EARTH SCIENCES
RESEARCH JOURNAL
Earth Sci. Res. J. Vol. 29, No. 3 (September, 2025): 261 - 273

UNIVERSIDAD NACIONAL DE COLOMBIA

GEOLOGICAL ENGINEERING

Susceptibility Assessment of Deep-seated Landslides in Sub-Himalayan Galiat Region, Pakistan

Qaiser Mehmood, Tolga Çan*
Department of Geological Engineering, Cukurova University, Adana, Turkey
qmehmood@student.cu.edu.tr
* Corresponding author: tolgacan@cu.edu.tr

ABSTRACT

Landslides are the most prevalent natural hazards in the Sub-Himalayan region, posing extensive socio-economic losses. Their occurrence is highly influenced by weak geological formations, steep and dissected topography, irregular land-use, high seismic activity, and seasonal precipitation and snowmelt. Despite the high threat, there is an absence of landslide susceptibility maps for most of northern Pakistan, hindering effective measures for landslide hazard prevention. In this study, a relevant deep-seated landslide inventory for landslide susceptibility assessment of the Galiat Region was prepared based on field studies and multi-temporal Google Earth images, identifying 68 landslide polygons. Due to the localized nature of landslides, substantial predictions cannot be made with classical statistical modelling. Therefore, the landslide susceptibility map of the study area was modelled using the maximum entropy method, which allows predictions based on limited observational data. The analyses were repeated, with three randomly selected data sets being 30% and 70% for training and testing data, respectively. Fourteen environmental variables were considered, including geology, digital elevation model (DEM), and first and second DEM derivatives. The accuracy of the obtained models reached 0.80 ±0.002, evaluated by the AUC technique. The high to very high susceptible classes correspond to 26.16 % of the study area, including 74.3 % of the mapped landslides. The resultant landslide susceptibility map will raise understanding of dynamic and potential landslides for citizens, engineers, and land-use agencies.

Keywords: Landslide inventory; deep-seated landslide; landslide susceptibility map; maximum entropy model.

Evaluación de la susceptibilidad de d**eslizamientos de tierra profundos en la región Galiat del sub-Himalaya, Pakistán**

RESUMEN

Los deslizamientos de tierra son la amenaza natural más prevalente en la región del sub-Himalaya, y generan grandes perdidas socioeconómicas. Su ocurrencia está fuertemente influenciada por las débiles formaciones geológicas, la topografía inclinada y diseccionada, el uso irregular del suelo, la alta actividad sísmica y las lluvias estacionarias y el descongelamiento de nieve. A pesar de esta amenaza hacen falta mapas de susceptibilidad de deslizamientos de tierra para gran parte del norte de Pakistán, lo que dificulta la implementación efectiva de medidas para la mitigacion de estas amenazas. Para este estudio se preparó un inventario de evaluación de la susceptibilidad de deslizamientos de tierra profundos en la región de Galiat con base en estudios de campo y de imágenes multitemporales de Google Earth, donde se identificaron 68 polígonos de deslizamiento. Debido a la naturaleza localizada de los deslizamientos no se pudieron hacer predicciones sustanciales con los modelos estadísticos clásicos. Además, el mapa de susceptibilidad de deslizamientos de tierra del área de estudio se modeló con el método de máxima entropía, lo que permite predicciones con base en datos observacionales limitados. Los análisis se repitieron con tres secuencias de datos seleccionadas aleatoriamente, cada una con el 30 por ciento de entrenamiento y el 70 por ciento de prueba. Se consideraron 14 variables ambientales que incluyen geología, modelo de elevación digital, y la primera y segunda derivada de los modelos de elevación digital. La exactitud de los modelos obtenidos alcanzó 0.80 ± 0.002, según se evaluó con la técnica AUC. La clasificación de alta a muy alta susceptibilidad corresponde al 26.16 % del área de estudio, y que incluye el 74.3 % de los deslizamientos mapeados. El mapa de susceptibilidad de deslizamientos de tierra resultante incrementará el conocimiento de las dinámicas y potenciales deslizamientos para ciudadanos, ingenieros y agencias de uso del suelo.

Palabras clave: inventario de deslizamientos de tierra; deslizamientos profundos; mapa de deslizamientos de tierra; modelo de máxima entropía.

Record
Manuscript received: 05/08/2024
Accepted for publication: 25/08/2025

How to cite this item:
Mehmood, Q., & Can, T. (2025). Susceptibility Assessment of Deep-seated Landslides in Sub-Himalayan Galiat Region, Pakistan. *Earth Sciences Research Journal, 25*(3), 261-273. https://doi.org/10.15446/esrj.v29n3.116114

ISSN 1794-6190 e-ISSN 2339-3459
https://doi.org/10.15446/esrj.v29n3.116114

1. Introduction:

Deep-seated landslides generally occur at slopes between 5° and 30° (some examples in steep slopes) with a comparatively deeper sliding surface, lower velocity, less disturbance during movement, and an arc-shaped sliding surface. To effectively manage landslide hazards, it is necessary to comprehend the primary situations and governing mechanisms of existing landslides in a region (Mehmood et al., 2024). A precise landslide inventory dataset is a fundamental prerequisite for quantitative landslide susceptibility, hazard, and risk modeling. The preparation of landslide inventory maps is essential for documenting the scope of landslide events in a region as well as evaluating the dissemination, classes, structures, recurring, and statistics of slope failures (Guzzetti et al., 2012; Cömert et al., 2018; Rosi et al., 2023). The accuracy and completeness of the maps are crucial for dependable applications and are mostly controlled by the extent of the study area, available resources, and utilized mapping techniques (Galli et al., 2008; Van Westen et al., 2008; Guzzetti et al., 2012).

The development of landslide susceptibility assessment maps is a complex process that requires a significant amount of data collection and analysis. Recent improvements in geographic information systems (GIS) and remote sensing equipment have made it easier to collect and analyse data over large areas (Chen et al., 2016). consequently, the accuracy and precision of landslide susceptibility assessment maps have improved. Landslide susceptibility assessments rely primarily on GIS frameworks and diverse mathematical approaches to analyse the correlation between various parameters and the likelihood of landslide occurrence (Chen and Wang 2007; Chen et al., 2016). Several methods have achieved positive outcomes in evaluating landslide susceptibility, such as logistic regression (Conoscenti et al., 2015; Duman and Çan 2023), frequency ratio (Sun et al., 2018; Xiao et al., 2020), neural network (Zhang et al., 2019), information value (Melo et al., 2012), index entropy (Wu et al., 2020), support vector machines (Pourghasemi and Rossi 2017), analytic hierarchy process (Cao et al., 2017; Mehmood et al., 2021), certainty factor (Tsangaratos and Ilia 2016; Xiao et al., 2019), and random forest (Xiao et al., 2020). Numerous variables are essential to be considered, including land cover, hydrology, geology, and geomorphology (Lee, 2005).

Landslides are among the most prevalent natural hazards in the Sub-Himalayan Range of Pakistan, causing substantial losses of life and property. The region is characterized by inept geological units, steep topography, erratic land-use activities, high seismicity, and excessive rainfall, all of which are major landslide triggering and preparatory factors. Despite the high threat, most of the northern regions of Pakistan lack landslide susceptibility maps, which are essential for hazard mitigation. To address this gap, a deep-seated landslide susceptibility assessment was carried out in the Galiat region of the Sub-Himalayan Range. A comprehensive landslide inventory map was developed by integrating field surveys with multi-temporal Google Earth imagery, and landslide susceptibility was evaluated using 68 mapped landslide polygons.

Since deep-seated landslides usually occur in small and limited areas, classical statistical modeling cannot provide significant predictions. Therefore, we modelled the landslide susceptibility map using the maximum entropy method, which allows predictions from limited observation data. Maximum entropy is a statistical-probabilistic machine learning method that has been used in a variety of disciplines, including ecology and geological and environmental sciences, particularly the estimation of species distribution (Phillips et al., 2004; Elith et al., 2011; Mert et al., 2016). Recently, several researchers have used the maximum entropy method to study the susceptibility of landslides (Vorpahl et al., 2012; Convertino et al., 2013; Kim et al., 2015; Park 2015; Chen et al., 2017; Pandey et al., 2020). Despite its complex mathematics, this method doesn't require high precision or extensive surveys. In this method, the association between the observed landslide areas and the environmental variables is related to the optimum probability density. Fourteen environmental variables, including geology, 12.5m spatial resolution digital elevation model (DEM), and first and second DEM derivative parameters, have been used in this study.

The main objectives of this study are to identify the key factors that contribute to landslide happenings in the study area, to collect and analyse data on these factors using GIS and Remote Sensing and field investigation, and lastly, to use the data collected and apply MaxEnt model to develop deep-seated landslides susceptibility assessment map for Galiat region of NW Himalaya.

1.1 Study Area

Geographically, the research area is in the northeast of Islamabad, the capital territory, around the Murree area, a hilly tourist destination positioned in the Galiat region of the Himalayan Mountains of Pakistan. The geographical location of the study area based on GIS is given in Figure 1.

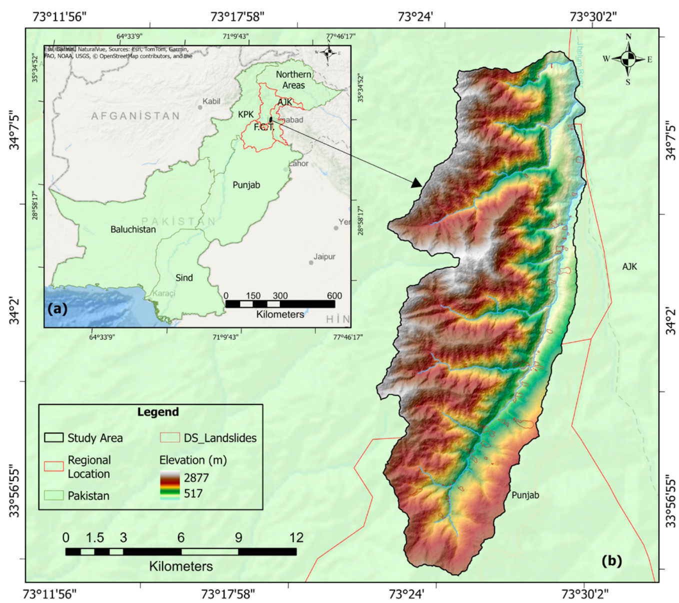


Figure 1. Geographical location map, (a) Regional location, (b) Study area.

