

Article

Geohazards Susceptibility Assessment along the Upper Indus Basin Using Four Machine Learning and Statistical Models

Hilal Ahmad ^{1,2,†}, Chen Ningsheng ^{1,2,*}, Mahfuzur Rahman ^{1,2,3,†}, Md Monirul Islam ³,
Hamid Reza Pourghasemi ⁴, Syed Fahad Hussain ^{1,2}, Jules Maurice Habumugisha ^{1,2}, Enlong Liu ^{5,6},
Han Zheng ⁷, Huayong Ni ⁸ and Ashraf Dewan ⁹

- ¹ Key Laboratory for Mountain Hazards and Earth Surface Process, Institute of Mountain Hazards and Environment, Chinese Academy of Sciences (CAS), Chengdu 610041, China; hilal@imde.ac.cn (H.A.); mfz.rahman@iubat.edu (M.R.); fahadishah6@mails.ucas.ac.cn (S.F.H.); maurice@imde.ac.cn (J.M.H.)
 - ² University of Chinese Academy of Sciences, Beijing 100049, China
 - ³ Department of Civil Engineering, International University of Business Agriculture and Technology (IUBAT), Dhaka 1230, Bangladesh; mmislam@iubat.edu
 - ⁴ Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz, Iran; hr.pourghasemi@shirazu.ac.ir
 - ⁵ State Key Laboratory of Hydraulics and Mountain River Engineering, College of Water Resource and Hydropower, Sichuan University, Chengdu 610065, China; liuenlong@lzb.ac.cn
 - ⁶ State Key Laboratory of Hydraulics and Natural River Engineering, College of Water Resource and Hydropower, Sichuan University, Chengdu 610065, China
 - ⁷ School of Civil Engineering, Central South University, Changsha 410075, China; zheng_han@csu.edu.cn
 - ⁸ Chengdu Institute of Geology and Mineral Resources, China Geological Survey, Beijing 100037, China; nhuayong@cgs.cn
 - ⁹ School of Earth and Planetary Sciences, Curtin University, Kent St, Bentley, WA 6102, Australia; A.Dewan@curtin.edu.au
- * Correspondence: chennsh@imde.ac.cn
 † These authors contributed equally to this work and should be considered as co-first authors.



Citation: Ahmad, H.; Ningsheng, C.; Rahman, M.; Islam, M.M.; Pourghasemi, H.R.; Hussain, S.F.; Habumugisha, J.M.; Liu, E.; Zheng, H.; Ni, H.; et al. Geohazards Susceptibility Assessment along the Upper Indus Basin Using Four Machine Learning and Statistical Models. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 315. <https://doi.org/10.3390/ijgi10050315>

Academic Editors: Wolfgang Kainz and Massimiliano Cannata

Received: 3 February 2021
 Accepted: 1 May 2021
 Published: 7 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The China–Pakistan Economic Corridor (CPEC) project passes through the Karakoram Highway in northern Pakistan, which is one of the most hazardous regions of the world. The most common hazards in this region are landslides and debris flows, which result in loss of life and severe infrastructure damage every year. This study assessed geohazards (landslides and debris flows) and developed susceptibility maps by considering four standalone machine-learning and statistical approaches, namely, Logistic Regression (LR), Shannon Entropy (SE), Weights-of-Evidence (WoE), and Frequency Ratio (FR) models. To this end, geohazard inventories were prepared using remote sensing techniques with field observations and historical hazard datasets. The spatial relationship of thirteen conditioning factors, namely, slope (degree), distance to faults, geology, elevation, distance to rivers, slope aspect, distance to road, annual mean rainfall, normalized difference vegetation index, profile curvature, stream power index, topographic wetness index, and land cover, with hazard distribution was analyzed. The results showed that faults, slope angles, elevation, lithology, land cover, and mean annual rainfall play a key role in controlling the spatial distribution of geohazards in the study area. The final susceptibility maps were validated against ground truth points and by plotting Area Under the Receiver Operating Characteristic (AUROC) curves. According to the AUROC curves, the success rates of the LR, WoE, FR, and SE models were 85.30%, 76.00, 74.60%, and 71.40%, and their prediction rates were 83.10%, 75.00%, 73.50%, and 70.10%, respectively; these values show higher performance of LR over the other three models. Furthermore, 11.19%, 9.24%, 10.18%, 39.14%, and 30.25% of the areas corresponded to classes of very-high, high, moderate, low, and very-low susceptibility, respectively. The developed geohazard susceptibility map can be used by relevant government officials for the smooth implementation of the CPEC project at the regional scale.

Keywords: China–Pakistan economic corridor; landslides; debris flows; geohazards; remote sensing

1. Introduction

The Himalayan–Karakoram Mountain (HKM) in northern Pakistan are among the most important mountain ranges globally, with changing topographic and climatic characteristics [1,2]. The irregular topography makes this region extremely vulnerable to geohazards, including landslides, debris flow, glacial erosion, flash floods, river incision, and periglacial action [3,4]. These geohazards cause severe damages to human lives and infrastructure [5–8]. The China–Pakistan Economic Corridor (CPEC) project extends from China to Pakistan through this mountain range. This project has been a driver of the recent rapid growth of urban and rural areas in Pakistan. However, the expected development of new infrastructure in this region faces many challenges as geological surveying and assessment activities have been impeded by geohazards.

