earth

MDPI

Article

Deep Learning-Driven Flood Susceptibility Mapping in Northwest Iran: A Comparative Study of CNN, MLP, and DNN Approaches

Raoof Mostafazadeh [1,*], Ali Nasiri Khiavi [2] and Shahnaz Mirzaei [3]

[1] Department of Natural Resources, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil 5619911367, Iran
[2] Ardabil Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education and Extension Organization (AREEO), Ardabil 5695157451, Iran; a.nasiri@areeo.ac.ir
[3] Department of Watershed Management, Faculty of Range and Watershed Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan 4913815739, Iran; mirzaeishahnaz@gmail.com
* Correspondence: raoofmostafazadeh@uma.ac.ir; Tel.: +98-9144815743; Fax: +98-4533512204

Abstract

Floods are frequent and major natural hazards in watersheds, causing severe environmental, economic, and societal damage. This study compares convolutional neural networks (CNN), multilayer perceptrons (MLP), and deep neural networks (DNN) for flood susceptibility mapping in the snow-fed Qareh-Sou watershed, northwest Iran. The approach integrates 24 candidate geo-environmental and hydrological predictor variables, of which 10 were retained by principal component analysis (PCA) and used as inputs to the deep learning models, and validates the resulting maps against observational discharge data. CNN showed superior performance, with R^2 = 0.87, mean absolute error (MAE) = 0.19, and area under the receiver operating characteristic curve (AUC-ROC) = 0.93, outperforming MLP (R^2 = 0.72, MAE = 0.25) and DNN (R^2 = 0.64, MAE = 0.55). The CNN-derived map classified 28.88% of the watershed as very high risk, 33.45% as high risk, and 10.31% as very low risk. Watershed morphology, particularly slope and drainage density, dominated flood generation, while climatic factors were secondary. These results demonstrate the effectiveness of a CNN-based flood risk assessment and provide practical implications for watershed management and disaster mitigation. The framework can be adapted to other flood-prone regions, offering a tool for policymakers and environmental planners. Its novelty lies in comparing three deep learning algorithms within a PCA-based framework and validating CNN-derived zones with gauge-based maximum instantaneous discharge records.

Keywords: convolutional neural network; deep learning; flood susceptibility mapping; geo-environmental factors; watershed management

Academic Editor: Sajjad Ahmad

Received: 20 September 2026
Revised: 19 November 2026
Accepted: 21 November 2026
Published: 25 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Floods are among the most important and common natural phenomena, especially in watershed areas, and can have devastating and widespread impacts on the environment, economy, and society. Flooding typically results from heavy rainfall, snowmelt, or a combination of meteorological and human factors. These events can lead to infrastructure damage, destruction of farms and agricultural lands, water source pollution, and

even loss of life [1,2]. Additionally, climate change and improper management of water and land resources increase the intensity and frequency of floods, indicating the need for effective resource planning and measures to mitigate their risks [3,4]. Anthropogenic climate change and improper water and land resource management have made watersheds highly susceptible to extreme events such as floods, droughts, and heatwaves [5]. Traditional methods and empirical models for spatial and distributed flood modeling in watersheds typically rely on simplified assumptions and linear relationships. This approach can lead to limitations in accurately and reliably predicting flood behaviors [6]. One of the main challenges facing these models is their inability to understand and analyze complex, non-linear patterns present in Hydrological Systems [7]. These systems are influenced by various factors such as soil type, vegetation cover, terrain slope, and weather variations, making it difficult for conventional empirical models to effectively capture the intricate relationships among these variables [8].

Therefore, the use of innovative techniques like Machine Learning (ML) algorithms and Deep Learning (DL) in Spatial Flood Modeling (SFM) and prediction represents a significant advancement in crisis management and risk reduction related to floods [9]. Given the inherent complexity and diversity of factors influencing flooding, deep learning algorithms are capable of analyzing and processing large volumes of high-dimensional and complex data, including geographic, climatic, hydrological, and land use change information [10].

Extensive research has been conducted on the application of Artificial Intelligence (AI) algorithms across various scientific fields, particularly in SFM within watersheds. Additionally, studies have examined the effects of Land Use (LU) changes and climatic factors on runoff. For example, Shang et al. [11] investigated the impact of LU and Climate Change (CC) on surface runoff in the Upper Heihe River. Their results indicated that CC contributed approximately 87.15% and LU change about 12.85% to variations in surface runoff, with a greater influence from climate factors. Regarding this, Hu and Shrestha [12] analyzed how LU change affected surface runoff in a Chinese urban area using the Soil Conservation Service-Curve Number (SCS-CN) model. They found that the trend in surface runoff increased during 1984–2009, coinciding with the expansion of impervious surfaces in central Beijing, and decreased from 2009 to 2019. Avand et al. [13] compared game theory and machine learning for flood vulnerability in the Tajan watershed, resulting in similar prioritization across algorithms but more detailed and accurate assessments from distributed versus semi-distributed modeling. Dawson et al. [14] developed flood susceptibility maps for England using ML techniques. They employed algorithms like Classification and Regression Trees (CART), Random Forest, and XGBoost for classification. The results showed that XGBoost, with an AUC of 0.93, was the optimal algorithm. The study concluded that the model effectively identified and classified diverse regions based on flood susceptibility. Similarly, Al-Sheriadeh and Daqdouq [15] examined flood susceptibility mapping in Jordan using ML algorithms. The study focused on two contrasting climatic regimes, mountainous and desert regions. Results indicated that the Random Forest (RF) algorithm outperformed others in producing accurate flood susceptibility maps, with 11% of the Al-Buaida watershed classified as high to very high risk, and 2.5% of the Zarqa watershed falling into these categories. Ultimately, they concluded that ML algorithms can effectively produce flood susceptibility maps and are suitable alternatives to physical modeling approaches. In another study, Nasiri Khiavi et al. [16] employed game theory algorithms to identify flood-prone sub-watersheds considering flood potential generation in the north of Iran. Their results showed that in physical and hydrological modeling frameworks, the prioritization of sub-watersheds primarily depended on physical factors like slope, area, and concentration time.

A review of the research background indicates that in recent decades, numerous researchers have employed various tools to map floods from different perspectives, such as flood potential (flood generation potential) and flood control. For example, studies such as [11,12] primarily focused on the hydrological aspect and physical modeling of floods, with relatively few utilizing advanced DL algorithms. Furthermore, many previous studies have applied individual machine learning models without systematically comparing deep learning architectures (e.g., [13–16]), and few have validated their outputs against observed discharge data [16]. Furthermore, the combined effects of watershed morphometry, land use, and climatic factors on flood susceptibility in snow-fed watersheds of Iran remain insufficiently explored. Additionally, many of these studies have relied on a limited set of criteria and variables for distributed or semi-distributed flood modeling. In contrast, this study aims to incorporate a broader range of variables (based on a comprehensive review of relevant sources), including physiographic, climatic, and human-related factors, totaling 24 variables. Accordingly, the present study aimed to model and map SFM in northwest Iran using the latest DL algorithms, including Convolutional Neural Networks (CNN), Multilayer Perceptrons (MLP), and Deep Neural Networks (DNN). A key distinguishing aspect of this study is the comparative assessment of three DL architectures (CNN, MLP, and DNN) specifically within a snow-fed mountain watershed, complemented by an independent validation against gauge-measured discharge records (MIFD and AAD) rather than relying exclusively on discrete flood inventory points. Accordingly, the main objectives of this study are: (1) to identify the key geo-environmental and hydrological factors controlling flood susceptibility using PCA; (2) to model and map flood susceptibility using CNN, MLP, and DNN algorithms; (3) to compare the predictive performance of these algorithms using multiple error metrics; and (4) to validate the resulting susceptibility maps against observational discharge data from river gauge stations.

2. Materials and Methods

2.1. General Characteristics of the Studied Watershed

The Qareh-Sou Watershed is situated within the Ardabil province, Iran. The Qareh-Sou River is a significant tributary of the Aras River in Iran, collecting water from a broad region of Ardabil Province and transferring it to the Aras River. The minimum elevation in the watershed is 1290 m, recorded at the Samian gauging station at the outlet, while the maximum elevation reaches 4781 m, with an average elevation of 1787 m. The average slope of the watershed is about 11.4% [17]. LU in this area predominantly includes rainfed and irrigated agriculture, rangelands, forests, and residential zones. The highest seasonal rainfall occurs in spring, whereas summer experiences the lowest. Overall, the Qareh-Sou watershed functions primarily as a snow-fed and flood-prone system, with a substantial portion of river flow resulting from snowmelt stored during winter and late March at higher elevations. Geologically, the watershed is dominated by volcanic and sedimentary formations, including andesitic lavas, tuffs, and limestone, which influence infiltration and runoff generation. The climate is semi-arid to cold, with mean annual precipitation ranging from approximately 300 to 500 mm, mostly occurring as snow at higher elevations. The predominant land uses are rangelands and agricultural lands, with limited forest cover and scattered residential areas.

The Qareh-Sou watershed has a history of destructive floods. A 25-year flood event generates a peak discharge of approximately 255 m^3/s at the Doost-Beyglou station, inundating 1085 ha of agricultural lands and orchards with a floodplain width up to 550 m. The river's meandering pattern (sinuosity index 1.51) and agricultural encroachment into the riverbed have exacerbated flood risks, prompting recent dredging of a 14.5-km reach (Aladizgeh to Namin) to mitigate inundation. Notable events include the June 2023 floods

(causing 250 billion Rials in damage to Nir's Aglaghan Bridge and inundating Samian) as well as historical and recent hailstorm floods in Namin County that damaged orchards, residential units, and utility networks. These records justify the necessity for robust flood susceptibility mapping in this morphologically vulnerable basin.

The high-elevation snowpack gradually melts in summer, feeding the rivers [18]. The location map of the study sub-watersheds is shown in Figure 1.