The altitude of the study area is between 520 – 2877 m above sea level. Landslides frequently occur in the study area, especially during the rainy season, and inflict enormous socio-economic damage by destroying settlements and businesses in the vicinity. An inventory map of deep-seated landslides was prepared for the 187 km² Galiat region using field surveys and high-resolution imagery from Google Earth Pro (Figure 2). The landslides susceptibility map and its application can significantly reduce the socio-economic losses of the study area.

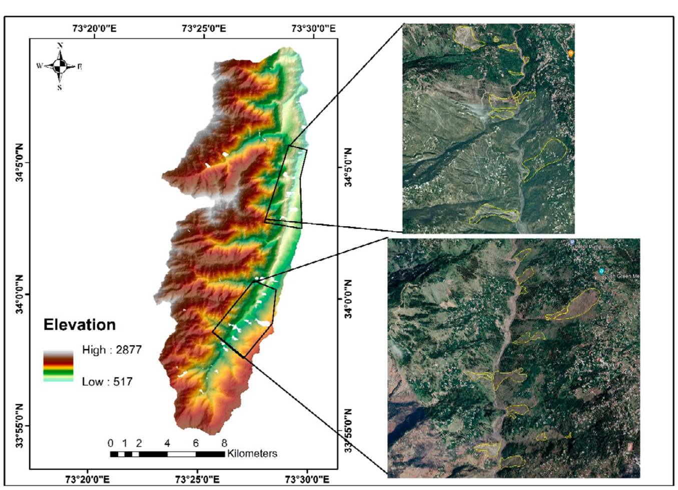


Figure 2. Inventory map of deep-seated landslides compiled using Google Earth and ArcMap.

1.2 Geological Settings of the Area

The study area is influenced significantly by the active seismic and tectonic structure, which plays a substantial role in triggering landslides. The Galiat region primarily comprises the Himalayan Mountains, originated in the

Eocene period through the collision of the Indian with the Eurasian Plate along the Main Mantle Thrust (MMT), as documented by (Tahirkheli 1979; Coward et al., 1986), and it is the highest uplifting region on the globe (Zeitler 1985). Himalayan foothills are characterized by various rocks bounded by major thrust faults, comprising the Main Frontal Thrust (MFT) and the Main Boundary Thrust (MBT). The MBT has shown heightened activity in recent decades, triggered by the 1977 and 2005 earthquakes (Farooq and Malik 1996; Kamp et al., 2008). These seismic events contribute to ground movement, dynamic loading, increased pore water pressure, and shear stress, exacerbating slope instability.

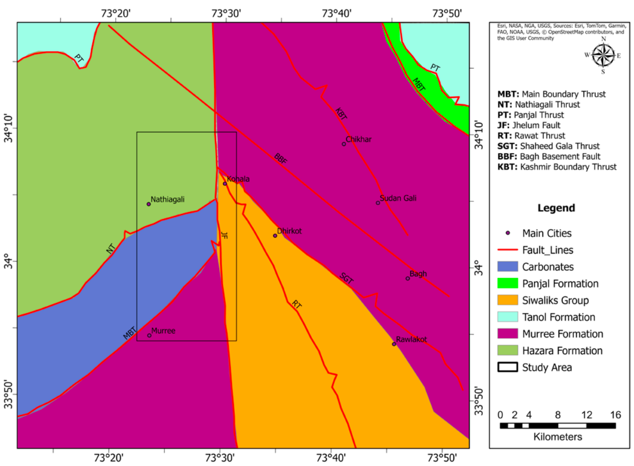


Figure 3. Regional geological map of the study area revised after (Khan et al., 2016).

The geological stratigraphy of the study area also plays a key role in mass movement activities. Along the Himalayan Karakoram Syntaxis (HKS), the Panjal Formation is crammed between the MBT and PT. The Cambrian Panjal Formation has been thrust over the Miocene Murree Formation along the MBT (Mehmood et al., 2023). The PT delineates the tectonic boundary dividing the Precambrian Tanol and the Carboniferous-Triassic Panjal Formation on the eastern limb of HKS (Khan and Ali 1994). Similarly, along the western limb of the HKS, the PT splits the Precambrian Tanol Formation from the Precambrian Hazara Formation. In the north, the MBT isolates Jurassic-Cretaceous and Palaeocene-Eocene periods from the Miocene Murree Formation. The Hazara Formation comprises slate, phyllite, and shale, while the Tanol Formation comprises quartzite and schist. Carbonates, shales, and sandstones from the Samana Suk, Chichali, and Kawagarh formations make up the Jurassic to Cretaceous sequence. Furthermore, the carbonates and shales of the Lockhart, Patala, Margalla Hill, and Chorgali formations make up the Palaeocene to Eocene sequence. The core of the HKS is dominated by Eocene Kuldana and Miocene Murree shale and sandstone formations (Rehman et al., 2020). The Miocene Siwalik Group rocks, composed primarily of sandstone, shale, and conglomerate, are subjected to the south of the HKS. The fine-grained lithology of the Kuldana and Murree formations provides a weak zone for fault localization (Mughal et al., 2018). The study area is characterized by structural complexity and sophistication, resulting from the impact of MBT and its corresponding effects (Iqbal and Bannert 1998). The structural framework is responsible for the development of an extensive water reservoir, which feeds multiple springs and seepages at critical formation toes (Niederer et al., 1989), influencing slope instability of rock and soil structures, resulting in mass movement. The study area's regional geological map is shown in Figure 3.

1.3 Climate

Pakistan is located on the western fringe of the monsoon zone, although the Galiat region is classified as Semi-Continental Highlands. Variations in altitude, quantity of snowfall, and level of snow cover are responsible for considerable climatic diversity. The study area receives an average annual total of 1,900 mm of rainfall and snow over approximately 92 days, though precipitation amounts vary significantly across different parts of the region. The maximum rainfall typically happens in July, August, and September, when the rains are in heavy bursts. The mean monthly precipitation varies from 33 mm in November to 418 mm in July. In winter, snow falls above 1100m and retains above 1600m altitude during January and February. Increasing the moisture content on a relatively unstable slope performs a decisive role in causing mass movement by reducing the shear strength of the material. However, the sandstone of the Murree Formation indicates a reduction of about 20 percent in strength due to moisture, while the overlaid silt/claystone shows a significant reduction in strength due to remolding (Farooq and Malik 1996). The climate information of the study area is shown in Figure 4 and Table 1. Landslides most often happen around monsoon time; however, certain damage is also linked with snowmelt in spring.

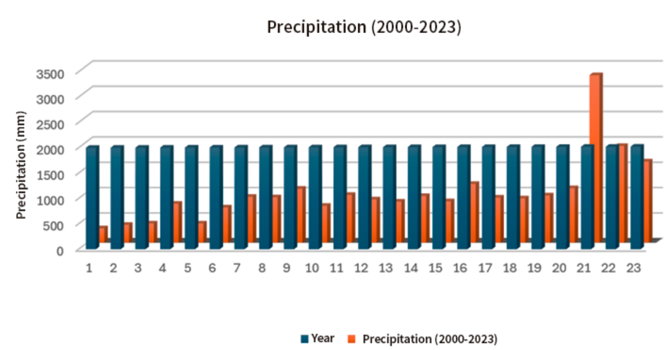


Figure 4. Annual rainfall record from 2000 to 2023 (PMD, 2023)

Table 1. Climate data of the study area 1990-2017 (Nasa, 2023)

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Yearly Mean
Daily mean °C	3.7	4.0	8.0	13.2	17.3	20.6	19.1	18.4	17.2	14.3	10.3	6.3	12.7
Rainy days	4	6	7	8	9	10	14	13	9	5	3	3	92(days)
Humidity (%)	62	64	61	57	50	56	78	82	73	61	60	60	63

2. Materials and Methods:

2.1. Data Used

This research comprises five main steps; (I) The first step was data collection and material regarding the study area. Geotechnical, geological, topographical, and hydrogeological data are among these data and materials. (II) The second step was the development of deep-seated landslides inventory map utilising field investigation and Google Earth Pro. (III) The third step was choosing environmental variables utilising the fieldwork and prior study practices. In susceptibility mapping, the selection and evaluation of environmental parameters is critical for representing the physical processes driving hazard occurrence. The fourteen variables applied in this study cover a broad spectrum of topographic, geological, and hydrological controls. Elevation and slope directly influence gravitational forces and slope stability, with higher gradients typically increasing susceptibility (Orefice and Innocenti, 2025). Elevation captures the influence of altitudinal gradients on gravitational forces and climatic conditions, while slope, slope derivatives, and curvature metrics describe terrain geometry, influencing runoff concentration, erosion, and stress distribution. Lithology dictates material strength and permeability, strongly governing failure mechanisms. Hydrological indices such as CTI and IMI indicate areas of potential moisture accumulation, often linked to slope failures. The transformed aspect accounts for directional exposure of solar radiation and its effect on moisture and weathering. SRR, SAR, and the dissection index quantify drainage energy, fluvial erosion, and landscape dissection intensity. Within the MaxEnt framework, these parameters were weighted by their contribution to the model's gain, enabling identification of the most influential factors for susceptibility in the study area. (IV) The fourth step was calculating the value and the controlling power of each variable in landslide instigation, (V) Finally, to evaluate the susceptibility level of the study area employing the MaxEnt model. The flowchart of the susceptibility evaluation of the study area is given in figure 5.