In particular, landslides and debris flow are the most prominent geohazards affecting human lives and infrastructure in this region. These hazards are considered to be catastrophic natural hazards with potential for causing significant damage to human lives, infrastructure, agricultural livestock, and forest growth; thus, they have profound effects on social activity and economic growth [9–11]. The occurrence of these geohazards varies according to different environmental settings, from mountainous to coastal areas, tectonically active to tectonically stable and inactive regions, and very dry to humid areas [12,13]. Therefore, identification of the spatial distribution of these geohazards is crucial for disaster management.

Geographical Information System (GIS) and Remote Sensing (RS) approaches are useful for determining the spatial distribution of geohazards, which facilitated the development of susceptibility maps [14–17]. GeoHazard Susceptibility Maps (GHSMs) are used to determine the relative probability of hazard occurrence by considering the causes of previous events. The term susceptibility can be defined as the probability of a particular hazard in a particular area. In our case, the probability of an area to experience landslides and debris flows occurrence [18]. In GHSMs, an area is divided into different classes depending on the probability of events, such as slope failure and mass movements. According to the availability of data, GHSMs can be prepared using different approaches, including the fuzzy logic [19–21], Multi-Criteria Decision Making (MCDM) [22], Weights-of-Evidence (WoE) [23,24], Logistic Regression (LR) [25], Support Vector Machine (SVM) [26], neuro-fuzzy [27], Artificial Neural Networks (ANNs) [28], Frequency Ratio (FR) [29], and Shannon Entropy (SE) models [30]. Among these listed models, LR, FR, WoE, and SE were found to be more accurate, easy to implement, and highly reliable [31–33].

Thus far, many researchers have performed geohazard susceptibility analyses for the development of susceptibility maps. However, only Ahmed et al. [34] have carried out a regional-level study on the susceptibility of northern Pakistan to geohazards; they used Weighted Overlay Model (WOM) and fuzzy logic methods. Some individual studies have also been conducted on smaller-scale areas in this region: Derbyshire et al. [35] studied the terrain type and geomorphic processes of a 200 km section of the Karakoram Highway (KKH) from Gilgit to Khunjerab Pass. Kamp et al. [36] developed a Landslide Susceptibility Map (LSM) using the Multi-Criteria Evaluation (MCE) and the Analytical Hierarchy Process (AHP) method. Bacha et al. [37] developed a detailed LSM of Hunza-Nagar district using WoE and FR models. Recently, Ali et al. [38] developed an LSM of the KKH from Hassanabdal to Khunjerab pass, using AHP and WOM models. Khan et al. [39] studied the spatial distribution of landslides and developed an LSM using the FR approach (Haramosh Valley and Bagrote Valley).

Although geospatial techniques have shown enormous potentials for assessing hazard susceptibility, their use remains relatively low in the context of the study area. In addition, comprehensive multi-hazard (landslides and debris flows) risk management and susceptibility mapping is lacking at a regional scale, especially in areas frequently affected by geohazards. Previously, Ahmed et al. [35] developed an LSM for the upper Indus basin (Pakistan) by considering elevation, slope, aspect, curvature, seismic hazards, geological structure, precipitation (monsoon), drainage network, and Normalized Differ-

ence Vegetation Index (NDVI). However, to generate a detailed hazard map, all factors that control their spatial distribution of geohazards need to be considered. The addition of more conditioning factors will improve the precision and accuracy of GHSMs. Though there have been several studies assessing individual landslide or debris flow susceptibility in the study area, none of the works have reported a significant role in controlling the spatial distribution of geohazards in the case study site. The present study aims to fill this gap. Furthermore, the implementation of only one model is not sufficient for the susceptibility assessment of an area because every model has drawbacks.

To address these limitations, thirteen conditioning factors, namely, slope (degree), distance to faults, geology, elevation, distance to rivers, slope aspect, distance to road, annual mean rainfall, NDVI, profile curvature, Stream Power Index (SPI), Topographic Witness Index (TWI), and Land Cover (LC), were considered in this study. Moreover, four machine learning and statistical models, SE, LR, WoE, and FR, were adopted. The main objectives of this research are (i) to analyze factors influencing the spatial distribution of geohazards (landslides and debris flows) and (ii) to compare the performances of the four models. The results of the study are not only relevant to other researchers in the field, but also policymakers, in that the study provides useful insights for managing geohazards.

2. Materials and Methods

2.1. Study Area

The study area is located at 31°0' N–38°0' N longitude and 70°0' E–80°0' E latitude, covering the upper Indus basin (Pakistan) and southwestern part of the Tashkurgan Valley (China) (Figure 1). The transboundary study area was selected based on the importance of the CPEC project. The main route of the CPEC project (KKH) passes through the northern areas of Pakistan and connects to China through the Khunjerab Pass. The main-stream channel in this area is the Indus River, which flows along the KKH, from north to south. The river originates from the Tibetan Plateau (China), entering Pakistan in the Dardistan region through Jammu Kashmir (Indian Territory), crossing the Himalayas, and finally flows into the Arabian Sea. This region is highly prone to geohazards, including debris flow and landslides, because of its uneven terrain, high seismicity, and high intensity of monsoon precipitation (200–500 mm), spanning from July to mid-September (<http://www.pmd.gov.pk/en/> (accessed on 2 February 2021)).