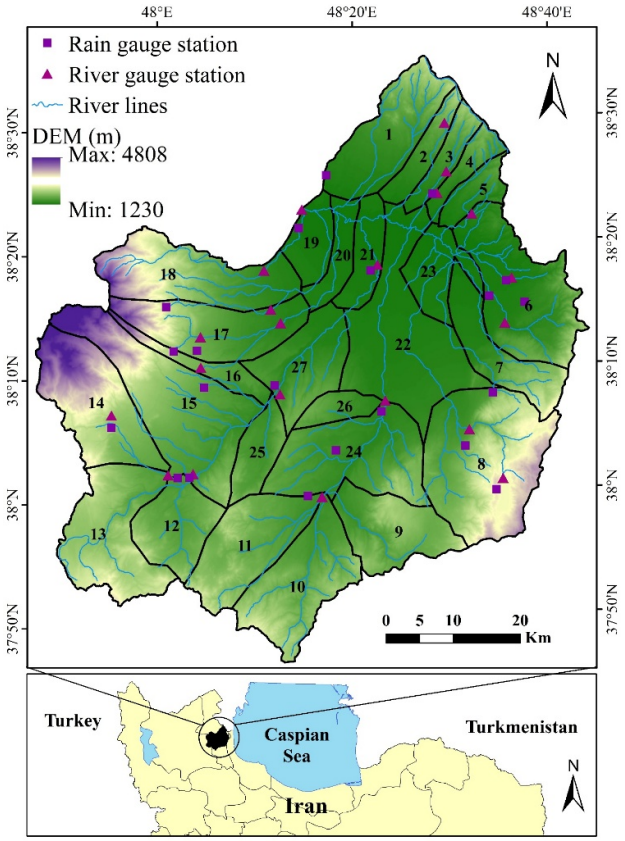


Figure 1. Geographical location map of study sub-watersheds.

2.2. Research Methodology

2.2.1. Preparation of Baseline Maps and Geo-Environmental Criteria

To support the analysis, various spatial and geo-environmental data layers were obtained. Maps of watershed boundaries, topography, regional drainage network, vegetation and LU, demographic information, and roads were collected from the General Directorate of Natural Resources and Watershed Management of Ardabil Province. Additionally, daily discharge data and maximum instantaneous flood discharge data (1971-09-23 to 2021-08-23) for all river gauge stations, as well as rainfall and temperature data (1981-09-23 to 2021-08-23), were sourced from the Regional Water Company of Ardabil Province and the National Meteorological Organization, respectively.

The data sources, types, spatial scales/temporal periods, and roles of the variables used in this study are summarized in Table 1.

Table 1. Data sources, types, spatial scales/temporal periods, and roles of the variables used in this study.

Variable	Source	Type	Scale/Period	Role
Elevation	SRTM DEM (GEE)	Continuous	30 m	Predictor
Slope	Derived from DEM	Continuous	30 m	Predictor
Plan curvature	Derived from DEM	Continuous	30 m	Predictor
Profile curvature	Derived from DEM	Continuous	30 m	Predictor
Drainage density	ArcGIS 10.8	Continuous	Sub-watershed	Predictor
TWI	SAGA-GIS	Continuous	30 m	Predictor
CN	ArcGIS 10.8	Continuous	Sub-watershed	Predictor
Land use	Sentinel-2 (GEE)	Categorical	10 m	Predictor
Soil type	Natural Resources Dept.	Categorical	1:250,000	Predictor
Max annual rainfall	MODIS (GEE)	Continuous	1 km	Predictor
Morphometric parameters	ArcGIS/Excel	Continuous	Sub-watershed	Predictor
MIFD	Gauge stations	Continuous	1971–2021	Validation
AAD	Gauge stations	Continuous	1971–2021	Validation

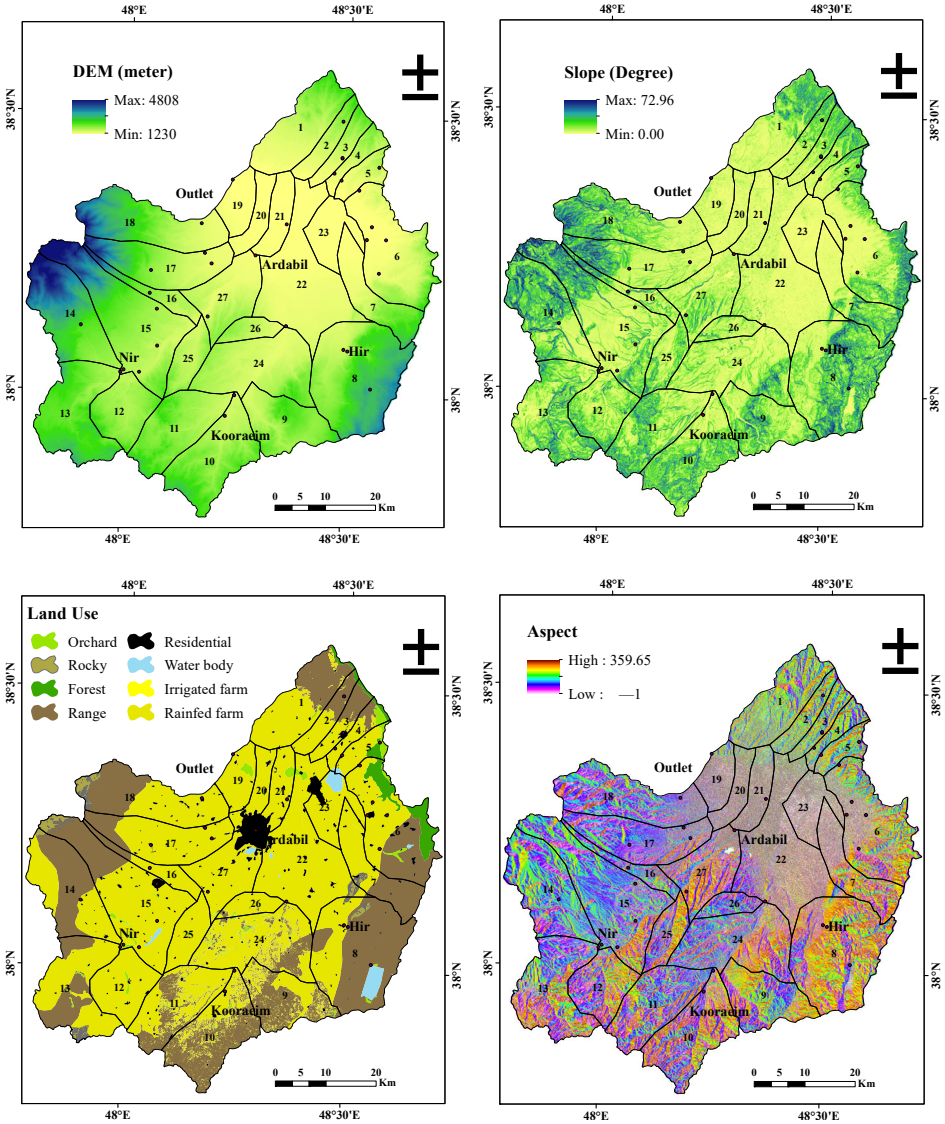
The data sources, types, spatial scales, and temporal periods of all variables are summarized in Table 1. A total of 24 geo-environmental and hydrological predictor variables were selected for flood susceptibility modeling [13,19]. To provide a clear summary of these variables, they are categorized into five groups: (i) topographic attributes (n = 3), including elevation, plan curvature, and profile curvature; (ii) hydrological variables (n = 5) comprising the Topographic Wetness Index (TWI), Curve Number (CN), drainage density, impervious lands, and distance from the river; (iii) morphometric parameters (n = 14), including area, perimeter, watershed width, main stream length, total channel length, main channel slope, watershed slope, form factor, compaction coefficient, concentration time (Kirpich's method), equivalent rectangle length, equivalent rectangle width, iso-surface circle diameter, branching ratio, elongation ratio, and roundness ratio; (iv) climatic factors (n = 1) represented by maximum annual rainfall; and (v) land use/land cover (n = 1). The target variable for model training was the flood generation potential (FGP) of each sub-watershed, derived from the same set of criteria using PCA-based weighting. For validation, independent observational data including maximum instantaneous flood discharge (MIFD) and average annual discharge (AAD) from 21 gauge stations were used.

In this study, the dependent (target/response) variable is the flood generation potential (FGP) of each sub-watershed. FGP is a continuous index (0–1), derived from PCA-weighted geo-environmental and hydrological criteria, and was used as the response variable for training and testing the CNN, MLP, and DNN models. The 24 geo-environmental and hydrological variables listed in Table 1 are the independent (predictor) variables. Maximum instantaneous flood discharge (MIFD) and average annual discharge (AAD) are not predictors; they are independent observational variables used only for external validation of the final flood susceptibility maps. For the binary ROC-AUC evaluation, a binary dependent variable was derived from FGP using a threshold of 0.5 (FGP > 0.5 = flood-prone; FGP ≤ 0.5 = non-flood-prone).

The Digital Elevation Model (DEM) was generated using Google Earth Engine (GEE) based on 30-m SRTM data [20]. Slope was mapped and classified in ArcGIS 10.8. For LU, Sentinel-2 images (1–28 June 2024) were corrected in GEE. Several classifiers were tested (minimum distance to mean, decision tree, RF, SVM, and LibSVM [21,22]) using training samples randomly collected via on-site GEE monitoring for residential, rangeland, agricultural, and water body classes. Following an accuracy assessment, the most suitable algorithm was chosen for land use classification [19].