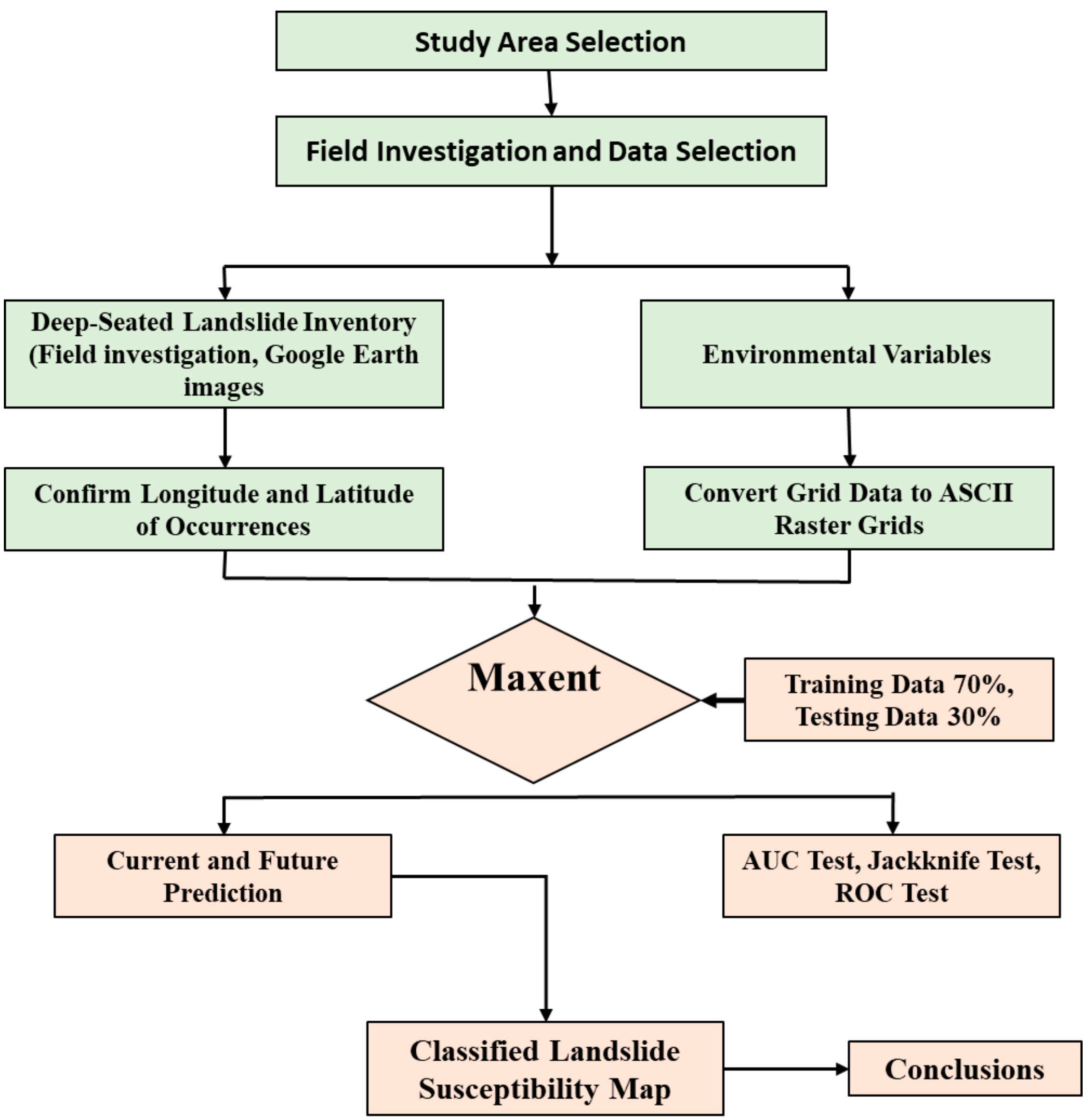


Figure 5. Flowchart of the research work

2.2 Model Selection

In this study, the MaxEnt model has been utilised to evaluate the landslide susceptibility map of the study area. The MaxEnt model is based on the principle of maximum entropy, which states that the best model is the one that makes the fewest assumptions while still being consistent with the available data. The MaxEnt model can handle both continuous and categorical data, making it suitable for the evaluation of a broad range of environmental variables.

2.3 Maximum Entropy Model (MaxEnt)

MaxEnt is a statistical-probabilistic machine learning modeling approach based on maximum entropy. The MaxEnt program was created by AT and T Labs' Steven J. Phillips, Miroslav Dudik, and Robert Schapire for the American Museum of Natural History's Centre for Biodiversity and Conservation (Phillips et al., 2004). The MaxEnt model uses the principle of maximum entropy to predict the probability distribution of an unknown variable. MaxEnt estimates the probability density function (PDF) for each environmental variable based on presence data and contrasts it with the background distribution for the unknown variable. In the case of shallow landslide susceptibility assessment, the unknown variable is the occurrence of landslides. These PDFs help determine how strongly a variable is associated with landslide presence.

The equation used by MaxEnt:

$$P\left(\frac{y}{x}\right) = \exp(\sum \lambda_i F_i(x))/Z \quad (1)$$

where *P(y|x)* is the likelihood of landslide happening *y* assigned a set of environmental variables *x*, *λi* is a Lagrange equation related to an attribute *Fi(x),* and *Z* is the normalizing constant that ensures the sum of all probabilities equals 1. The feature functions *Fi(x)* represents the environmental variables used to model the occurrence of landslide. The Lagrange multipliers λi represent the weights assigned to each feature function and are optimized during the model training process to maximize the likelihood of the observed data.

The MaxEnt model seeks to find the probability distribution of landslide occurrence consistent with the observed environmental variable while making the fewest assumptions possible (Phillips et al., 2009; Elith et al., 2011). The model can handle continuous and categorical data and corporate connections among variables to capture complex relationships (Baldwin 2009; Halvorsen 2012). Entropy, originating from information theory, serves as an important indicator to quantify the level of disorder within a dataset (Shannon 1948). It represents the expected information content within the data necessary for describing a pattern or process (Jaynes 1982; Kleidon et al., 2010).

In the context of landslide occurrence estimation, the MaxEnt approach utilizes the maximum entropy probability distribution, which approximates a uniform distribution across the study area. By doing so, it transforms the incomplete information derived from presence points and associated geoenvironmental data into a less arbitrary distribution for representation. MaxEnt aims to identify the factors with the maximum information value for predicting designs, such as the distribution of landslides in this study. This technique has also been reconstructed as an aspect of logistic regression employing presence/absence data (Renner and Warton 2013). MaxEnt employs an iterative approach to adjust the distribution pattern; the adjustment system consists of a random-walk within the factors space with the allocation of an attribute value. The software offers new parameter settings per iteration, attempting to raise the gain up to the number of iterations chosen by the individual in the advanced factors tab of MaxEnt. The objective is to maximize the gain and obtain the best-fitted model. The more discriminating the distribution, the bigger the gain and the random-walk algorithm begins with an identical probability development in the geographical space. Subsequently, it deviates from this distribution due to data constraints, proceeding through successive iterations until the gain rise falls under a convergence threshold, which is configured by default in MaxEnt but can be specified by the operator (Phillips et al., 2004). The output of the MaxEnt model is a probability map that represents the likelihood of landslide occurrence across the study area. This map can identify areas of high and low susceptibility to landslides and inform decision-making in the management of landslide risk.

2.4 Validation of susceptibility models

The Maxent software provides diverse outputs that facilitate the assessment of various aspects of the results (Elith et al., 2011). One such output is the receiver operating characteristic (ROC), which represents sensitivity on the vertical axis and 1-specificity on the horizontal axis. Generally, the area under curve (AUC) of a model with poor precision and predictive ability is 0.6, whereas it is 1 for complicated systems with great simplification ability. Another fundamental question in modeling is identifying the most influential factor. The MaxEnt software solves this question through several means. The gain of the variables is calculated as the model strives to fit and estimate well in each iteration. As a result, the model uses an enhanced coefficient on elements with significant information, facilitating the model's advancement, and a reduced coefficient on variables with minimal or no information. The coefficient changes with each iteration, relying on the reconsideration of the model results.

Furthermore, the jackknife test enhances the assessment of the importance of variables by presenting the AUC results utilising the test set (Pontius Jr and Schneider 2001). Additionally, the response curves demonstrate how variations in factor values influence landslide susceptibility, providing further insights into the spatial interface between landslide probability and controlling variables, as presented in this study. This provides valuable insights into the relative significance of different variables in the model's predictive performance.