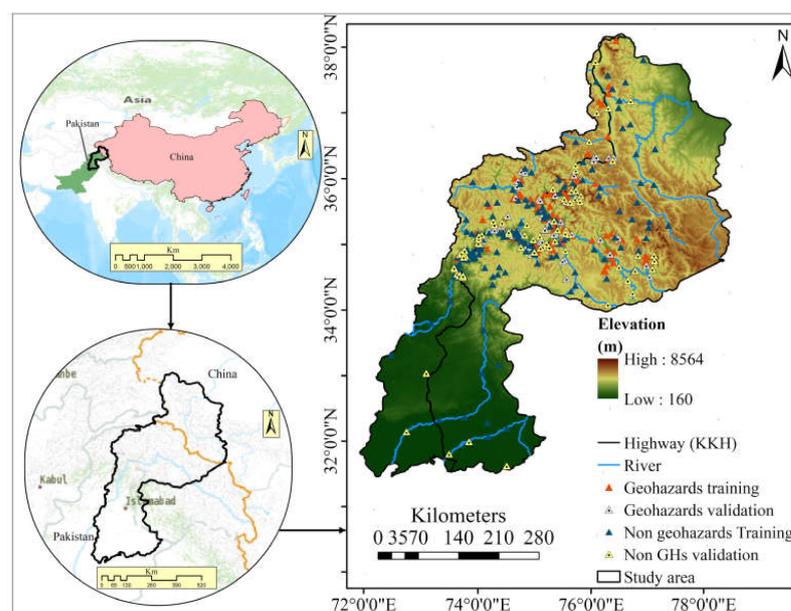


Figure 1. Location of the study area with geohazards (GHs: geohazards).

Topographically, the study area is part of the Himalaya, Hindukush, and Karakoram (HHK) mountain ranges. The elevation of the study area ranges from 160 m to 8564 m above sea level (a.s.l.), with an average altitude of 3500 m. Slope angles in this region range from 0 to 87°. The common geomorphological features in the lower reach of the valleys are alluvial fans and flood terraces with some big debris fans, while snow and glaciers cover the higher slopes and peaks.

The study area features different geological formations of various ages, exposed along the Indus River and KKH (Figure 2a and Table 1). Tectonically, this region is characterized by orogenic features that began to form with the start of the Indo-Eurasian collision (50 million years). Active faults, crustal shorting, and subduction are ongoing processes with an uplift rate of 7 mm/year and a convergence rate of 4–5 cm/year along the Indo-Eurasian collision zone [40,41]. Important tectonic features of this region include the main boundary thrust, main mantle thrust, main Karakoram thrust, Kamila Jal shear zone, Raikot fault, and Karakoram fault (Figure 2b) [42,43]. All these tectonic features are responsible for high seismicity, ultimately causing brittle deformation along the KKH. Over the last two centuries, frequent earthquakes with magnitudes greater than 6.5 Mw (1897, 1905, 1934, 1950, and 2005) have been recorded in this region [44]. The geological, geomorphological, and hydrogeological environment of the area is favorable for geohazards [34,45–47]. Valley hills in the Indus River basin exhibit an extensive amount of mass movement, rockslide avalanches (volumes $>1 \times 10^{10} \text{ m}^3$), and slope failures, especially along the KKH [48–50]. Considering that the main route of the CPEC project passes through this highly active region, susceptibility assessment of geohazards is essential.