The maps related to surface curvature, distance from the river, and drainage density were prepared using ArcGIS 10.8 software. TWI was calculated using SAGA-GIS software [23]. Soil and soil erodibility maps were obtained from the General Department of Natural Resources and Watershed Management of Ardabil Province. To generate maps of average and maximum annual precipitation, data from the MODIS sensors were used within the GEE platform [24,25]. Finally, physiographic criteria, including the concentration time based on Kirpich's method [2], main channel slope, shape coefficient, compactness coefficient, equivalent rectangle length, equivalent rectangle width, equivalent circle diameter, branching ratio, elongation ratio, and roundness ratio, were quantified in Microsoft Excel. These were derived from various shape files exported from ArcGIS 10.8 for each sub-watershed.

Some important maps, including DEM, slope, geographical orientation, land use, soil, climatology and geological units, are presented in Figure 2. Also, the values of some important criteria related to the shape of the watershed are presented in Table 2.



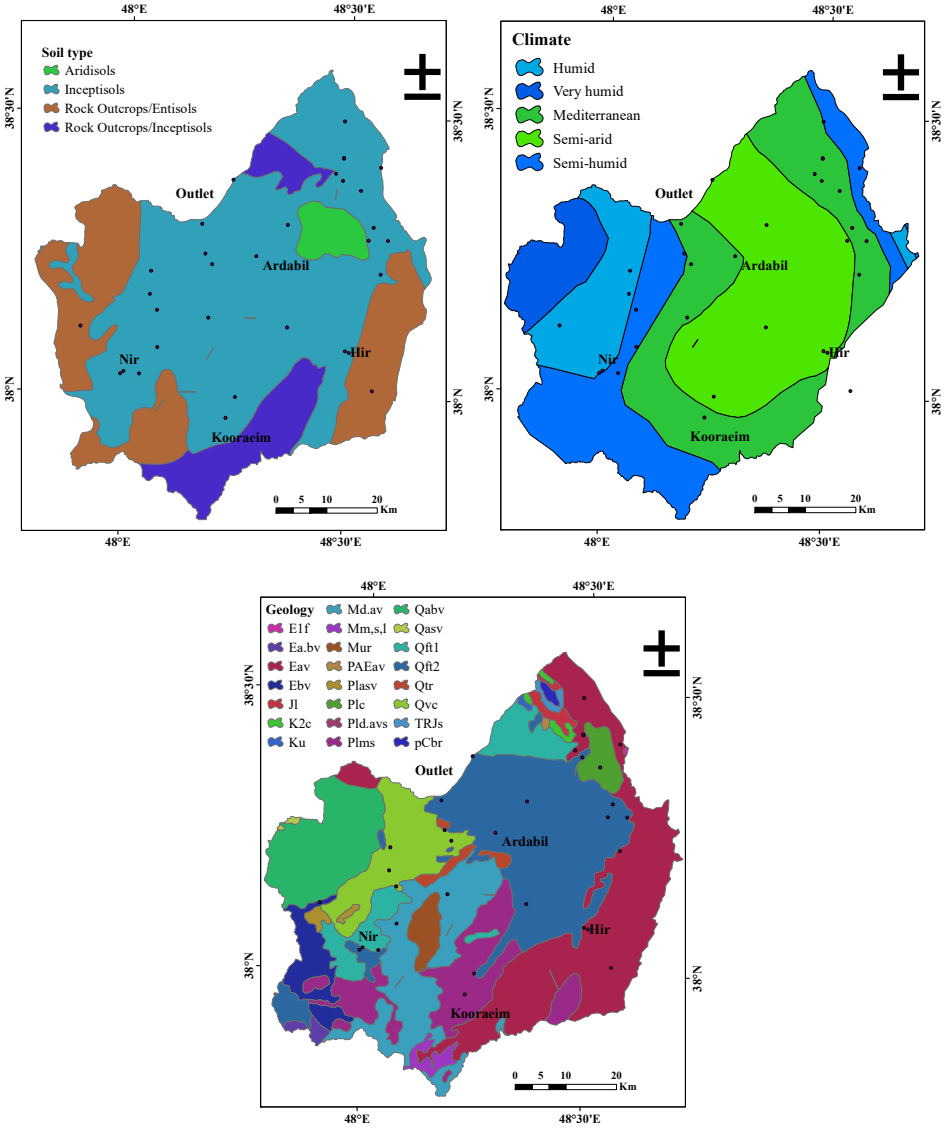


Figure 2. Some important environmental maps in the Qareh-Sou watershed, Iran.

Table 2. Quantitative values of some important criteria in the Qareh-Sou watershed, Iran.

SWs	Area (Km^2)	Perimeter (Km)	Stream Length (Km)	Form Factor (-)	Roundness Ratio (-)	Elongation Ratio (-)	Compactness Coefficient (-)
1	186.71	64.40	81.26	0.03	0.57	0.57	1.33
2	61.48	44.93	18.26	0.18	0.38	0.42	1.62
3	56.10	43.33	36.90	0.04	0.38	0.42	1.63
4	48.27	41.32	24.18	0.08	0.36	0.40	1.68
5	51.75	37.90	24.32	0.09	0.45	0.48	1.49
6	262.26	83.08	94.18	0.03	0.48	0.50	1.45
7	132.16	68.12	42.22	0.07	0.36	0.41	1.67
8	348.10	77.78	81.85	0.05	0.72	0.75	1.18
9	157.11	58.34	21.50	0.34	0.58	0.59	1.31
10	215.49	73.59	73.10	0.04	0.50	0.52	1.41
11	189.71	61.33	43.98	0.10	0.63	0.64	1.26
12	130.20	46.61	34.83	0.11	0.75	0.81	1.15
13	225.72	80.86	50.02	0.09	0.43	0.46	1.52

14	224.29	72.21	20.99	0.51	0.54	0.55	1.36
15	325.68	99.03	94.05	0.04	0.42	0.45	1.55
16	55.54	59.72	14.45	0.27	0.20	0.28	2.26
17	189.25	69.76	48.40	0.08	0.49	0.51	1.43
18	228.10	78.94	98.57	0.02	0.46	0.48	1.47
19	71.20	35.78	25.59	0.11	0.70	0.72	1.20
20	62.38	45.52	32.44	0.06	0.38	0.42	1.63
21	74.43	43.00	24.64	0.12	0.51	0.52	1.41
22	327.32	111.32	74.89	0.06	0.33	0.39	1.74
23	92.97	55.16	60.49	0.03	0.38	0.43	1.61
24	248.72	75.25	100.88	0.02	0.55	0.56	1.35
25	75.39	41.69	11.40	0.58	0.54	0.56	1.35
26	44.29	31.66	13.41	0.25	0.55	0.56	1.34
27	150.99	57.59	41.69	0.09	0.57	0.58	1.32

2.2.2. Application of PCA in the Selection of Selected Criteria in SFM

As a crucial first step, the criteria values for SFM conditioning were meticulously quantified. Then, Principal Component Analysis (PCA) was employed as a sophisticated technique to systematically determine the main hydrological and geo-environmental factors that strongly influence sub-watershed prioritization according to SFM [26–28].

PCA is a reliable method for pinpointing influential variables and identifying those with the greatest impact on the analysis [29]. This dimensionality reduction is particularly valuable in environmental studies, where datasets are often complex and contain interrelated or auto-correlated variables. All statistical analyses, including PCA, were performed using Python (version 3.13.5) through the Anaconda distribution, specifically utilizing the scikit-learn library for the PCA implementation. Prior to PCA, the overall Kaiser–Meyer–Olkin (KMO) measure was examined to confirm the factorability of the entire dataset (overall KMO > 0.6). The selection rule was based on the Kaiser criterion, retaining principal components with eigenvalues greater than 1. Variables with factor loadings greater than 0.5 were retained for subsequent DL modeling.

Main components were determined by examining the cumulative variance explained by each, following standard procedures [30]. Once the primary components were identified, relevant criteria were selected for each based on the significance of their correlation coefficients [31].

To quantitatively construct the target variable (FGP), a systematic five-step procedure was implemented based on the PCA results. (1) Standardization: First, to eliminate the effects of different measurement units (e.g., meters for elevation, mm for rainfall), all 24 predictor variables were standardized to a common 0–1 range. (2) Component Extraction: Second, PCA was performed on the standardized dataset. Principal components with eigenvalues greater than 1 were retained, as they captured the most significant variance and addressed multicollinearity among the predictors. (3) Weight Assignment: Third, a composite weight was calculated for each variable by multiplying its factor loading (its correlation with the component) by the percentage of total variance explained by that specific component. This ensured that variables with higher contributions received greater weights. (4) Normalization: Fourth, the absolute values of these weights were normalized so that their sum equaled one, allowing for a consistent and comparable scale across all 24 predictors. (5) Final Calculation: Finally, the FGP for each sub-watershed was computed as the sum of its standardized predictor values, each multiplied by its respective normalized weight. This yielded a continuous quantitative index representing the inherent flood generation potential of each sub-watershed based on its unique geo-environmental characteristics.

2.2.3. Application of DL Algorithms in FS Modeling and Mapping

Convolutional Neural Networks (CNNs) are a type of deep learning algorithm designed to process structured grid data, such as images. They are especially effective for tasks like image classification, object detection, and image segmentation [32,33]. Implementing CNNs in Python for spatial flood modeling was straightforward using the PyTorch library [34]. The CNN architecture consisted of two convolutional layers with 32 and 64 filters (3 × 3 kernels), each followed by a 2 × 2 max-pooling layer and a fully connected layer with 128 neurons. To apply the CNN to tabular data, the 1D vector of 24 predictors was grouped and reshaped into a 4 × 6 pseudo-image tensor (Number of samples, 4, 6, 1). This arrangement allows 3 × 3 convolutional kernels to capture spatial interactions between adjacent variable groups, which flattened MLP/DNN inputs cannot easily model. To rule out overfitting, the architecture was kept intentionally shallow (two convolutional layers), evaluated on an independent 30% test set, and externally validated against observed discharge data (MIFD/AAD), confirming the model learned genuine hydrological patterns rather than memorizing the training data.