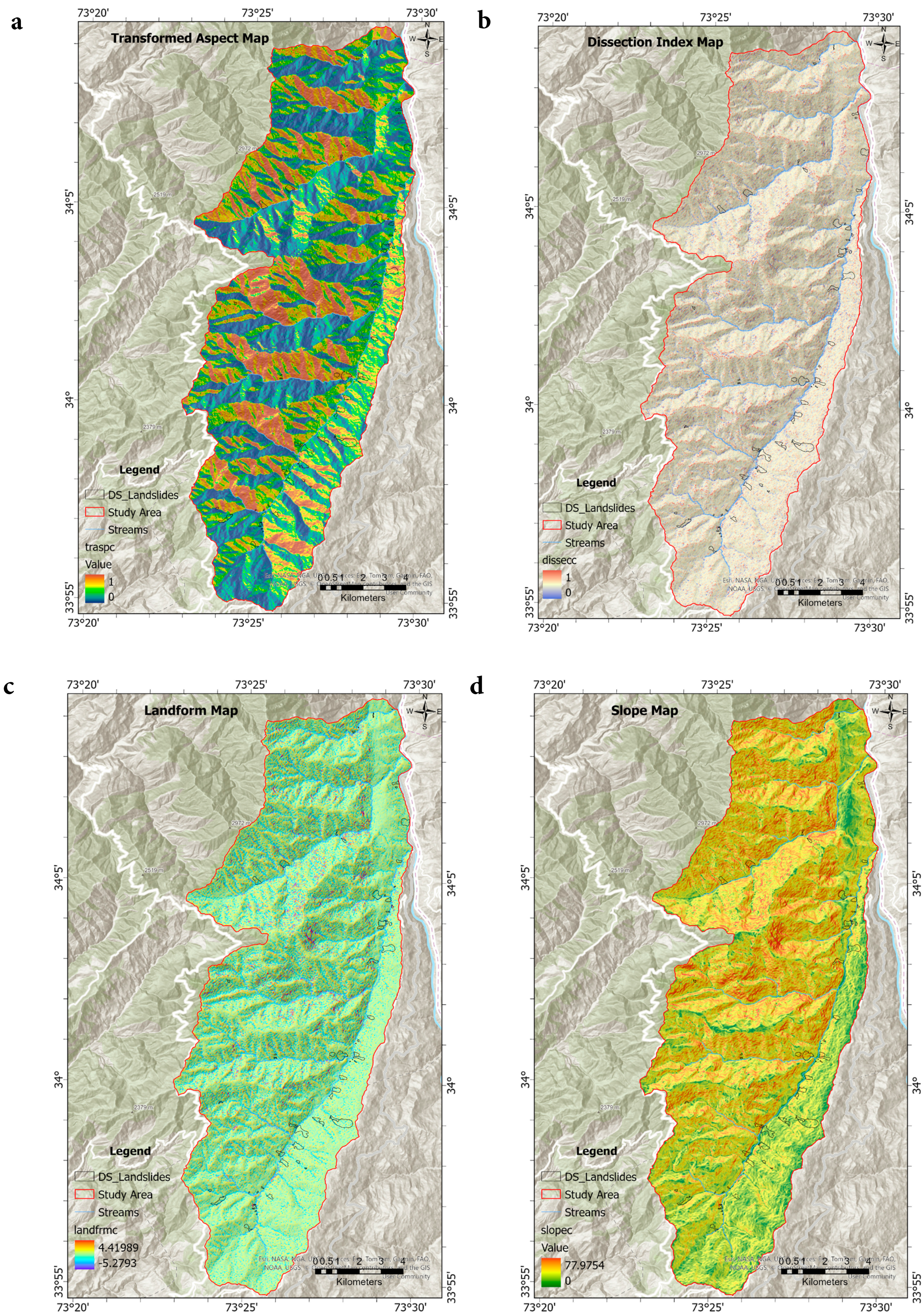
3. Susceptibility Assessment Map of Deep-seated Landslides

3.1. Environmental Variables

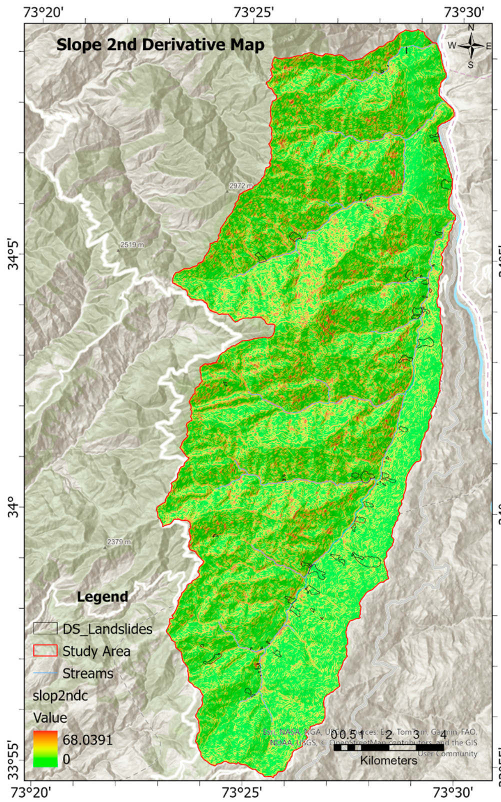
The identification of environmental variables is an essential prerequisite for pre-processing when assessing the susceptibility of landslides. With the recent advancements in GIS software and improved computer power, it has become possible to use a significantly large number of individual factors in data-driven landslide susceptibility assessment. Landslides can be caused by a range of external ecological and internal geological factors (Xiao et al., 2019). Researchers have widely used various environmental variables for landslide susceptibility assessment. The preliminary principle for landslide susceptibility is the collection and development of environmental variables (Li et al., 2020). Fourteen environmental parameters, including Transformed Aspect, Dissection Index, landform, slope, slope 2^{nd} derivative, average slope, curvature plan, curvature profile, SAR, SRR, CTI, and IMI, elevation, and lithology, were considered in the susceptibility mapping of the study area (Figure 6).

3.2 Variable Contribution

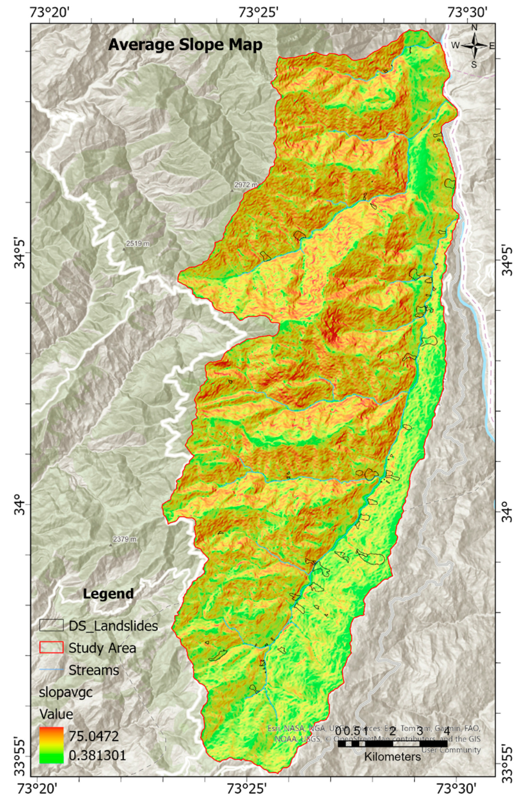
The process of selecting causative factors has identified several critical parameters that may play a significant role in determining the susceptibility of landslides. It is crucial to combine these parameters to obtain a consistent descriptive value for susceptibility analysis. Otherwise, they will remain distinct variables that provide distinct indications. The first estimate is determined by incorporating the increase in regularized gain, either adding it to the corresponding factor's contribution or subtracting it if the alteration in the mean value of lambda (λi) is negative. For the second approximate, the importance of each environmental variable on the training presence and background data are randomly permuted, and the model is reconsidered on the permuted data. Factor contributions, like the factor jackknife, should be taken with prudence when the predictor factors are associated. The values displayed are averages of duplicate runs. Figure 7(a) depicts the findings of the jackknife test of factor significance. When employed alone, the environmental variable with the highest gain is elevation, which appears to contain the most meaningful information. The MaxEnt analysis showed that susceptibility is primarily influenced by elevation, lithology, and slope orientation, reflecting the combined effects of terrain position, material strength, and microclimatic conditions. These variables also reduced the gain the most when omitted, suggesting they contain most of the information not present in the other variables. Hydrological indices such as CTI and IMI, along with slope, landform, and curvature measures, provided secondary insights into the influence of the moisture accumulation and terrain geometry. Other parameters had minimal impact under the current model settings. Overall, susceptibility patterns are driven by a combination of topographic position, geological structure, slope orientation, and hydrological conditions. Figure 7(b) depicts the identical jackknife test using test gain instead of training gain.



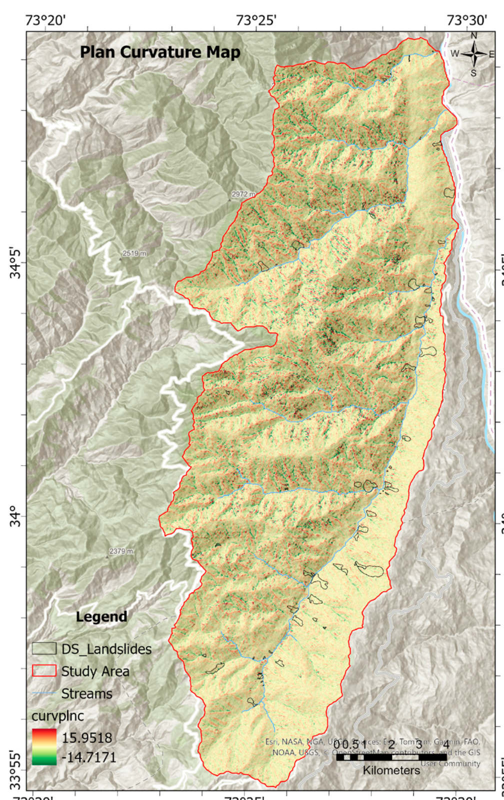
e


Slope 2nd Derivative Map
73°20'
73°25'
73°30'
34°5'
34°
33°55'
Legend
DS_Landslides
Study Area
Streams
slop2ndc
Value
68.0391
0
0 0.5 1 2 3 4
Kilometers

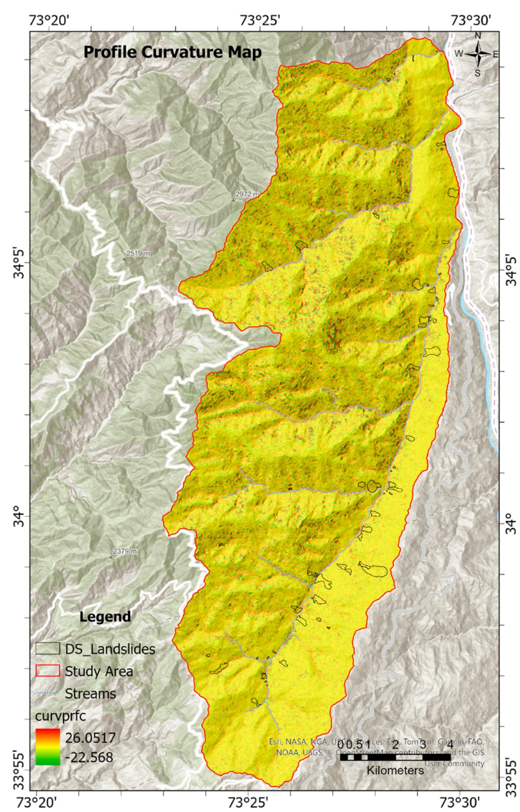
f


Average Slope Map
73°20'
73°25'
73°30'
34°5'
34°
33°55'
Legend
DS_Landslides
Study Area
Streams
slopavgc
Value
75.0472
0.381301
0 0.5 1 2 3 4
Kilometers