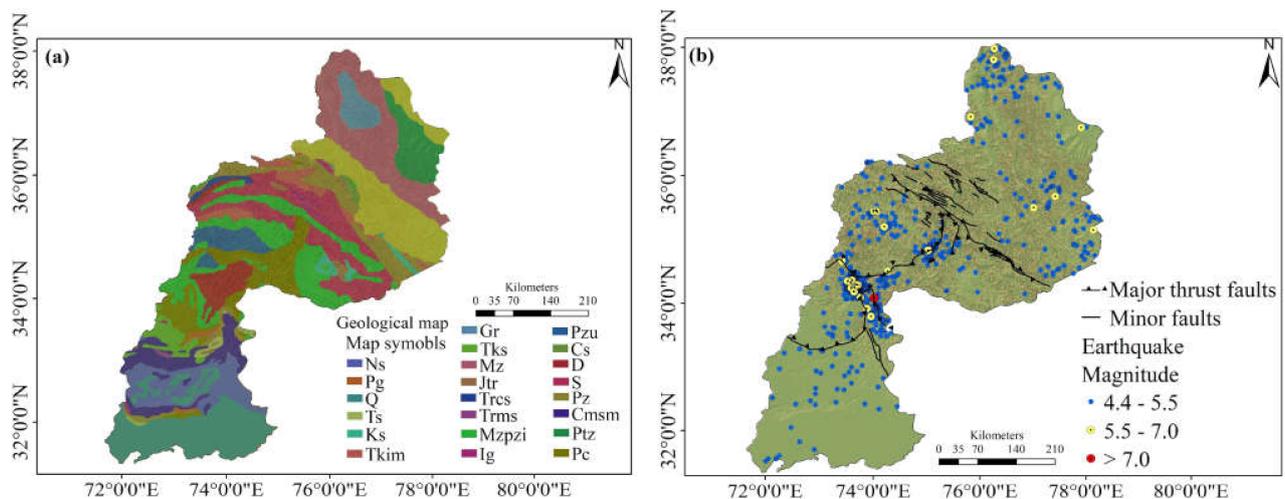


Figure 2. (a) Geological map of the study area (<http://gsp.gov.pk> and <http://en.cgs.gov.cn/> (accessed on 2 February 2021)) and (b) tectonic map with epicenter of earthquakes from last 10 decades (<http://gsp.gov.pk> and <http://en.cgs.gov.cn/> (accessed on 2 February 2021)). Ns = Neogene sedimentary rocks, Pg = Paleogene metamorphic and igneous rocks, Q = Quaternary sedimentary rocks, Ts = Tertiary meta-sedimentary rocks, Ks = Miocene to cretaceous batholith and plutons, Tkim = Miocene to cretaceous igneous and metamorphic rocks, Gr = Undifferentiated tertiary rocks, Tks = Miocene to cretaceous metamorphic rocks, Mz = Jurassic and cretaceous sedimentary rocks, Jtr = Jurassic and Triassic igneous rocks, Trcs = Lower Triassic to upper carboniferous igneous rocks, Trms = Mesozoic to Paleozoic meta-sedimentary rocks, Mzpz = Mesozoic and Paleozoic intrusive igneous and metamorphic rocks, Ig = Permian to carboniferous igneous and metamorphic rocks, Pzu = Upper Paleozoic igneous rocks, Cs = Carboniferous meta-sedimentary and sedimentary rocks, D = Devonian igneous rocks, S = Undivided Silurian rocks, Pz = Paleozoic metamorphic rocks, Cmsm = Cambrian Sedimentary and metamorphic rocks, Ptz = Archean to Proterozoic metamorphic rocks, Pc = Undivided Precambrian.

Table 1. Geological formations of the study area with major rock units and dominant lithology (<http://gsp.gov.pk> and <http://en.cgs.gov.cn/> (accessed on 2 February 2021)).

Map Symbol (Figure 2)	Geological Age	Major Rocks	Dominant Lithology
Ns	Neogene	Sedimentary rocks	Sandstone, siltstone, conglomerates and mudstone
Pg	Paleogene	Metamorphic and igneous rocks	Mostly Granite and finely foliated gneisses
Q	Quaternary	Sedimentary rocks	Sandstone, siltstone, conglomerates and loess
Ts	Tertiary	Meta-sedimentary rocks	Garnet-bearing graphitic phyllite, and marble
Ks	Miocene to cretaceous	Batholith and plutons	Granite, granodiorites, hornblende gabbro
Tkim	Miocene to cretaceous	Igneous and metamorphic rocks	granite, granodiorites, diorite and granitic gneiss
Gr	Tertiary	Undifferentiated	Granites
Tks	Miocene to cretaceous	Metamorphic rocks	Younger tertiary Gabbros, diorite and granite and pegmatites
Mz	Jurassic and cretaceous	Sedimentary rocks	Sandstone, shale and limestone
Jtr	Jurassic and Triassic	Igneous rocks	Arc-related calc-alkaline andesite and volcanic formations
Trcs	Lower Triassic to upper carboniferous	Igneous rocks	Alkaline and tourmaline Granite, Syenite with miner carbonates
Trms	Mesozoic to Paleozoic	Meta-sedimentary rocks	Sandstone, shale quartzite, limestone, slates, phyllite
Mzpsi	Mesozoic and Paleozoic	Intrusive igneous and metamorphic rocks	Volcanic rocks, limestone, red shale, sandstone, quartzite, conglomerate
Ig	Permian to carboniferous	Igneous and metamorphic rocks	Alkaline and tourmaline granite, syenite with miner carbonates with medium to low grade metamorphic rocks
Pzu	Upper Paleozoic	Igneous rocks	Norite, Pyroxene-gabbros, dunite and peridotites
Cs	Carboniferous	Meta-sedimentary and sedimentary rocks	Slates, phyletic slates with subordinate limestone and quartzite
D	Devonian	Igneous rocks	Mafic and ultramafic rocks
S	Undivided Silurian	Metamorphic and igneous rocks	Volcano-clast sediments metamorphosed green schist, interbedded slates
Pz	Paleozoic	Metamorphic rocks	Slate, phyllite and quartzite with gabbro, diorite and granite intrusions
Cmsm	Cambrian	Sedimentary and metamorphic rocks	Granite and gneisses
Ptz	Archean to Proterozoic	Metamorphic rocks	Gneiss, quartzite and marble
Pc	Undivided Precambrian	Metamorphic Rocks	Schistose to Phyllitic quartzite, schist, slate marble

2.2. Data Collection

The first step for the development of GHSMs is the determination of causative factors. Based on previous studies and the characteristics of geohazards in the study area [29,34,38], NDVI, TWI, distance to faults, elevation, annual mean rainfall, slope (degree), slope aspect, profile curvature, distance to rivers, distance to road, LC, geology, and SPI were selected as conditioning factors (Figure 3). The elevation of the study area was extracted from a Shuttle Radar Topography Mission (SRTM) 30 m digital elevation model (<https://earthexplorer.usgs.gov/> (accessed on 2 February 2021)), and the slope, slope aspect, profile curvature, SPI, TWI, and drainage network layers were developed from the elevation map in the ArcGIS platform. The annual mean rainfall map was prepared based on 27 weather stations (from 2000 to 2015) in the study area, using the Inverse Distance Weighted (IDW) tool in ArcGIS [51,52]. Geology and faults of the study area were collected from a regional geological map of the study area (at the scale of 1:1,000,000) (<http://gsp.gov.pk> and <http://en.cgs.gov.cn/> (accessed on 2 February 2021)). A brief description of these conditioning parameters is listed in Table 2.

