENCES

References (in Persian)

Esmaili, R., & Taheri, M. (2022). Evaluation of flood hazard areas with fuzzy approach, Case study: Downstream of Neka catchment, Mazandaran province. Journal of Natural Environmental Hazards, 11(34), 145-158. [In Persian]

References (in English)

Tabari, H. (2020). Climate change's impact on floods and extreme precipitation increases with water availability. Scientific reports, 10(1), 13768.

El-Rawy, M., Elsadek, W. M., & De Smedt, F. (2023). Flood hazard assessment and mitigation using a multi-criteria approach in the Sinai Peninsula, Egypt. Natural Hazards, 115(1), 215-236.

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

Ha, H., Luu, C., Bui, Q. D., Pham, D. H., Hoang, T., Nguyen, V. P., ... & Pham, B. T. (2021). Flash flood susceptibility prediction mapping for a road network using hybrid machine learning models. Natural hazards, 109(1), 1247-1270.

Riazi, M., Khosravi, K., Shahedi, K., Ahmad, S., Jun, C., Bateni, S. M., & Kazakis, N. (2023). Enhancing flood susceptibility modelling using multi-temporal SAR images, CHIRPS data, and hybrid machine learning algorithms. Science of The Total Environment, 871, 162066.