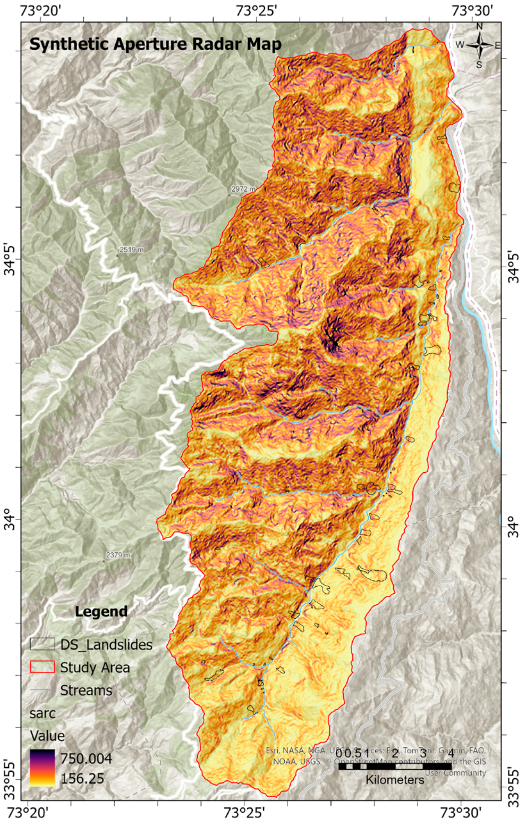
g


Plan Curvature Map
73°20'
73°25'
73°30'
34°5'
34°
33°55'
Legend
DS_Landslides
Study Area
Streams
curvplnc
15.9518
-14.7171
0 0.5 1 2 3 4
Kilometers

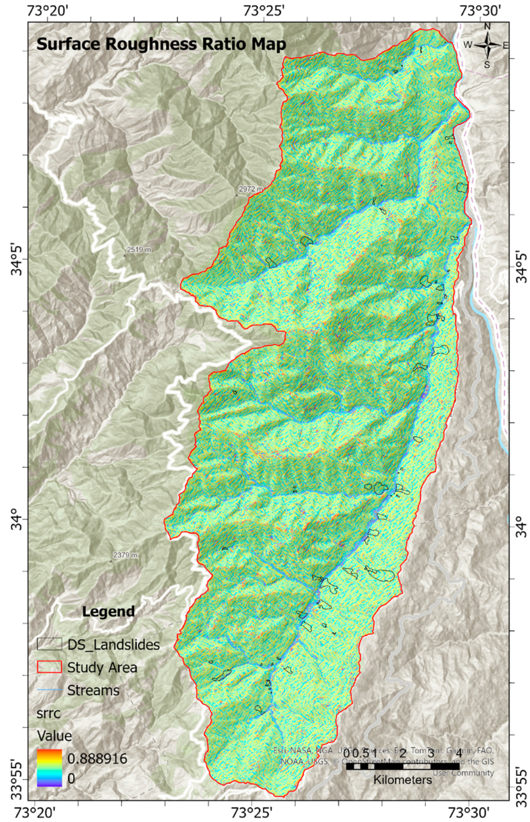
h


Profile Curvature Map
73°20'
73°25'
73°30'
34°5'
34°
33°55'
Legend
DS_Landslides
Study Area
Streams
curvprfc
26.0517
-22.568
0 0.5 1 2 3 4
Kilometers

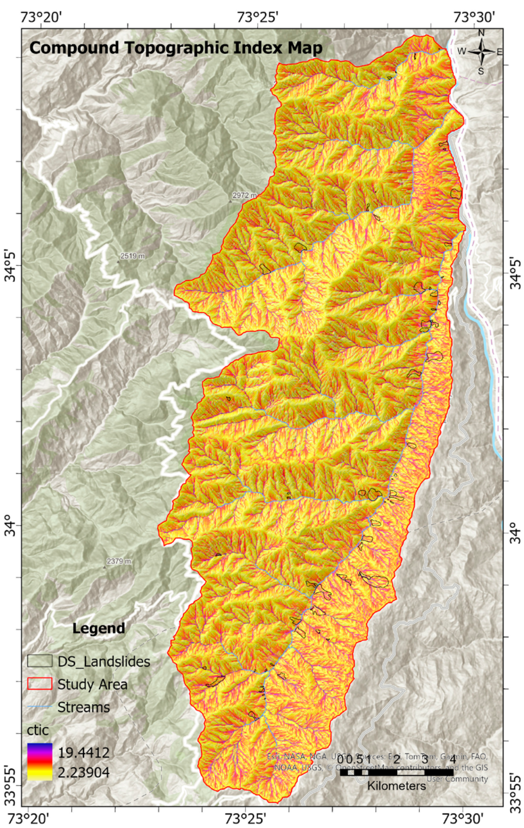
i


Synthetic Aperture Radar Map
73°20'
73°25'
73°30'
34°5'
34°
33°55'
Legend
DS_Landslides
Study Area
Streams
sarc
Value
750.004
156.25
Kilometers

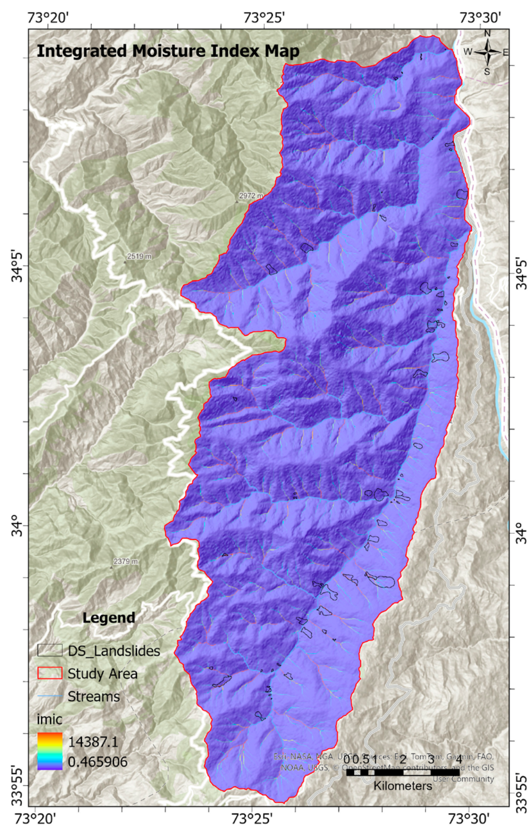
j


Surface Roughness Ratio Map
73°20'
73°25'
73°30'
34°5'
34°
33°55'
Legend
DS_Landslides
Study Area
Streams
srrc
Value
0.888916
0
Kilometers

k


Compound Topographic Index Map
73°20'
73°25'
73°30'
34°5'
34°
33°55'
Legend
DS_Landslides
Study Area
Streams
ctic
19.4412
2.23904
Kilometers

l


Integrated Moisture Index Map
73°20'
73°25'
73°30'
34°5'
34°
33°55'
Legend
DS_Landslides
Study Area
Streams
imic
14387.1
0.465906
Kilometers

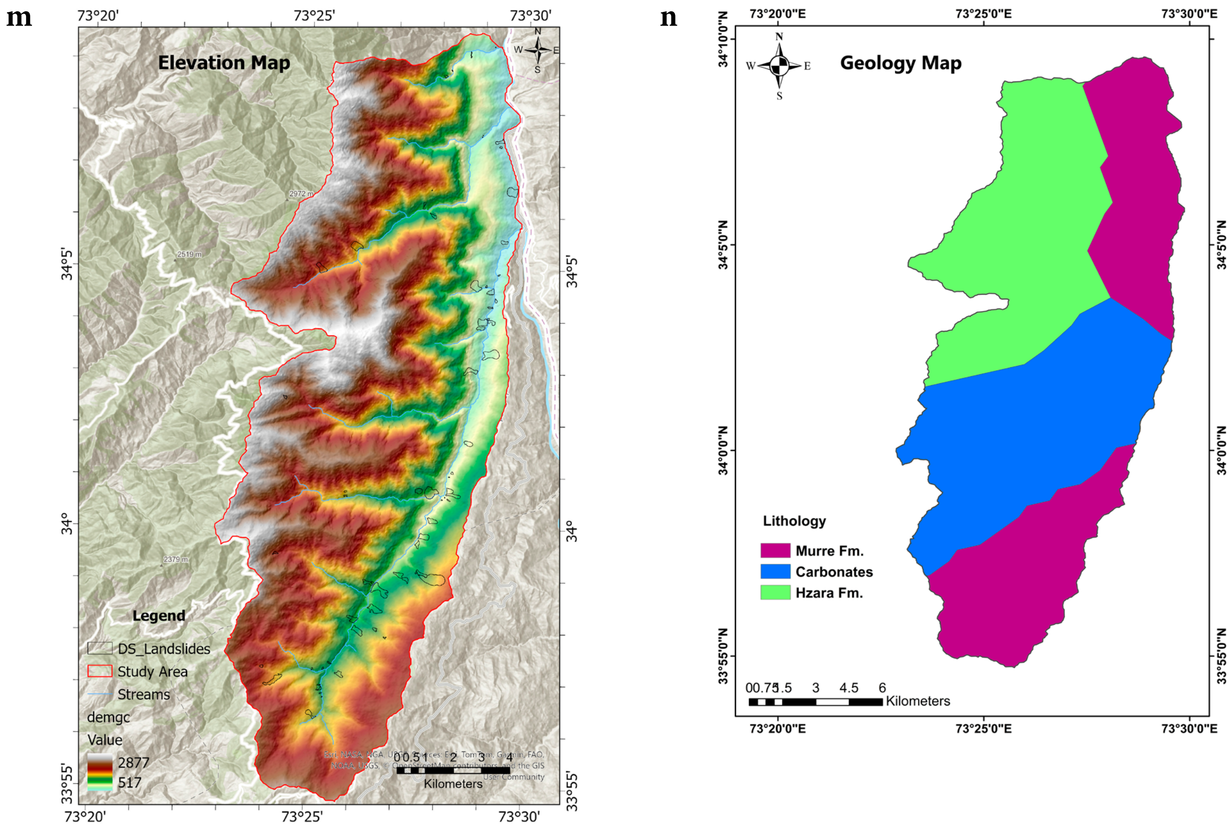


Figure 6. Acquired maps of the fourteen environmental variables used in this study ((a) Transformed Aspect, (b) Dissection Index, (c) landform, (d) slope, (e) slope 2nd derivative, (f) average slope, (g) curvature plan, (h) curvature profile, (i) SAR, (j) SRR, (k) CTI, (l) IMI, (m) elevation, and (n) lithology).