The MLP model comprised three hidden layers with 64, 32, and 16 neurons, respectively, while the DNN model consisted of four hidden layers with 128, 64, 32, and 16 neurons. All hidden layers used the ReLU activation function, and the output layer used a sigmoid activation. Models were trained using the Adam optimizer (learning rate = 0.001), a batch size of 8, and 100 epochs. Land use and soil type were one-hot encoded, and continuous variables were normalized to the range [0, 1]. The loss function was mean squared error (MSE) for regression and binary cross-entropy for classification. After quantifying the key factors affecting FS modeling and mapping, various DL algorithms, including CNN, MLP and DNN, were employed using the Python programming language to model SFM. The target was sub-watershed flood generation potential (FGP), derived from PCA-weighted criteria. It should be noted that the error metrics (R^2, MAE, MSE, RMSE) were computed against the target variable (FGP) derived from PCA-weighted predictors, rather than an observed flood inventory map. The final susceptibility map was then independently validated using observed discharge data (MIFD and AAD) from gauge stations. Since the same predictors were used to construct FGP and train the DL models, some circularity is possible, a common limitation in data-scarce regions. FGP was normalized to 0–1 and divided into five equal-interval classes: very low (0–0.2), low (0.21–0.4), moderate (0.41–0.6), high (0.61–0.8), and very high (0.81–1). To compute the AUC-ROC for binary classification, a threshold of 0.5 was applied to the continuous FGP values. This threshold was selected because FGP was normalized to a 0–1 scale; thus, 0.5 represents the theoretical median of the index, serving as the natural conceptual boundary between areas with low (<0.5) and high (>0.5) inherent flood generation potential. About 70% of the data was used for training, while 30% was reserved for testing. Given the limited number of sub-watersheds (n = 27), a 70/30 split was adopted to ensure sufficient training data while retaining an independent test set for evaluation. This split ratio is consistent with common practice in small-sample hydrological modeling. Additionally, to determine the optimal model, error metrics such as the correlation coefficient, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), and ROC curves [8] were used. For regression tasks, MAE, MSE, RMSE, and R^2 were used. For classification, precision, recall, F1-score, balanced accuracy, confusion matrix, and AUC-ROC were calculated. Given the five-class susceptibility scheme, class-specific metrics were also computed to assess the model's ability to distinguish among susceptibility levels. After modeling based on the selected criteria related to SFM, different DL approaches were mapped using ArcGIS 10.8 software. Since the DL models generated discrete FGP values for each of the 27 sub-watersheds, a spatial interpolation step was required to produce the continuous high-resolution susceptibility map This was performed using the Inverse Distance

Weighting (IDW) technique in ArcGIS 10.8, which estimates pixel values based on the weighted average of neighboring sub-watershed centroids. While IDW effectively visualizes the spatial continuum of flood potential, it may introduce slight edge-smoothing artefacts at sub-watershed boundaries; however, the overall spatial patterns remain highly representative of the discrete model outputs. Because the IDW output can extend beyond the watershed boundary, the class proportions derived from the CNN map were subsequently applied to the reference watershed area to obtain the final. It is important to note that, due to differences in the range and scope of the classification numbers (and as a result of the functional and structural differences in each approach), the weights were initially normalized between zero and one. This normalization process was carried out to standardize and homogenize the weights, enabling effective comparisons among them.

Next, SFM was classified into five different categories: very low (0–0.2), low (0.21–0.4), moderate (0.41–0.6), high (0.61–0.8), and very high (0.81–1). This classification was based on the weights and priorities derived from sub-watersheds. To enable a more detailed analysis, spatial models were developed to identify and understand SFM within the Qareh-Sou watershed.

2.2.4. Validation of DL Algorithms for SFM and Selection of the Optimal Approach

After preparing the SFM maps, observational and real data on the average annual discharge (measured in cubic meters per second) and the maximum instantaneous flood discharge (also in cubic meters per second) within the sub-watersheds were used for validation and optimization of the employed DL methods. These observational data served as the benchmark for assessing and comparing the performance of different DL approaches. The validation was performed by comparing the CNN-derived susceptibility classes with observed MIFD and AAD values at each station. Stations with higher discharge values were expected to fall within high or very high susceptibility zones. The agreement between predicted classes and observed discharge ranks was assessed qualitatively.

2.2.5. Sample Size, Model Complexity, and Model Stability

The limited number of sub-watersheds (n = 27) is a recognized constraint. DL architectures typically require larger datasets; however, our models were deliberately kept shallow (CNN with only two convolutional layers; MLP and DNN with three and four hidden layers) to minimize trainable parameters and prevent overfitting. Furthermore, models were trained on aggregated sub-watershed features (24 predictors) rather than raw spatial grids, drastically reducing effective dimensionality. While k-fold cross-validation is a standard check, applying it to spatial data can cause spatial autocorrelation leakage between folds. Thus, a 70/30 spatial split was used to ensure an independent test set, consistent with small-sample hydrological practices. Finally, all reported metrics are site-specific and should not be over-generalized without independent validation. Given the limited number of sub-watersheds (n = 27) and the 70/30 train–test split, the test set contains about eight sub-watersheds. At this sample size, formal significance tests would likely be underpowered and confidence intervals too wide to reliably distinguish between models. Therefore, robustness was assessed using a multi-metric consistency framework: model ranking was evaluated across MAE, MSE, RMSE, R^2, AUC, precision, recall, F1-score, balanced accuracy, and the confusion matrix.

The overall framework and sequential steps of the research methodology are illustrated in Figure 3.

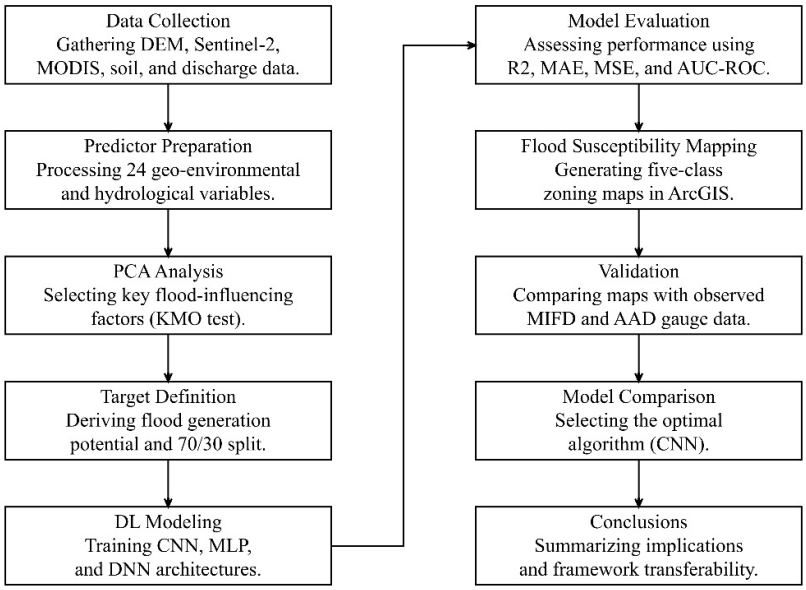


Figure 3. Flowchart illustrating the main research steps for deep learning-driven flood susceptibility mapping.

3. Results and Discussion

3.1. Spatial Characterization of Land Use Predictors Across Sub-Watersheds

The LU areas based on sub-watersheds within the Qareh-Sou watershed are provided in Table 3.

Table 3. The LU areas based on sub-watersheds in the Qareh-Sou watershed (in hectares), Iran.

SWs	Forest	Range	Orchard	Rainfed Agri.	Irrigated Agri.	Residential	Water Body
1	366.76	7212.45	101.48	10,387.7	472.51	116.66	-
2	170.58	2416.39	-	3486.67	52.69	18.36	-
3	276.78	2461.51	29.63	2360.89	233.92	245.2	-
4	335.66	1167.73	556.97	1146.59	1479.94	91.76	47.1
5	746.6	-	627.15	29.04	3560.99	60.03	133.16
6	5343.29	8757.82	450.5	-	10,864.9	689.61	63.93
7	442.5	5837.12	-	1324.39	5398.19	217.41	-
8	695.32	22,824.08	121.59	992.38	8179.2	189.65	1835.05
9	-	9867.29	133.16	2260.13	3340.34	104.49	-
10	-	15,630.5	195.8	4821.15	761.02	127.03	0.12
11	0.01	6925.57	163.05	9303.15	2321.99	237.05	0.44
12	-	56.26	207.42	12,591.1	7.8	149.13	-
13	813.61	7536.82	159.03	13,642.38	294.5	58.94	-
14	235.71	11,399.36	9.48	10,643.56	-	99.71	-
15	1313.57	7603.52	576.12	20,303.91	1704.77	752.97	260.52
16	-	1395.04	11.59	3221.64	890.54	26.65	-
17	1.46	3115.73	-	5055.9	10,514.9	232.17	-
18	1555.87	8798.05	-	-	12,128.2	312.08	-
19	-	-	404.73	57.65	6547.83	109.14	-
20	-	-	-	337.44	4910.49	989.91	-
21	-	-	339.52	1096.33	4867.68	1139.1	-
22	526.29	114.24	288.82	8000.08	22,530.17	1211.26	69.84
23	-	-	-	4749.87	3989.18	762.45	1013.64

24	-	6333.26	857.25	14,924.4	2558.68	189.56	7.58
25	-	29.78	102.9	7109.82	265.97	32.88	-
26	-	247.86	106.35	4035.03	-	40.05	-
27	-	39.18	54.88	11,152.45	1506.02	2140.19	199.78

Land use (LU) varied considerably across the 27 sub-watersheds (Table 3). Forest cover was generally limited; Sub-watershed 6 had the highest forest area (5343.29 ha, 20.4%), whereas Sub-watersheds 9, 10, 12, 17, 19, 20, 21, 23, 25, 26 and 27 had negligible or no forest. Rangelands dominated many sub-watersheds, particularly Sub-watershed 8 (22,824.08 ha, 65.6%) and Sub-watershed 10 (15,630.5 ha, 72.5%). Sub-watersheds 5 and 19 lacked rangelands entirely.