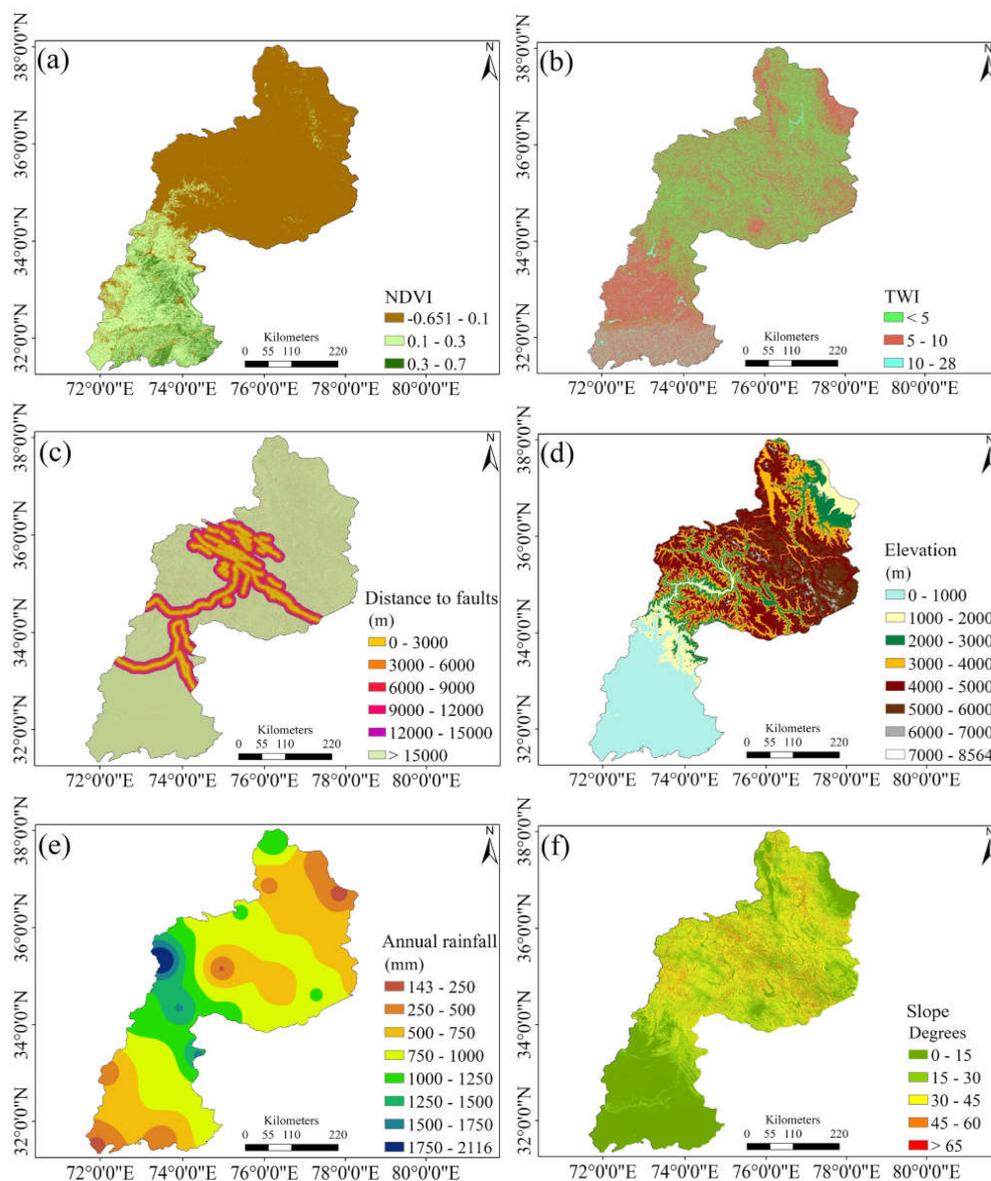


Figure 3. Cont.