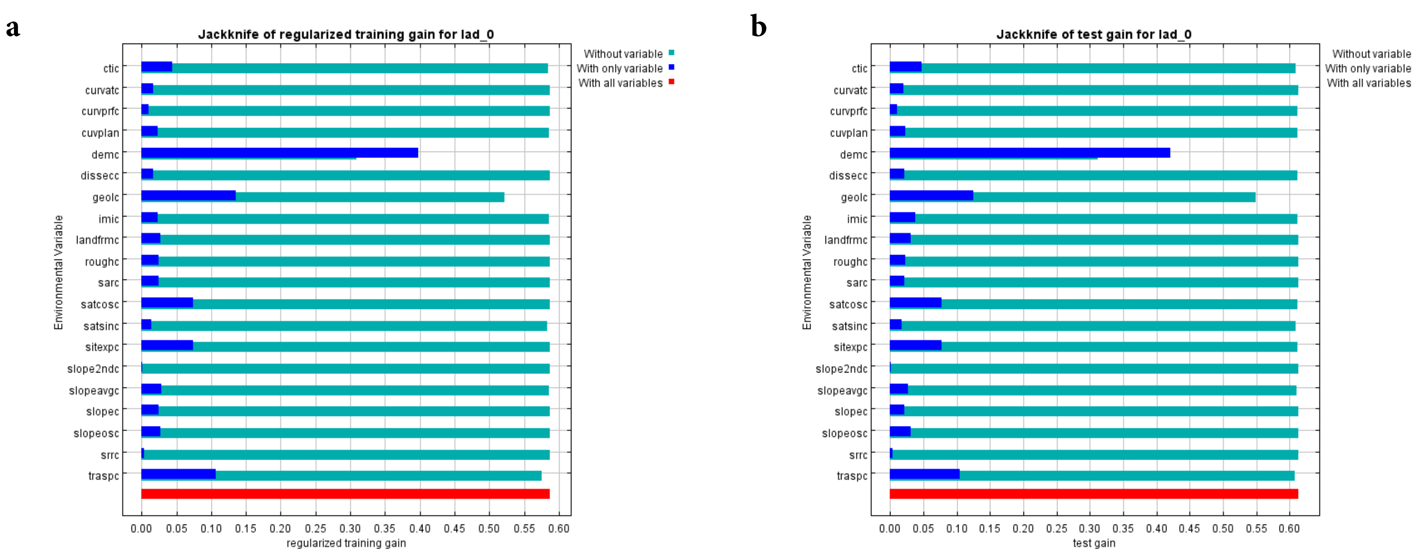


Figure 7. (a) Jackknife test of variable importance. (b) jackknife test, using test gain.

3.3. Susceptibility Assessment Map

The susceptibility of landslides is calculated using the environmental variables and their weights. The choice of probability intervals in susceptibility classes exhibits significant variation, including options like natural Jenks and equal intervals. This selection greatly affects the division and distribution of susceptibility classes within the study area. For this study, a Natural Breaks (Jenks) approach was employed for probability classes, leading to a situation where a small portion of the study area encompasses a substantial percentage of the existing landslides classified as high and very high susceptibility. The weight of individual variables is calculated in the MaxEnt model. The cumulative map obtained with the MaxEnt program is reclassified into five classes, i.e., very low, low, moderate, high, and very high, respectively, to develop a landslide susceptibility map for the study area and the graphical representation of the map classification is shown in Figure 8a. Notably, the areas categorized as high to very high susceptibility cover 26.16% of the total area, yet they contain 74.3% of the mapped landslide occurrences, indicating a strong spatial correlation between high susceptibility zones and actual landslide events (Fig. 8b).

3.4. Validation of susceptibility models

In this study, various metrics can be used to assess and confirm the accuracy of a susceptibility model. Specifically, we define evaluation as the assessment of the agreement between the Susceptibility classification system and the frequency of landslides utilized for model calibration or training, while confirmation is an assessment of the model's projecting ability using landslides not utilized during model calibration or training. One effective statistical tool for assessing the discriminant capability of the model and comparing the relative validity among models is the ROC-AUC provided in different MaxEnt output plots. The AUC estimates the model's accuracy, with values ranging from 0.6 to 1, where higher AUC values indicate better prediction. Several conventional AUC interpretation scales classify results based on their accuracy or goodness. According to the (Metz, 1978). criterion, AUC-ROC values above 0.90 are regarded as excellent, those from 0.80 to 0.89 as good, those between 0.70 and 0.79 regular, and those less than 0.70 as bad. In this study, the obtained AUC value is 0.80, regarded as reliable for landslide susceptibility mapping (figure 9).

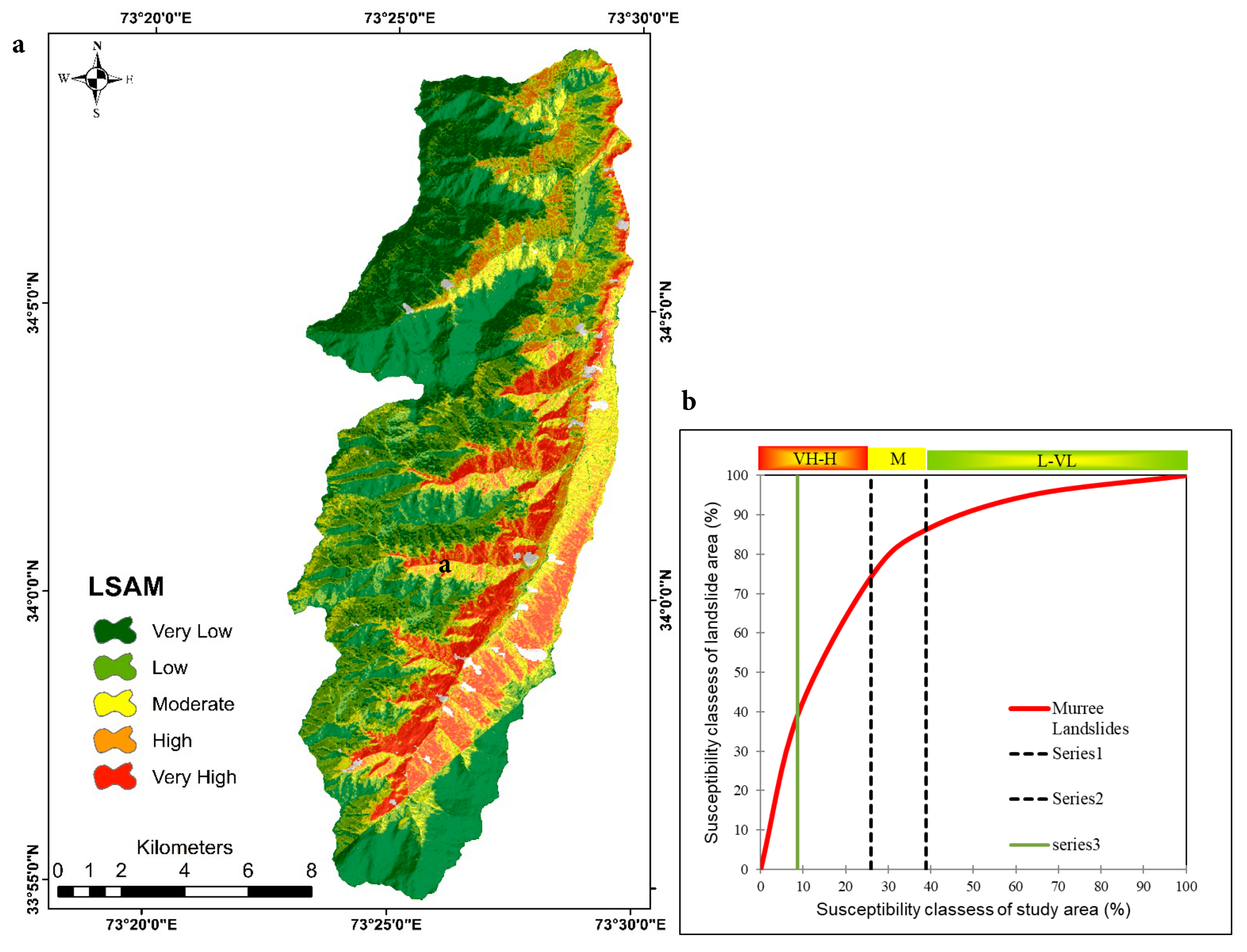


Figure 8. Acquired deep-seated landslide susceptibility map (a), the percentage in the different landslide susceptibility classes (b).