Agriculture was the other major LU. Rainfed farming dominated by absolute area in Sub-watersheds 11 (9303.15 ha, 49.0%) and 15 (20,303.91 ha, 62.3%), while the highest proportional rainfed cover occurred in Sub-watershed 12 (96.7%). Irrigated agriculture was concentrated in Sub-watersheds 22 (22,530.17 ha, 68.8%) and 17 (10,514.9 ha, 55.6%), with the highest proportional cover in Sub-watershed 19 (92.0%). Orchards were minor, mainly in Sub-watersheds 5 (12.1%) and 4 (11.5%). Residential areas were small but relatively prominent in Sub-watersheds 22 (3.7%) and 27 (14.2%). Water bodies were scarce, except in Sub-watersheds 8 (5.3%) and 23 (10.9%).

Overall, the LU pattern reflects a mix of natural ecosystems and human-modified landscapes. As noted by Baral et al. [35], land-cover transitions can involve complex, bi-directional swaps and persistence rather than linear and irreversible changes, which supports the cautious interpretation of the non-monotonic vegetation trends observed in this study. The dominance of rangelands and agriculture indicates strong dependency on agro-pastoral activities, while limited forest and water bodies suggest potential vulnerability to climate change and land degradation [36,37].

3.2. Analysis of PCA Results for SFM in the Qareh-Sou Watershed, Iran

Based on PCA, the final results for each component, based on the percentage of variance and the correlation coefficient, are shown in Table 4. Additionally, the correlation matrix illustrating the relationships between the final criteria influencing SFM is presented in Figure 4.

Table 4. PCA results for the selected criteria used in SFM, the Qareh-Sou watershed, Iran.

Components	Factor Loading	Selected Variable
1	0.70	Aspect
2	0.94	Plan curvature
3	0.93	Profile curvature
4	0.64	Distance from the river
5	0.61	Soil type
6	0.85	Area
7	0.94	Watershed slope
8	0.78	Compactness Coefficient
9	0.78	Main stream length
10	0.57	Maximum annual rainfall

The factor loadings of the retained variables are shown in Table 4, and all exceeded the retention threshold of 0.5. PCA was conducted to identify the most influential criteria in flood modeling. The overall KMO values confirmed the suitability of the dataset for PCA. Plan curvature (0.94), profile curvature (0.93), and watershed slope (0.94) exhibited the highest factor loadings, suggesting their strong contribution to explaining variance in

FSM. The Kaiser–Meyer–Olkin (KMO) values, which measure sampling adequacy, indicated that most of the selected variables were highly suitable for PCA, with values exceeding 0.6. Especially, plan curvature (0.94), profile curvature (0.93), and watershed slope (0.94) exhibited the highest KMO values, suggesting their strong contribution to explaining variance in FSM. These geomorphological factors likely play a critical role in controlling water flow direction, accumulation, and velocity, thus significantly influencing flood dynamics (Table 4) [38,39].

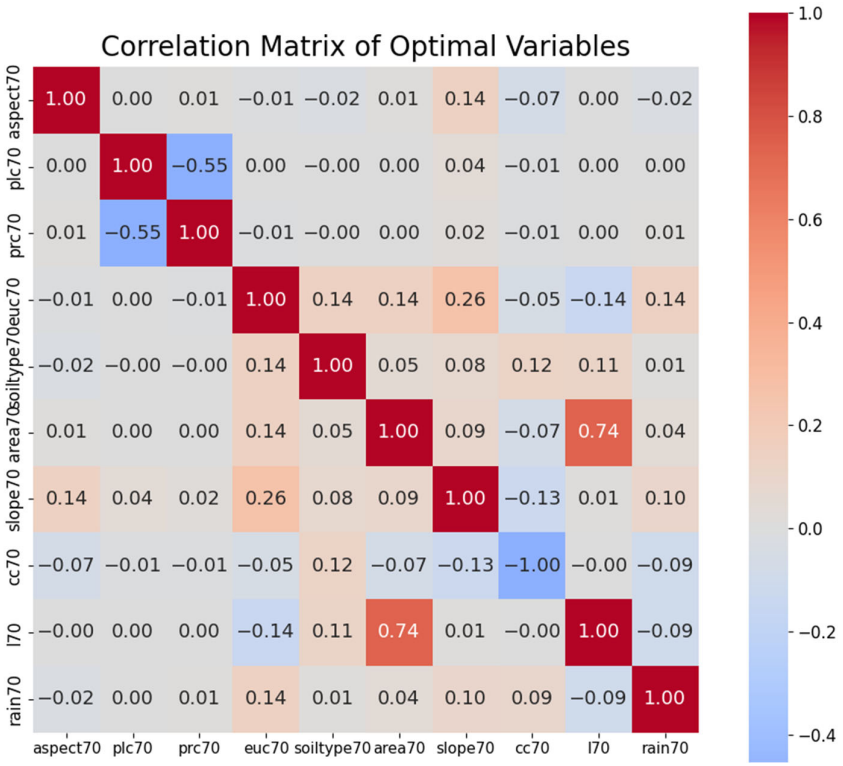


Figure 4. The correlation matrix related to the influential criteria affecting SFM. The numbered axes correspond to 1—Aspect, 2—Plan curvature, 3—Profile curvature, 4—Distance from the river, 5—Soil type, 6-Area, 7—Watershed slope, 8—Compactness Coefficient, 9—Main stream length, and 10—Maximum annual rainfall.

Variables such as compactness coefficient (0.78), main stream length (0.78), and area (0.85) also showed relatively high factor loadings, reflecting the importance of watershed shape and drainage network structure in flood generation [40–42]. In contrast, annual maximum rainfall (0.57) and soil type (0.61) had the lowest factor loadings among the retained variables, although both still exceeded the 0.5 retention threshold, implying that their explanatory power was secondary compared to topo-hydrological factors. The distance from the river (0.64) showed moderate adequacy, reinforcing the expected relationship between proximity to river channels and flood risk (Table 4).

The results indicate that watershed morphology (slope, curvature, compactness) and hydrological connectivity (stream length, drainage area) were the dominant controls on flood modeling [43] in the Qareh-Sou watershed, Iran. However, the lower influence of climatic (rainfall) and pedological (soil) factors suggests that in this region, flood generation was more structurally controlled by landscape features rather than precipitation variability or infiltration capacity.

This correlation matrix (Figure 4) revealed various relationships among the key variables influencing SFM. Especially, “distance from the river” and “soil type” showed a

strong positive correlation, indicating that these factors tended to increase together in the data. Conversely, "plan curvature" and "profile curvature" exhibited a significant negative correlation, suggesting that higher levels of one variable corresponded to lower levels of the other. The variables "aspect" and "slope" revealed a moderate positive correlation, implying some degree of association between terrain orientation and slope. Overall, the matrix highlighted both strong and weak relationships, which could help in understanding how different factors interacted to affect SFM.

3.3. Comparative Evaluation of DL Algorithms for SFM

The performance of the DL algorithms was evaluated using two sets of metrics: (1) regression metrics (MAE, MSE, and R^2) to assess the accuracy of the continuous FGP predictions, and (2) classification metrics (AUC-ROC) to evaluate the ability of the models to discriminate flood-prone areas based on the 0.5 threshold. The results are presented in Table 5.

Table 5. Regression and classification performance metrics for each DL algorithm used in SFM.

Algorithm	Regression Metrics			Classification Metric
	MAE	MSE	R^2	AUC
CNN	0.19	0.11	0.87	0.93
MLP	0.25	0.18	0.72	0.88
DNN	0.55	0.32	0.64	0.86

The results revealed notable differences in how well the three deep learning algorithms predicted flood susceptibility. CNN demonstrated superior performance, but MLP and DNN exhibited lower accuracy. The CNN model revealed superior performance across all error metrics, achieving a MAE of 0.19, MSE of 0.11, and an R^2 value of 0.87, indicating that 87% of the variance in SFM was explained by the model. Furthermore, its AUC score of 0.93 confirmed outstanding discrimination capacity between flood-prone and non-flood-prone areas (Table 5).

In contrast, the MLP model exhibited substantially higher errors, with an MAE of 0.25 and MSE of 0.18, suggesting limitations in handling spatial patterns or non-linear relationships within the dataset. While its R^2 (0.72) and AUC (0.88) remained acceptable, the inflated MAE implies systematic biases in magnitude prediction. The DNN model exhibited the weakest performance, with the highest MAE (0.55) and MSE (0.32) and the lowest R^2 (0.64) and AUC (0.86) among the three models. This indicates that while DNNs captured general trends, their precision was inferior to both CNNs and MLPs, likely due to less effective feature extraction from geospatial inputs (Figure 5). These results are consistent with previous studies. For example, Dawson et al. [14] reported an AUC of 0.93 for XGBoost in flood susceptibility mapping in England, which is comparable to the CNN performance in this study. Youssef et al. [44] also found that CNN outperformed SVM in flood susceptibility mapping in Egypt. The superior performance of CNN can be attributed to its ability to capture spatial dependencies and local patterns in geospatial data, which are critical for flood modeling. The confusion matrix for the CNN model showed that the majority of misclassifications occurred between adjacent classes (e.g., high vs. very high), which is expected in continuous susceptibility mapping. The CNN's dominance can be attributed to its architectural advantage in processing spatial data (e.g., terrain curvature, slope), where convolutional layers efficiently extract localized flood-influencing features [44,45]. The higher MAE of the MLP model is likely attributable to its fully connected structure, which struggles with high-dimensional geospatial correlations. These results advocate for prioritizing CNN-based approaches in SFM applications, particularly where topographic and hydrological spatial dependencies are critical.