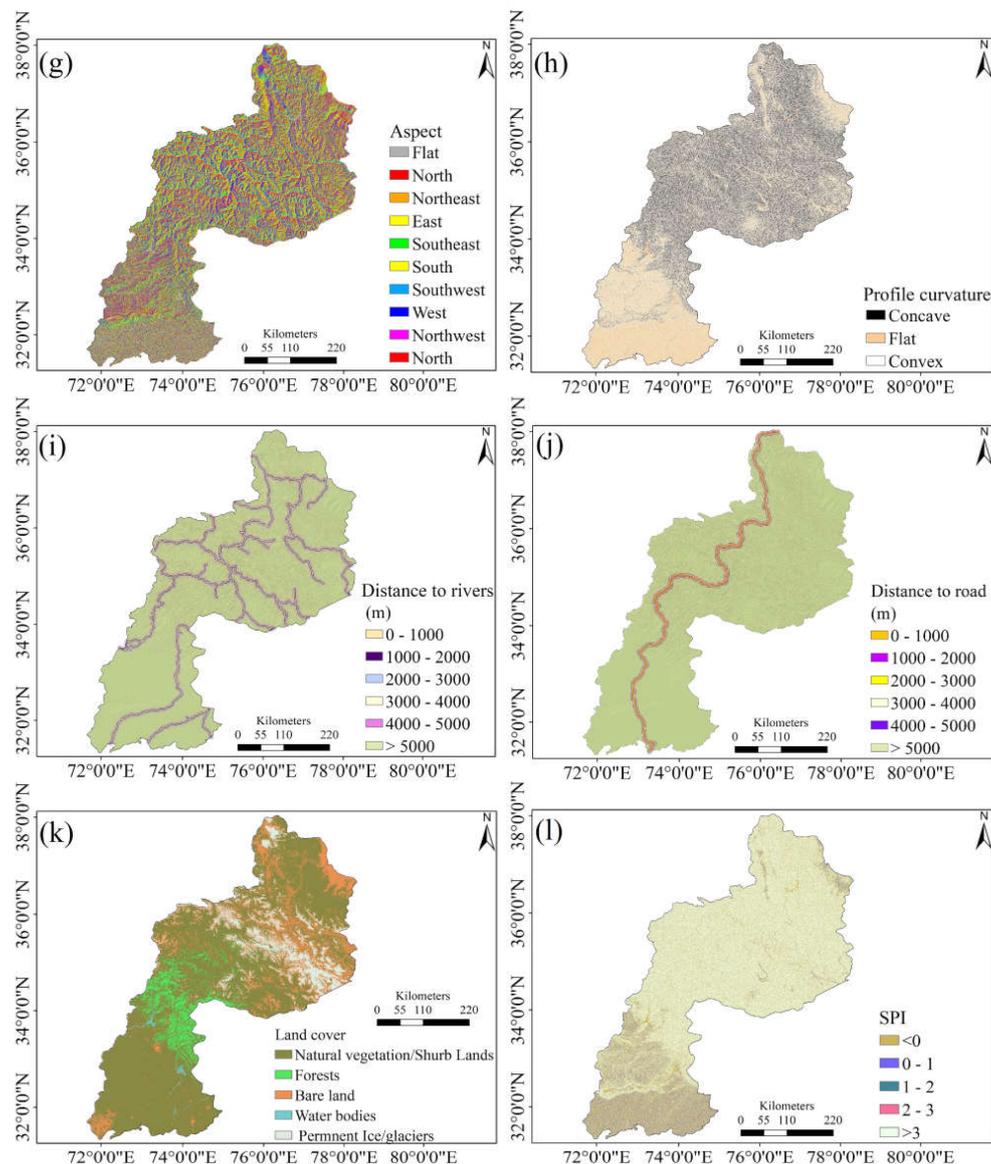


Figure 3. Conditioning factors used for geohazards susceptibility maps: (a) NDVI, (b) TWI, (c) Distance to faults, (d) Elevation, (e) Annual mean rainfall, (f) Slope, (g) Slope aspect, (h) Profile curvature, (i) Distance to rivers, (j) Distance to road, (k) Land cover, and (l) SPI.

2.3. Geohazards Inventory

A geohazard inventory map was prepared after a field visit, using the interpretation of satellite imageries (Landsat 7 and Landsat 8: <https://earthexplorer.usgs.gov/> (accessed on 2 February 2021)) and Google Earth images. During the field survey (from 16 October to 26 October 2019), the coordinates of the geohazards were taken and these coordinate points were digitized as polygons. The total numbers of sample points are 480, of which 240 are geohazard points (210 landslides and 30 debris flow events), and 240 are non-hazard points (Figure 1). All these points were used as sampling points. For both landslide and debris flow inventories, hazard pixels were assigned a value of 1, and non-hazard pixels were assigned a value of 0. To evaluate the predictability, field verified data were potentially used. Finally, the total hazard events were randomly divided into training models (70%, comprising 168 hazard points) and validation models (30%, comprising 72 hazard points) [53].

Table 2. Description of the conditioning factors used in geohazard susceptibility analysis.

Factors	Description	Source/Scale/Resolution	Reference
NDVI	Bare soil/rock is showing maximum exposure to weathering agents as compared to surfaces covered by vegetation.	$NDVI = \frac{(NIR-RED)}{NIR+RED}$ 30 m × 30 m	Li et al. [54], Intarawichian and Dasananda [55]
TWI	Higher TWI values indicate the infiltration of water bodies into slope-forming materials and decrease the shear strength of the slope.	$TWI = \log\left(\frac{A_s}{\beta}\right)$ 30 m × 30 m	Hengl and Reuter [56]
Distance to faults	Tectonic features affect the internal cohesion of the strata and hence reduce their strengths, weakening the geo-mechanical properties, and facilitating the occurrence of slope failures.	Geological survey of Pakistan, and geological survey of China 1:1,000,000	Peduzzi [57]
Elevation	3D topographic representative and is quite useful in the preparation of many geomorphological variants for hazard mapping	30 m × 30 m https://earthexplorer.usgs.gov	Ahmed et al. [35]
Annual mean rainfall	An important environmental triggering factor of landslides and debris flows	Pakistan Metrological Department	Hearn and Hart [58]
Slope (degrees)	Slope angle has a significant importance in influencing slope failure mechanisms.	Extracted from DEM 30 m × 30 m	Kayastha et al. [59]
Slope aspect	Slope aspect has an intensified involvement in geohazards as it affects factors such as weathering, exposure to sunlight, snow melt water, precipitation, soil conditions, and land cover	Extracted from DEM 30 m × 30 m	Gao and Sang [60]
Profile curvature	Affects the flow pattern on a slope and influence surficial erosion by converging or diverging the flow of runoff down the slope at a particular location	Extracted from DEM 30 m × 30 m	Kayastha et al. [59]
Distance to road	The presence of road/highway can cause many geohazards mainly due to changes in slopes stability and shear stress resulting from the excavation, undercutting of the slope, changes in hydrological conditions, and additional loads. The KKH is the main route of the CPEC project that connects Pakistan to China. The KKH is passing through northern Pakistan, where the reconstruction and re-routing are still ongoing. Therefore, only KKH was considered for this study to prepare the distance to the road thematic map.	Extracted from Google earth	Acharya and Lee [29]
Distance to rivers	River channels affect slope stability by slope toe erosion and saturating the lower part of the material.	Extracted from Google earth	Myronidis et al. [61]
Land cover	Forest plays a key role in minimizing hazards risk through various mechanisms. (roots reinforce soil layers, lowering soil moisture levels and transpiration)	http://maps.elie.ucl.ac.be/CCI/viewer/download.php 300 m × 300 m	Dolidon et al. [62] and Forbes et al. [63]
Geological map	Slope-forming materials are characterized by lithology that affects the strength and permeability of the slope.	Geological survey of Pakistan, and geological survey of China Scale- 1:1,000,000	Dai and Lee [64]
Stream Power Index (SPI)	The incidence of SPI shows flow conditions with a medium degree of erosion, a feature that is linked with abrupt topographical changes and steep slopes.	$SPI = A_s \times \tan\beta$ 30 m × 30 m	Valencia Ortiz and Martínez-Graña [65]