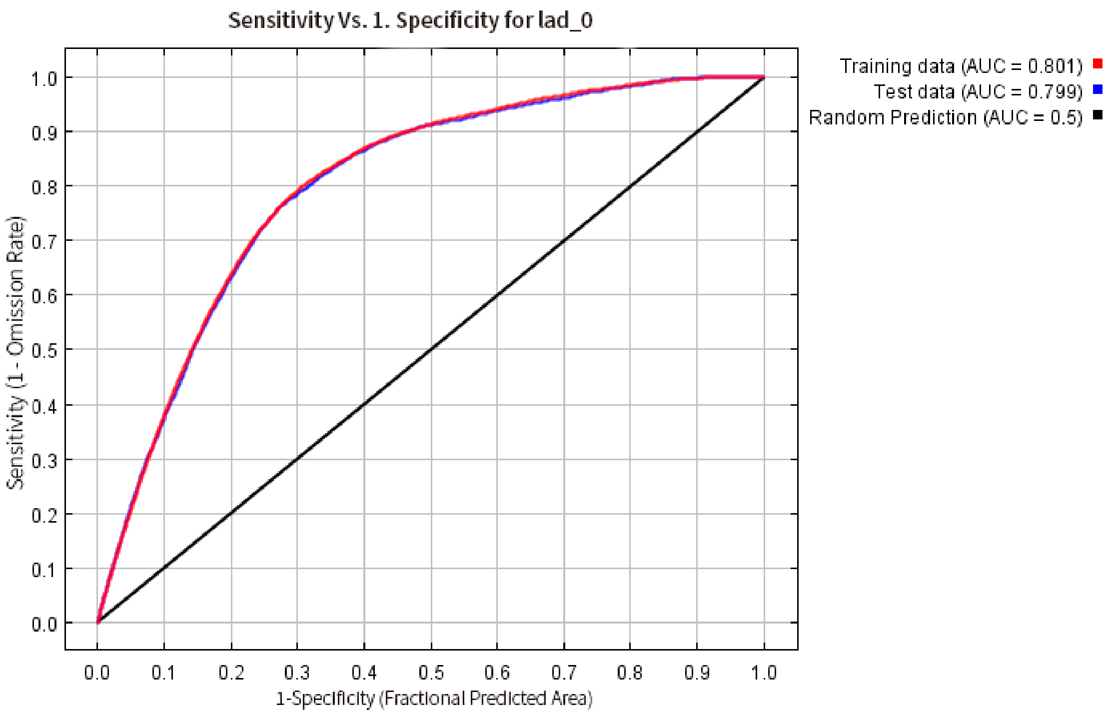


Figure 9. The area under Curve graph of the study area extracted from the MaxEnt program.

4. Discussion and Conclusion

This study was arranged to create a landslide susceptibility map covering the 187 km^2 area of the Galiat region around Murree in the NW-Himalaya, which is highly prone to landslides, especially in the last decade. The relevant deep-seated landslide inventory map was generated based on field investigation and multi-temporal Google Earth Pro images. 68 deep-seated landslide polygon features ranging in size from 348 m^2 to 258612 m^2 were determined to evaluate landslide susceptibility. Since the landslide mostly occurred in limited areas, significant predictions cannot be made with classical statistical modelling. Therefore, the landslide susceptibility map of the study area was modelled using the maximum entropy method, which permits predictions from limited observation data. The analyses were repeated with two randomly selected 30% and 70% data sets for testing and training data, respectively. Fourteen environmental variables were considered, including geology, DEM, first and second DEM derivative parameters. Our results indicated that the most critical parameters that contributed to the Maxent models were DEM, geology, topographic radiation aspect index, and aspect transformation. We evaluated the accuracy and sensitivity of the obtained models using the area under the receiver operating characteristic and the area under curve. The average area under the receiver operating characteristic curves was 0.80±0.002. We also found that the high to very high susceptible classes corresponded to 34.16% of the study area, including 74.3% of the mapped landslides.

MaxEnt is a machine learning model that specializes in presence-only data. This feature is advantageous in remote and impassable areas with limited access to presence and absence data (Pearson et al., 2007). In landslide studies, it is essential to consider that the absence of landslides does not necessarily indicate the absence of susceptibility. It is possible that an area with no recorded landslide may still have a high potential for landslide occurrence but has not been observed or properly captured by the researcher. Therefore, the maximum entropy model, which relies only on presence locations, can reduce many of these inconsistencies.

Moreover, the misidentification of landslides and inadequate consideration of environmental variables can jeopardize the efficiency and factors estimation of the model. A field survey was conducted to address these issues, including expert opinions and the selection of relevant environmental variables to better represent the landslide occurrence process. One other limitation of MaxEnt is that it simplistically accounts for landslide inventory by only including point features in ASCII format and disregarding information about their shape and dimensions. Despite the user-friendly interface, there is complex math involved in the MaxEnt software. It operates with a set of landslides, environmental variables, and two sets of incomplete random sample data that provide inadequate information regarding the objective phenomenon (Kornejady et al., 2017). MaxEnt randomly splits the sample data into two sets, 70% for training and 30% for results validation. The software aims to approximate the best probability distribution of the objective phenomenon using this incomplete information. To accomplish this, it starts with an identical likelihood distribution equation as a first guess and employs constraints to determine the best PDF of the phenomenon, known as the maximum entropy function, over a spatial context known as the feature. The features are the related functions and transformations of the variables that are connected by constraints. There are five types of features, including continuous layers that form linear, quadratic, product, and threshold components, and categorical layers that form binary features. The model imposes constraints on the mean, variance, and covariance of the variables to find the best PDF of the phenomenon.

The developed landslide susceptibility map can provide valuable information for residents, engineers, and land-use authorities to reduce the destruction caused by landslides. By identifying active and potential landslide-prone areas, land-use decisions can be made more effectively, and hazard mitigation strategies can be developed. Overall, our study highlights the importance of developing a landslide susceptibility map in high-risk areas and provides a useful framework for future studies on landslide susceptibility assessment in the Sub-Himalayan Range.

Acknowledgements:

The authors would like to thank the editor and anonymous reviewers for their comments and suggestions, which helped a lot in making this paper better.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest –

The authors declare no conflict of interest.

References

Baldwin, R. A. (2009). Use of maximum entropy modeling in wildlife research. *Entropy,* 11, 854-866. https://doi.org/10.3390/e11040854

Cao, C., Xu, P., Chen, J., Zheng, L., & Niu, C. (2017). Hazard assessment of debris-flow along the Baicha River in Heshigten Banner, Inner Mongolia, China. *International Journal of Environmental Research and Public Health,* 14, 30-30. https://doi.org/10.3390/ijerph14010030

Chen, W., Pourghasemi, H. R., Kornejady, A., & Zhang, N. (2017). Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma,* 305, 314-327. https://doi.org/10.1016/j.geoderma.2017.06.020

Chen, X., Chen, H., You, Y., Chen, X., & Liu, J. (2016). Weights-of-evidence method based on GIS for assessing susceptibility to debris flows in Kangding County, Sichuan Province, China. *Environmental Earth Sciences,* 75, 1-16. https://doi.org/10.1007/s12665-015-5033-z

Chen, Z., & Wang, J. (2007). Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada. *Natural Hazards,* 42, 75-89. https://doi.org/10.1007/s11069-006-9061-6

Cömert, R., Avdan, U., & Gorum, T. (2018). Rapid mapping of forested landslide from ultra-high resolution unmanned aerial vehicle data. *International Society for Photogrammetry and Remote Sensing.* https://doi.org/10.5194/isprs-archives-XLII-3-W4-171-2018

Conoscenti, C., Ciaccio, M., Caraballo-Arias, N. A., Gómez-Gutiérrez, Á., Rotigliano, E., & Agnesi, V. (2015). Assessment of susceptibility to earth-flow landslide using logistic regression and multivariate adaptive regression splines: A case of the Belice River basin (western Sicily, Italy). *Geomorphology,* 242, 49-64. https://doi.org/10.1016/j.geomorph.2014.09.020

Convertino, M., Troccoli, A., & Catani, F. (2013). Detecting fingerprints of landslide drivers: A MaxEnt model. *Journal of Geophysical Research: Earth Surface,* 118, 1367-1386. https://doi.org/10.1002/jgrf.20099

Coward, M. P., Rex, D. C., Asif Khan, M., Windley, B. F., Broughton, R. D., Luff, I. W., Petterson, M. G., Pudsey, C. J. (1986). Collision tectonics in the NW Himalayas. *Geological Society, London, Special Publications,* 19, 203-219. https://doi.org/10.1144/GSL.SP.1986.019.01.11

Duman, T. Y. & Çan, T. (2023). Characteristics of landslides and assessment of deep-seated landslide susceptibility in Northern Turkey. *Mediterranean Geoscience Reviews,* 1-27. https://doi.org/10.1007/s42990-023-00105-3

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions,* 17, 43-57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Farooq, S., & Malik, M. H. (1996). *Landslide hazard management and control in Pakistan: A review.* Kathmandu: ICIMOD, 68 p.

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., & Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Engineering Geology,* 102, 85-98. https://doi.org/10.1016/j.enggeo.2008.03.022

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., & Reichenbach, P. (2008). Comparing landslide inventory maps. *Geomorphology,* 94, 268-289. https://doi.org/10.1016/j.geomorph.2006.09.023

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. T. (2012). Landslide inventory maps: New tools for an old problem. *Earth-Science Reviews,* 112, 42-66. https://doi.org/10.1016/j.earscirev.2012.02.001

Halvorsen, R. (2012). A gradient analytic perspective on distribution modelling. *Sommerfeltia,* 35, 1-165. https://doi.org/10.2478/v10208-011-0015-3

Iqbal, M., & Bannert, D. (1998). Structural observations of the Margala hills, Pakistan and the nature of the Main Boundary Thrust. *Pakistan Journal of Hydrocarbon Research,* 10, 41-53.

Jaynes, E. T. (1982). On the rationale of maximum-entropy methods. *Proceedings of the IEEE,* 70, 939-952.