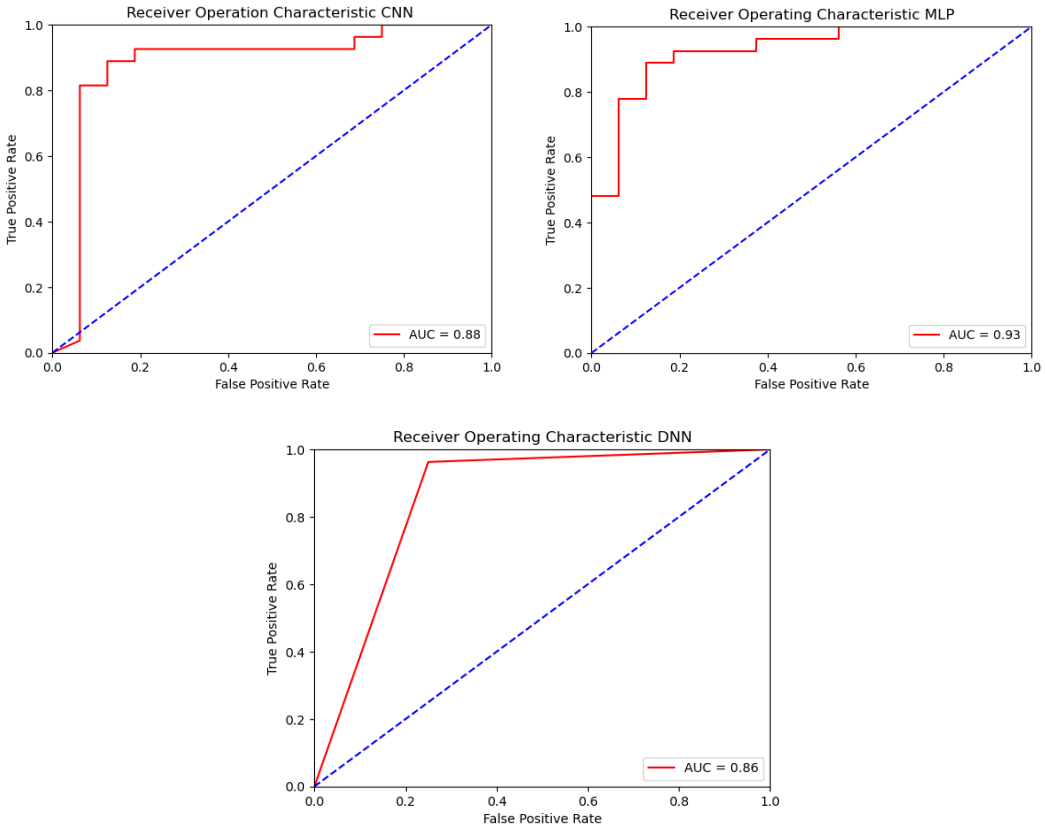


Figure 5. The ROC curves for each of the DL algorithms.

This aligns with recent DL flood studies (e.g., Youssef et al. [44]), which show CNNs outperform fully connected networks in complex terrain. CNNs excel when flood drivers show strong spatial autocorrelation, localized topographic features, and hydrologically meaningful predictor arrangement, as in the rugged, snow-fed Qareh-Sou watershed. MLP/DNN may match or outperform them in flat, homogeneous, or heavily urbanized basins driven by non-spatial factors, or with purely tabular inputs. Qareh-Sou's topographic heterogeneity thus explains the CNN's advantage in capturing localized flood-influencing features. Given the small test set, formal significance testing was not used as the primary basis for model comparison; instead, robustness was assessed through cross-metric consistency. CNN ranked best on every regression metric (lowest MAE and MSE, highest R^2) and on AUC. This consistent ranking across independent error measures provides supporting evidence against overfitting or test-split artifacts. The CNN confusion matrix showed that misclassifications occurred mainly between adjacent classes (e.g., high vs. very high), which is consistent with the model learning the ordinal structure of the target rather than memorizing training labels. Independent MIFD/AAD validation further supports the presence of genuine hydrological signal. Future studies with larger samples should apply k-fold cross-validation with paired tests.

3.4. Spatial Distribution and Quantitative Analysis of SFM Zones Based on CNN Algorithm

The flood generation potential zoning map is provided in Figure 6. Additionally, the area of flood-prone classes based on the CNN algorithm within the Qareh-Sou watershed, Iran is presented in Table 6.

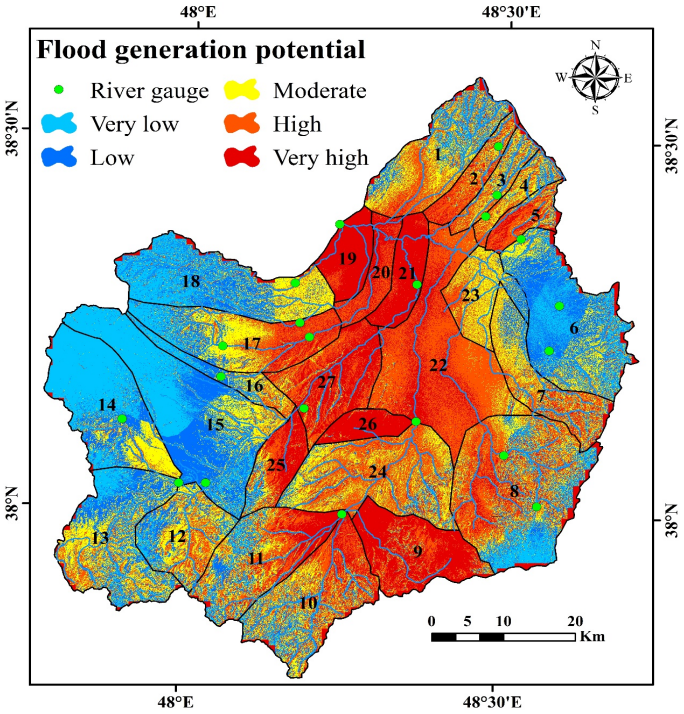


Figure 6. The SFM zoning map based on the CNN algorithm in the Qareh-Sou watershed, Iran.

Table 6. The area of flood-prone classes, CNN algorithm, the Qareh-Sou watershed, Iran.

SFM Classes	Area (Hectares)	Area (%)
Very low	43,669.04	10.31
Low	53,410.92	12.61
Moderate	62,517.46	14.76
High	141,680.82	33.45
Very high	122,324.13	28.88

The flood risk assessment for the Qareh-Sou watershed, based on the CNN algorithm, revealed significant spatial variation in flood potential zones. Approximately 10.31% of the area was classified as having a very low flood potential, predominantly located in the southwestern and southeastern parts of the watershed, where the terrain is relatively flatter and hydrologically less active. The low flood potential zone accounted for 12.61% of the total area and was mainly distributed in the peripheral regions along the watershed's edges, where elevation and slope were moderate to low. Radwan et al. [46] also concluded in their flood risk assessment study that flood risk categories including very high, high, moderate, low, and very low included 17, 41, 33, 8, and 1 percent of the watershed, respectively.

The moderate flood risk zone comprised about 14% of the watershed, mainly situated in regions with intermediate topographical features, serving as transitional zones between lower and higher risk areas. These areas often exhibited moderate slope and elevation, facilitating limited flood propagation under extreme conditions. High flood potential zones covered approximately 33% of the watershed, primarily located in the central and northern parts, where the topography is more rugged and the flow accumulation is higher, thus increasing SFM [47].

Approximately 28% of the area was classified as very high SFM, predominantly found in the northern and northeastern sections of the watershed.

In terms of geographic considerations, the upper watershed with steep slopes and a dense hydrographic network exhibited greater SFM, aligning with the high percentage of high and very high flood potential zones. Conversely, the lower plains with flatter terrain showed lower flood risk percentages. Hydrologically, regions with higher flow accumulation and steeper slopes correlated with increased flood potential, confirming the spatial distribution patterns observed in the map.

Table 6 summarizes the area distribution among flood risk categories, showing that the largest proportion was classified as high risk (33.45%), followed by very high (28.88%), moderate (14.76%), low (12.61%), and very low (10.31%).

3.5. Validation of FS Modeling Map with Observational Data

The quantitative values of the maximum instantaneous flood discharge and the annual average flow at the outlet of each sub-watershed, based on gauging stations and observational data, are presented in Table 7.

Table 7. The values of the maximum instantaneous flood discharge and the annual average flow at each gauging station, based on observational data for the Qareh-Sou watershed, Iran.