2.4. Methodology

2.4.1. Multicollinearity

The Tolerance (TOL) and Variance Inflation Factor (VIF) were used to test the multicollinearity of all conditioning parameters, as linear collinearity among the conditioning parameters decreases the predictive capability of a model [66]. Generally, a TOL less than 0.10 or a VIF greater than 5.0 indicates multicollinearity among the conditioning factors [67].

Multicollinearity analysis was performed, and the values of TOL and VIF were found to be >0.10 and <5.0, respectively, which indicate that the variables are free from collinearity (Table 3).

Table 3. Multicollinearity diagnosis for conditioning factors.

Factors	Collinearity Test		Factors	Collinearity Test	
	TOL	VIF		TOL	VIF
NDVI	0.849	1.178	Profile curvature	0.673	1.486
TWI	0.430	2.327	Distance to rivers	0.576	1.735
Distance to faults	0.760	1.316	Distance to road	0.701	1.427
Elevation	0.462	2.163	LC	0.809	1.236
Rainfall	0.732	1.365	Geology	0.885	1.130
Slope	0.480	2.082	SPI	0.547	1.829
Slope aspect	0.892	1.121	-	-	-

2.4.2. Logistic Regression (LR)

LR is a multivariate statistical approach used to evaluate the association between dependent variable and independent variables [68,69]. The advantage of this model is that the data need not have a normal distribution and the variables can be either categorical, continuous, or have any combination of these two [70,71]. The mathematical expression of the LR model is given below [26]:

$$\rho = \frac{1}{1 + e^{-z}} = \frac{1}{[1 + e\{- (l_0 + l_{1 \times 1} + l_{2 \times 2} + \dots l_{n \times n})\}]} \quad (1)$$

where ρ is the probability of hazards or non-hazards occurrence, z is the linear combination, n represents the number of conditioning variables, x_i ($i = 0, 1, 2, \dots, n$) are the conditioning variables, l_0 is the intercept of this model, and l_i ($i = 0, 1, 2, \dots, n$) are the regression coefficients for the independent variables of the LR model.

2.4.3. Weights-of-Evidence (WoE)

WoE is a quantitative “data-driven” approach [72]. This method was first used in the field of medicine [73] and then applied to the identification of mineral potential [74]. WoE is based on Bayes’ hypothesis—the concept of previous and subsequent probability to predict the relationship between the spatial distribution of hazard affected areas and analyzed hazard factors as reported by Bonham-Carter [75] as follows:

$$W_h^+ = \ln \frac{p\left(\frac{B_h}{A_h}\right)}{p\left(\frac{B_h}{\bar{A}_h}\right)} \quad (2)$$

$$W_h^- = \ln \frac{p\left(\frac{\bar{B}_h}{A_h}\right)}{p\left(\frac{\bar{B}_h}{\bar{A}_h}\right)} \quad (3)$$

where p is the probability of occurrence of a particular event, and \ln is the natural log. Similarly, B_h denotes the presence of a potential geohazard predictive factor, \bar{B}_h is the absence of a potential geohazard predictive factor, A_h is the presence of geohazards, and \bar{A}_h denotes the absence of geohazards. In Equations (2) and (3), W_h^+ and W_h^- show the relationship of each factor's class to geohazard occurrence, where positive weights are directly proportional for the occurrence of geohazards and vice-versa with negative weights. To calculate the weights of each conditioning factor, Equations (4) and (5) were used [76]:

$$W_h^+ = \ln \left\{ \left(\frac{[N_{pix1}]}{[N_{pix1}] + [N_{pix2}]} \right) / \left(\frac{[N_{pix3}]}{[N_{pix3}] + [N_{pix4}]} \right) \right\} \quad (4)$$

$$W_h^- = \ln \left\{ \left(\frac{[N_{pix1}]}{[N_{pix1}] + [N_{pix2}]} \right) / \left(\frac{[N_{pix3}]}{[N_{pix3}] + [N_{pix4}]} \right) \right\} \quad (5)$$

where N_{pix1} is the number of pixels showing the presence of geohazards and conditioning factors; N_{pix2} is the presence of geohazards and absence of conditioning factors; N_{pix3} is the absence of geohazards and the presence of conditioning factors; N_{pix4} is the absence of both geohazards and conditioning factors. The final weight W_hC was estimated using the following equation:

$$W_hC = (W_h^+) - (W_h^-) \quad (6)$$

2.4.4. Frequency Ratio (FR)

FR is a widely used probability-based approach for GHSM. The FR is defined as the ratio of the probability of an event occurrence to the probability of non-occurrence for given attributes [75,77]. The spatial relationship between the location of geohazards and each parameter was obtained. To obtain the FR values for each class of the conditioning parameter, a combination between the geohazard inventory map and factors map was established through Equation (7).