Kamp, U., Growley, B. J., & Khattak, G. A., & Owen, L. A. (2008). GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. *Geomorphology,* 101, 631-642. https://doi.org/10.1016/j.geomorph.2008.03.003

Khan, M. R., Hameed, F., Mughal, M. S., Basharat, M., & Mustafa, S. (2016). Tectonic study of the Sub-Himalayas based on geophysical data in Azad Jammu and Kashmir and northern Pakistan. *Journal of Earth Science,* 27(6), 981–988. https://doi.org/10.1007/s12583-016-0681-9

Khan, M. R., & Ali, M. (1994). Preliminary modeling of the western Himalaya, Kashmir. *Journal of Geology,* 11, 59-66.

Kim, H. G., Lee, D. K., Park, C., Kil, S., Son, Y., & Park, J. H. (2015). Evaluating landslide hazards using RCP 4.5 and 8.5 scenarios. *Environmental Earth Sciences,* 73, 1385-1400. https://doi.org/10.1007/s12665-014-3775-7

Kleidon, A., Malhi, Y., & Cox, P. M. (2010). Maximum entropy production in environmental and ecological systems. *The Royal Society,* 365. https://doi.org/10.1098/rstb.2010.0018

Kornejady, A., Ownegh, M., & Bahremand, A. (2017). Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. *Catena,* 152, 144-162. https://doi.org/10.1016/j.catena.2017.01.010

Lee, S. (2005). Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *International Journal of Remote Sensing,* 26, 1477-1491. https://doi.org/10.1080/01431160412331331012

Li, Y., Chen, J., Tan, C., Li, Y., Gu, F., Zhang, Y. & Mehmood, Q. (2020). Application of the borderline-SMOTE method in susceptibility assessments of debris flows in Pinggu District, Beijing, China. *Natural Hazards.* https://doi.org/10.1007/s11069-020-04409-7

Mehmood, Q., Qing, W., Chen, J., & Rahman, G. (2024). Predicting debris flow runout with SFLOW model: Case study of Datong and Taicun gullies, Lijiang, China. *Geotechnical and Geological Engineering,* 43(1), 47. https://doi.org/10.1007/s10706-024-03020-y

Mehmood, Q., Khan, A., Chen, J., Rahman, G., Nasrullah, & Tosunlu, H. (2023). Susceptibility assessment of landslide using Analytical Hierarchy Process and Weighted Overlay Analysis, along N-75 highway, Pakistan. *Acta Geologica Slovaca,* 15, 23-34.

Mehmood, Q., Qing, W., Chen, J., Yan, J., Ammar, M., Rahman, G., & Nasrullah (2021). Susceptibility assessment of single gully debris flow based on AHP and extension method. *Civil Engineering Journal (Iran),* 7. https://doi.org/10.28991/cej-2021-03091702.

Melo, R., Vieira, G., Caselli, A., & Ramos, M. (2012). Susceptibility modelling of hummocky terrain distribution using the information value method (Deception Island, Antarctic Peninsula). *Geomorphology,* 155, 88-95. https://doi.org/10.1016/j.geomorph.2011.12.027

Mert, A., Özkan, K., Şentürk, Ö., & Negiz, M. G. (2016). Changing the potential distribution of Turkey Oak (Quercus cerris L.) under climate change in Turkey. *Polish Journal of Environmental Studies,* 25, 1633-1638. https://doi.org/10.15244/pjoes/62230

Metz, C. E. (1978). *Basic principles of ROC analysis.* In Seminars in Nuclear Medicine (pp. 283-298). Elsevier.

Mughal, M. S., Zhang, C., Du, D., Zhang, L., Mustafa, S., Hameed, F., Khan, M. R., Zaheer, M., & Blaise, D. (2018). Petrography and provenance of the Early Miocene Murree Formation, Himalayan Foreland Basin, Muzaffarabad, Pakistan. *Journal of Asian Earth Sciences,* 162, 25-40. https://doi.org/10.1016/j.jseaes.2018.04.018

NASA. (2023). NASA POWER. https://power.larc.nasa.gov/data-access-viewer/

Niederer, S., & Schaffner, U. R. (1989). *Landslide problems and erosion control in Murree and Kohat tehsils of Rawalpindi Dist: Results of the fact-finding mission.* Swiss Development Cooperation, Ministry of Foreign Affairs, Government of Switzerland.

Niederer, S., Wagner, A., Khan, S. R., & Rafiq, M. (1989). *Murree erosion control: Results of the facts-finding mission.* SDC, Ministry of Foreign Affairs, Government of Switzerland.

Orefice, S., & Innocenti, C. (2025). Regional assessment of coastal landslide susceptibility in Liguria, Northern Italy, using MaxEnt. *Natural Hazards,* 121(3), 2613–2639. https://doi.org/10.1007/s11069-024-06833-5

Pandey, V. K., Pourghasemi, H. R., & Sharma, M. C. (2020). Landslide susceptibility mapping using maximum entropy and support vector machine models along the Highway Corridor, Garhwal Himalaya. *Geocarto International,* 35, 168-187. https://doi.org/10.1080/10106049.2018.1510038

Park, N. W. (2015). Using maximum entropy modeling for landslide susceptibility mapping with multiple geoenvironmental data sets. *Environmental Earth Sciences,* 73, 937-949. https://doi.org/10.1007/s12665-014-3442-z

Pearson, R. G., Raxworthy, C. J., Nakamura, M., & Townsend Peterson, A. (2007). Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *Journal of Biogeography,* 34, 102-117. https://doi.org/10.1111/j.1365-2699.2006.01594.x

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications,* 19, 181-197. https://doi.org/10.1890/07-2153.1

Phillips, S. J., Dudík, M., & Schapire, R. E. (2004). *A maximum entropy approach to species distribution modeling.* Proceedings of the Twenty-First International Conference on Machine Learning (pp. 83-83). https://doi.org/10.1145/1015330.1015412

PMD. (2023). Pakistan Meteorological Department. In Rainfall Reports. Pakistan. https://www.pmd.gov.pk/en/

Pontius, R. G. Jr., & Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment,* 85, 239-248. https://doi.org/10.1016/S0167-8809(01)00187-6

Pourghasemi, H. R., & Rossi, M. (2017). Landslide susceptibility modeling in a landslide prone area in Mazandarn Province, north of Iran: A comparison between GLM, GAM, MARS, and M-AHP methods. *Theoretical and Applied Climatology,* 130, 609-633. https://doi.org/10.1007/s00704-016-1919-2

Rehman, G., Zhang, G., Rahman, M. U., Rahman, N. U., Usman, T., & Imraz, M. (2020). The engineering assessments and potential aggregate analysis of Mesozoic carbonates of Kohat Hills Range, KP, Pakistan. *Acta Geodaetica et Geophysica*, 55, 477-493. https://doi.org/10.1007/s40328-020-00301-9

Renner, I. W., & Warton, D. I. (2013). Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. *Biometrics,* 69, 274-281. https://doi.org/10.1111/j.1541-0420.2012.01824.x

Rosi, A., Frodella, W., Nocentini, N., Caleca, F., Havenith, H. B., Strom, A., Saidov, M., Bimurzaev, G. A., & Tofani, V. (2023). Comprehensive landslide susceptibility map of Central Asia. *Natural Hazards and Earth System Sciences,* 23, 2229-2250. https://doi.org/10.5194/nhess-23-2229-2023

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Technical Journal,* 27, 379-423.

Sun, X., Chen, J., Bao, Y., Han, X., Zhan, J., & Peng, W. (2018). Landslide susceptibility mapping using logistic regression analysis along the Jinsha River and its tributaries close to Derong and Deqin County, southwestern China. *ISPRS International Journal of Geo-Information,* 7, 438. https://doi.org/10.3390/ijgi7110438

Tahirkheli, R. A. K. (1979). The India-Eurasia suture zone in northern Pakistan: Synthesis and interpretation of recent data at plate scale. *Geodynamics of Pakistan,* 125-130.

Tsangaratos, P., & Ilia, I. (2016). Landslide susceptibility mapping using a modified decision tree classifier in the Xanthi Perfection, Greece. *Landslides,* 13, 305-320. https://doi.org/10.1007/s10346-015-0565-6

Van Westen, C. J., Castellanos, E. & Kuriakose, S. L. (2008). Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology,* 102, 112-131. https://doi.org/10.1016/j.enggeo.2008.03.010

Vorpahl, P., Elsenbeer, H., Märker, M. & Schröder, B. (2012). How can statistical models help to determine driving factors of landslides? *Ecological Modelling,* 239, 27-39. https://doi.org/10.1016/j.ecolmodel.2011.12.007

Wu, S., Chen, J., Xu, C., Zhou, W., Yao, L., Yue, W., & Cui, Z. (2020). Susceptibility assessments and validations of debris-flow events in meizoseismal areas: Case study in China's Longxi River watershed. *Natural Hazards Review,* 21, 05019005. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000347

Xiao, T., Segoni, S., Chen, L., Yin, K., & Casagli, N. (2020). A step beyond landslide susceptibility maps: A simple method to investigate and explain the different outcomes obtained by different approaches. *Landslides,* 17, 627-640. https://doi.org/10.1007/s10346-019-01299-0

Xiao, T., Yin, K., Yao, T., & Liu, S. (2019). Spatial prediction of landslide susceptibility using GIS-based statistical and machine learning models in Wanzhou County, Three Gorges Reservoir, China. *Acta Geochimica,* 38, 654-669. https://doi.org/10.1007/s11631-019-00341-1

Zeitler, P. K. (1985). Cooling history of the NW Himalaya, Pakistan. *Tectonics,* 4, 127-151.

Zhang, Y., Ge, T., Tian, W. & Liou, Y. A. (2019). Debris flow susceptibility mapping using machine-learning techniques in Shigatse area, China. *Remote Sensing,* 11, 2801. https://doi.org/10.3390/rs11232801