Station Name	X (°)	Y (°)	Elevation (m)	MIFD (m^3/s)	AAD (m^3/s)
Pol-e-Almas	48.19	38.14	1440.00	29.76	3.29
Gilandeh	48.36	38.31	1332.00	23.79	2.93
Koozeh Topraqi	48.37	38.12	1394.00	17.81	1.07
Yamchi	48.04	38.03	1584.00	14.32	2.07
Samian	48.25	38.38	1290.00	13.49	6.11
Hir	48.51	38.08	1575.00	9.35	0.31
Ahl-e-Iman	48.47	38.40	1107.00	8.73	0.23
Shams Abad	48.25	38.00	1493.00	7.89	0.32
Nir	47.99	38.03	1624.00	6.22	1.22
Atashgah	48.06	38.22	1773.00	5.55	0.11
Aladizgeh	48.59	38.28	1347.00	5.49	0.21
Viladaraq	48.06	38.18	1800.00	5.47	0.07
Nooran	48.19	38.23	1423.00	4.90	0.31
Amooghin	48.18	38.25	1385.00	4.89	0.20
Aeril	48.58	38.22	1375.00	4.75	0.28
Lay	47.90	38.12	2068.00	4.05	0.12
Baroogh	48.17	38.30	1410.00	3.97	0.20
Anbaran	48.49	38.49	1550.00	3.43	0.06
Namin	48.49	38.43	1459.00	2.06	0.10
Naneh Karan	48.53	38.37	1350.00	1.59	0.08
Neor dam	48.56	38.01	2500.00	1.08	0.18

We validated the flood susceptibility map (which came from the CNN model) by lining it up against actual discharge measurements from gauge stations across the Qareh-Sou watershed. The map classified areas into different flood potential zones, ranging from very low to very high, using the CNN-based modeling approach. To assess the accuracy of this classification, the maximum instantaneous discharges and the annual mean flows obtained from the observational data at various stations (such as Pol-e-Almas, Gilandeh, Kozeh-Topraqi, Yamchi, and others) were considered. The stations exhibited a wide range of flow magnitudes, with maximum instantaneous discharges spanning from 1.08 m^3/s at the Neor Dam outlet to 29.76 m^3/s at Pol-e-Almas, and average annual flows ranging from 0.06 m^3/s to 6.11 m^3/s at different locations (Table 7).

The analysis revealed that the regions identified as having a high or very high flood potential in the classified map corresponded well with the stations exhibiting higher maximum discharges, such as Pol-e-Almas and Samian river gauge stations. Conversely, areas zoned as medium or low flood potential aligned with stations showing lower flow values, like Neor outlet and Naneh-Karan. The results show a clear consistency: the CNN-based map matches what we measured in the watershed. This supports the reliability of the model's predictions (Figure 6). However, some discrepancies were noted in zones where observed flow data did not precisely match the expected categorization, possibly due to local geographic or hydrological variations not fully captured by the model. Overall, the validation confirmed the reliability of the model in predicting flood potential zones within the Qareh-Sou watershed, Iran, supporting its applicability for flood risk management and planning. Beyond the descriptive comparison with gauge data, the spatial pattern of the CNN-derived susceptibility map strongly aligns with field observations, historical flood damage zones, and comprehensive local knowledge of the region. Specifically, the high and very high-risk zones identified in the central and northern sub-watersheds are highly consistent with the spatial distribution of iso-severity flood patterns reported for Ardabil Province [3]. Their study confirmed that flood frequency is significantly higher in the central, northern, and highland regions of this province. Although point-based gauge records may show minor variations due to local micro-topography or dam regulation at the sub-watershed outlets, the overall spatial agreement with both field evidence and published regional flood patterns confirms the practical reliability of the CNN model.

The main limitation of this study is the absence of a spatially explicit, high-resolution historical flood inventory. Consequently, the model's target variable (FGP) was derived from PCA-weighted predictors rather than direct hydrodynamic depths, introducing a structural dependency and inherent uncertainty into the training process. Furthermore, the model relies on a static land use map and static climatic inputs, lacking time-variant meteorological forcing and historical land use change scenarios. Additionally, extending point-based gauge discharge data to continuous spatial maps introduces a spatial scale mismatch, and the limited number of sub-watersheds (n = 27) restricts the generalizability of the DL models. Despite these limitations (which are common in data-scarce regions), the model's independent validation against observed MIFD and AAD records confirms its reliability. Future studies should address these uncertainties by integrating 2D hydrodynamic models, high-resolution remote sensing, and larger independent flood inventories. Specifically, future work should incorporate time-series meteorological data, multitemporal land use maps, and independent historical flood event locations (e.g., field surveys or SAR data) to capture dynamic flood processes and eliminate dependency on synthetic labels**.** Sharing predictors between FGP and model inputs introduces a structural dependency that may partially inflate R^2 and AUC. Since FGP is derived from the same 24 predictors used as inputs, the models partly approximate a function of their own inputs, so the reported $R^2 = 0.87$ and AUC = 0.93 should be interpreted with caution. The R^2 value also indicates that 13% of the variance remains unexplained, and AUC may be affected because the 0.5 threshold partitions the same predictor space used in training. Nevertheless, the DL models can capture non-linear relationships beyond linear PCA weighting, and independent MIFD/AAD validation supports the presence of genuine hydrological signal. Future studies using independent flood inventories should further isolate pure predictive skill and clarify the influence of this circularity. Furthermore, the single 70/30 train–test split, dictated by the small sample size, provides limited evidence of absolute model robustness. However, the consistent ranking of models (CNN > MLP > DNN) across all error metrics suggests that the comparative conclusions remain stable. Future studies with larger datasets should apply k-fold cross-validation to further test model stability.

4. Conclusions

This study applied Deep Learning (DL) algorithms (Convolutional Neural Networks (CNN), Multilayer Perceptron (MLP), and Deep Neural Networks (DNN)) to model and map Spatial Flood Modeling (SFM) in northwest Iran. The results showed that CNN outperformed MLP and DNN in accurately predicting flood-prone areas, leveraging its ability to process spatial data effectively. The integration of geo-environmental and hydrological variables, refined through Principal Component Analysis (PCA), provided a robust framework for flood risk assessment. The resulting flood susceptibility maps showed spatial variations in flood potential, with high-risk zones concentrated in areas characterized by steep slopes, dense drainage networks, and specific land use patterns. The key results of this study can be summarized as follows: (1) CNN outperformed MLP and DNN, achieving $R^2 = 0.87$, MAE = 0.19, and AUC = 0.93; (2) approximately 28% of the watershed was classified as very high risk and 33% as high risk; (3) watershed morphology, particularly slope and drainage density, was the dominant control on flood susceptibility; and (4) validation against gauge-based discharge data confirmed the reliability of the CNN-based map.

The research also emphasized the critical role of watershed morphology and land use distribution in influencing flood dynamics. The results indicate that most of the Qareh-Sou watershed is covered by rangeland and farmland, with limited forest cover. This pattern may increase land degradation and runoff intensity. The validation of model outputs against observational discharge data confirmed the reliability of the CNN-based approach, supporting its utility for decision-making in flood mitigation and land use planning. Future studies could further refine these models by incorporating real-time climatic data and anthropogenic factors to improve predictive accuracy. It is important to emphasize that the model outcomes, predictor weights, and spatial susceptibility classes are specifically calibrated for the snow-fed Qareh-Sou watershed. While the methodological framework provides a reference for other regions, transferring these calibrated results to different watersheds would require rigorous model recalibration and independent validation. Incorporating independent flood inventory data, real-time climatic inputs, and anthropogenic factors would further improve the predictive accuracy and practical applicability of the proposed framework.

Author Contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by the authors. The first draft of the manuscript was written by R.M. A.N.K., and S.M. Authors R.M. A.N.K., and S.M., commented on previous versions of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work is based upon research funded by Iran National Science Foundation (INSF) under project No. 4037755.

Data Availability Statement: Some of the data generated or analyzed during this study are included in this published article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cvetković, V.M.; Renner, R.; Aleksova, B.; Lukić, T. Geospatial and Temporal Patterns of Natural and Man-Made (Technological) Disasters (1900–2024): Insights from Different Socio-Economic and Demographic Perspectives. *Appl. Sci.* **2024**, *14*, 8129.
2. Khiavi, A.N.; Vafakhah, M.; Sadeghi, S.H. Comparative prioritization of sub-watersheds based on Flood Generation potential using physical, hydrological and co-managerial approaches. *Water Resour. Manag.* **2022**, *36*, 1897–1917.

3. Mostafazadeh, R.; Nasiri Khiavi, A.; Talebi Khiavi, H. Iso-severity flood hazard patterns in diverse relief and climatic regions to support flood management. *Nat. Hazards* **2025**, *121*, 17793–17814.
4. Azizi, E.; Mostafazadeh, R.; Hazbavi, Z.; Esmali-Ouri, A.; Mirzaei, S. Geospatial flood risk assessment: Multi-criteria approach integrating hydro-climatic hazard, exposure, vulnerability. *Geomat. Nat. Hazards Risk* **2025**, *16*, 2551814.
5. Mohammadi, N.; Mostafazadeh, R. Intensification of Extreme and Compound Hazards in Urban Areas Under Climate Change in Iran: A Scoping Review. *Climate* **2026**, *14*, 126.
6. Kumar, V.; Sharma, K.V.; Caloiero, T.; Mehta, D.J.; Singh, K. Comprehensive overview of flood modeling approaches: A review of recent advances. *Hydrology* **2023**, *10*, 141.
7. Tahmasebi, P.; Kamrava, S.; Bai, T.; Sahimi, M. Machine learning in geo- and environmental sciences: From small to large scale. *Adv. Water Resour.* **2020**, *142*, 103619. https://doi.org/10.1016/j.advwatres.2020.103619.
8. Nasiri Khiavi, A.; Vafakhah, M. Using algorithmic game theory to improve supervised machine learning: A novel applicability approach in flood susceptibility mapping. *Environ. Sci. Pollut. Res.* **2024**, *31*, 52740–52757.
9. Ruidas, D.; Pal, S.C.; Saha, A.; Roy, P.; Pande, C.B.; Islam, A.R.M.T.; Islam, A. Flood hazard forecasting and management systems: A review of state-of-the-art modelling, management strategies and policy-practice gap. *Int. J. Disaster Risk Reduct.* **2024**, *108*, 104539.
10. Jeba, G.S.; Chitra, P. Exploring the Power of Deep Learning and Big Data in Flood Forecasting: State-of-the-Art Techniques and Insights. In *Intelligent Systems and Sustainable Computational Models*; Auerbach Publications: Boca Raton, FL, USA, 2024; pp. 14–33.
11. Shang, X.; Jiang, X.; Jia, R.; Wei, C. Land use and climate change effects on surface runoff variations in the upper Heihe River basin. *Water* **2019**, *11*, 344.
12. Hu, S.; Shrestha, P. Examine the impact of land use and land cover changes on peak discharges of a watershed in the midwestern United States using the HEC-HMS model. *Pap. Appl. Geogr.* **2020**, *6*, 101–118.
13. Avand, M.; Khiavi, A.N.; Khazaei, M.; Tiefenbacher, J.P. Determination of flood probability and prioritization of sub-watersheds: A comparison of game theory to machine learning. *J. Environ. Manag.* **2021**, *295*, 113040.
14. Dawson, G.; Butt, J.; Jones, A.; Fraccaro, P. Flood susceptibility mapping at the country scale using machine learning approaches. *Appl. AI Lett.* **2024**, *4*, e88.
15. Al-Sheriadeh, M.S.; Daqdouq, M.A. Robustness of machine learning algorithms to generate flood susceptibility maps for watersheds in Jordan. *Geomat. Nat. Hazards Risk* **2024**, *15*, 2378991.
16. Nasiri Khiavi, A.; Sadeghi, S.H.; Vafakhah, M. Comparative Prioritization of Sub-Watersheds in Flood Generation Using Co-Management Best-Worst Method and Game Theory Algorithm. *Water Resour. Manag.* **2024**, *38*, 4431–4453.
17. Naserabadi, F.; Esmali Ouri, A.; Akbari, H.; Rostamian, R. A sensitivity analysis of SWAT model in Ghareh Su watershed, Ardabil. *Watershed Eng. Manag.* **2014**, *5*, 255–265.
18. Yari, M.; Soltani-Gerdefaramarzi, S.; Ghasemi, M. Investigation of the Effect of to Land Use Changes on Flood Hydrograph and Fluctuations of Groundwater Level in Part of Ghareh Souh Basin. *J. Geogr. Environ. Hazards* **2019**, *8*, 41–58.
19. Nasiri Khiavi, A.; Vafakhah, M.; Sadeghi, S.H. Flood-based critical sub-watershed mapping: Comparative application of multi-criteria decision making methods and hydrological modeling approach. *Stoch. Environ. Res. Risk Assess.* **2023**, *37*, 2757–2775. https://doi.org/10.1007/s00477-023-02417-0.
20. Li, A.; Song, K.; Chen, S.; Mu, Y.; Xu, Z.; Zeng, Q. Mapping African wetlands for 2020 using multiple spectral, geo-ecological features and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2022**, *193*, 252–268.
21. Liu, L.; Zhang, Q.; Guo, Y.; Li, Y.; Wang, B.; Chen, E.; Li, Z.; Hao, S. Mapping Coniferous Forest Distribution in a Semi-Arid Area Based on Multi-Classifier Fusion and Google Earth Engine Combining Gaofen-1 and Sentinel-1 Data: A Case Study in Northwestern Liaoning, China. *Forests* **2024**, *15*, 288.
22. Munizaga, J.; Garc'ia, M.; Ureta, F.; Novoa, V.; Rojas, O.; Rojas, C. Mapping coastal wetlands using satellite imagery and machine learning in a highly urbanized landscape. *Sustainability* **2022**, *14*, 5700.
23. Mattivi, P.; Franci, F.; Lambertini, A.; Bitelli, G. TWI computation: A comparison of different open source GISs. *Open Geospat. Data Softw. Stand.* **2019**, *4*, 6.
24. Ebrahimy, H.; Aghighi, H.; Azadbakht, M.; Amani, M.; Mahdavi, S.; Matkan, A.A. Downscaling MODIS land surface temperature product using an adaptive random forest regression method and Google Earth Engine for a 19-years spatiotemporal trend analysis over Iran. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 2103–2112.
25. Kibret, K.S.; Marohn, C.; Cadisch, G. Use of MODIS EVI to map crop phenology, identify cropping systems, detect land use change and drought risk in Ethiopia—An application of Google Earth Engine. *Eur. J. Remote Sens.* **2020**, *53*, 176–191.

26. Bro, R.; Smilde, A.K. Principal component analysis. *Anal. Methods* **2014**, *6*, 2812–2831.
27. Bucherie, A.; Hultquist, C.; Adamo, S.; Neely, C.; Ayala, F.; Bazo, J.; Kruczkiewicz, A. A comparison of social vulnerability indices specific to flooding in Ecuador: Principal component analysis (PCA) and expert knowledge. *Int. J. Disaster Risk Reduct.* **2022**, *73*, 102897. https://doi.org/10.1016/j.ijdrr.2022.102897.
28. Foody, G.M.; Ghoneim, E.M.; Arnell, N.W. Predicting locations sensitive to flash flooding in an arid environment. *J. Hydrol.* **2004**, *292*, 48–58.
29. Adhami, M.; Sadeghi, S.H. Sub-watershed prioritization based on sediment yield using game theory. *J. Hydrol.* **2016**, *541*, 977–987.
30. Abdrabo, K.I.; Kantoush, S.A.; Esmaiel, A.; Saber, M.; Sumi, T.; Almamari, M.; Elboshy, B.; Ghoniem, S. An integrated indicator-based approach for constructing an urban flood vulnerability index as an urban decision-making tool using the PCA and AHP techniques: A case study of Alexandria, Egypt. *Urban Clim.* **2023**, *48*, 101426.
31. Martins, B.; Nunes, A. Exploring flash flood risk perception using PCA analysis: The case of Mindelo, S. Vicente (Cape Verde). *Geogr. J.* **2020**, *186*, 375–389.
32. Chimakurthi, V. Application of convolution neural network for digital image processing. *Eng. Int.* **2020**, *8*, 149–158.
33. Asghari Saraskanroud, S.; Riahinia, M.; Zeinali, B.; Mostafazadeh, R. Comparative assessing the performance of ANN, RF and CNN machine learning methods in identifying landslide prone areas. *Res. Earth Sci.* **2025**, *16*, 73–91.
34. Kim, S.; Wimmer, H.; Kim, J. Analysis of deep learning libraries: Keras, pytorch, and MXnet. In Proceedings of the 2022 IEEE/ACIS 20th International Conference on Software Engineering Research, Management and Applications (SERA), Las Vegas, NV, USA, 25–27 May 2022; pp. 54–62.
35. Baral, P.; Wen, Y.; Urriola, N.N. Forest cover changes and trajectories in a typical middle mountain watershed of western Nepal. *Land* **2018**, *7*, 72.
36. Pricope, N.G.; Husak, G.; Lopez-Carr, D.; Funk, C.; Michaelsen, J. The climate-population nexus in the East African Horn: Emerging degradation trends in rangeland and pastoral livelihood zones. *Glob. Environ. Change* **2013**, *23*, 1525–1541.
37. Tsegaye, A. Monitoring Rangeland Degradation and its Dynamism in Pastoral Areas of Afar Region, Ethiopia by using Earth Observation Data. *Int. J. Hortic. Agric. Food Sci.* **2024**, *8*, 593086.
38. Barbosa, R.A.; Spletozer, A.G.; da Silveira, L.J.; Lorenzon, A.S.; Tonello, K.C.; Barbosa, S.G.; Dias, H.C.T. Morphometric characterization as a tool to control flood and landslide risks in a portion of the Manhuaçu River Basin, MG, Brazil. *Rev. Bras. De Ciências Agrárias* **2022**, *17*, 1–9.
39. Liu, J.; Luo, M.; Chen, Z.; Gou, J.; Yan, Z. The effects of terrain factors on the drainage area threshold: Comparison of principal component analysis and correlation analysis. *Environ. Monit. Assess.* **2022**, *194*, 168.
40. Chethan, B.J.; Vishnu, B. Influence of Geomorphology on Runoff Characteristics of a Catchment. Master's Thesis, Department of Irrigation and Drainage Engineering, Kelappaji College of Agricultural Engineering and Technology, Tavanur, India, 2018.
41. Rana, V.K.; Suryanarayana, T.M.V. Estimation of flood influencing characteristics of watershed and their impact on flooding in data-scarce region. *Ann. GIS* **2021**, *27*, 397–418.
42. Roy, S.; Das, S.; Sengupta, S. Predicting terrain erosion susceptibility from drainage basin morphometry using ALOS-PALSAR DEM: Analysis from PCA-weighted AHP approach in a river system of Eastern India. *Environ. Dev. Sustain.* **2023**, *25*, 9589–9617.
43. Tola, S.Y.; Shetty, A. Flood susceptibility modeling based on morphometric parameters in Upper Awash River basin, Ethiopia using geospatial techniques. *Sustain. Water Resour. Manag.* **2022**, *8*, 49.
44. Youssef, A.M.; Pradhan, B.; Dikshit, A.; Mahdi, A.M. Comparative study of convolutional neural network (CNN) and support vector machine (SVM) for flood susceptibility mapping: A case study at Ras Gharib, Red Sea, Egypt. *Geocarto Int.* **2022**, *37*, 11088–11115.
45. Kabir, S.; Patidar, S.; Xia, X.; Liang, Q.; Neal, J.; Pender, G. A deep convolutional neural network model for rapid prediction of fluvial flood inundation. *J. Hydrol.* **2020**, *590*, 125481.
46. Radwan, F.; Alazba, A.A.; Mossad, A. Flood risk assessment and mapping using AHP in arid and semiarid regions. *Acta Geophys.* **2019**, *67*, 215–229.
47. Ogato, G.S.; Bantider, A.; Abebe, K.; Geneletti, D. Geographic information system (GIS)-Based multicriteria analysis of flooding hazard and risk in Ambo Town and its watershed, West shoa zone, oromia regional State, Ethiopia. *J. Hydrol. Reg. Stud.* **2020**, *27*, 100659.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

https://doi.org/10.3390/earth7060195