$$FR = \frac{Npix(HX_i) / \sum_{i=1}^m Npix(HX_i)}{Npix(NX_j) / \sum_{j=1}^n Npix(NX_j)} \quad (7)$$

where FR represents the Frequency Ratio (FR) of the class i of factor j ; $Npix(HX_i)$ is the number of pixels with geohazards within class i of the conditioning factor X ; $Npix(NX_j)$ is the number of pixels within the conditioning factor X_j ; m is the number of classes in the conditioning factor X_j ; n is the number of factors in this area [29].

The derived frequency ratio was used to calculate relative frequency (Equation (8)). Further, the predictive rate was calculated using Equation (9) [29].

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

$$PR = \frac{(RF_{max} - RF_{min})}{(RF_{max} - RF_{min})_{min}} \quad (9)$$

Finally, the GeoHazards Susceptibility Index (GHSI) was calculated using Equation (10).

$$GHSI = \sum (PR \times RF) \quad (10)$$

2.4.5. Shannon Entropy (SE)

The SE model is based on the instability, disorder, imbalance, and uncertainty of a system. SE indicates the most favorable class of a variable in a natural environment for hazard occurrence [78]. Entropy values were used for calculating the target weights of the

index system. The mathematical expression of weight, W_j , was used for calculating the information coefficient for this model (as a whole variable weight value) as follows [79]:

$$P_{ij} = \frac{FR}{\sum_{j=1}^{s_j} FR} \quad (11)$$

where, P_{ij} is the probability density and FR is the frequency ratio. Furthermore, based on Equation (11), each class of the factor is divided into different ranks (Table 4).

$$H_j = - \sum_{j=1}^{s_j} P_{ij} \log_2 P_{ij}, \quad j = 1, \dots, n \quad (12)$$

H_j and H_{jmax} represent entropy values in Equations (12) and (13)

$$H_{jmax} = \log_2 s_j, \quad s_j - \text{number of classes} \quad (13)$$

$$I_j = \frac{H_{jmax} - H_j}{H_{jmax}} \quad I = (0,1), j = 1, \dots, n \quad (14)$$

$$W_j = I_j P_j \quad (15)$$

I_j represents the information coefficient and W_j is the final weight value of the variables as a whole.

Table 4. Data used and spatial relationship between each conditioning parameter with geohazards derived using the Logistic Regression, Weight-of-Evidence, Frequency Ratio, and Shannon Entropy models.

Factor	Class	Values	FR	PR	LR Coefficients	$W_i C$	Wj (%)	Ranks
NDVI	(i)	−0.651–0.1	1.33			1.800	0.298	3
	(ii)	0.1–0.3	0.27	5.62	0.158	−1.497		2
	(iii)	0.3–0.7	0.01			−4.670		1
TWI	(i)	0.44–5	1.57			0.761	0.164	3
	(ii)	5–10	0.82	3.45	−0.494	−0.418		2
	(iii)	10–28	0.25			−1.483		1
Distance to faults (m)	(i)	0–3000	0.54			−1.026	0.372	5
	(ii)	3000–6000	2.11			0.831		4
	(iii)	6000–9000	1.20	1.22	−0.120	0.199		1
	(iv)	9000–12,000	1.82			0.647		3
	(v)	12,000–15,000	1.97			0.722		2
	(vi)	>15,000	1.26			0.327		6
Elevation (m)	(i)	0–1000	0.01			−4.600	0.571	3
	(ii)	1000–2000	3.88			1.618		7
	(iii)	2000–3000	4.25			1.892		8
	(iv)	3000–4000	1.45	2.98	−0.508	0.468		6
	(v)	4000–5000	0.15			−2.146		5
	(vi)	5000–6000	0.04			−3.281		4
	(vii)	6000–7000	0			−0.010		0
	(viii)	7000–8557	0			−0.001		0

Table 4. Cont.

Factor	Class	Values	FR	PR	LR Coefficients	$W_i C$	W_j (%)	Ranks
Rainfall (mm)	(i)	0–250	0			−0.010		0
	(ii)	250–500	1			−0.003		4
	(iii)	500–750	0.93			−0.112		7
	(iv)	750–1000	1.00			−0.001		8
	(v)	1000–1250	0.78	2.69	−0.388	−0.283	0.186	5
	(vi)	1250–1500	2.55			1.026		6
	(vii)	1500–1750	0.26			−1.344		3
	(viii)	1750–2157	0			−0.010		0
Slope angle (°)	(i)	0–15	0.17			−2.252		2
	(ii)	15–30	1			0.001		4
	(iii)	30–45	2	1.84	1.243	1.130	0.154	5
	(iv)	45–60	2.03			0.796		3
	(v)	>60	1.83			0.609		1
Slope aspect	(i)	Flat	0		0	−5.631		1
	(ii)	North	0.94		−20.071	−0.067		2
	(iii)	Northeast	0.74		−0.664	−0.339		4
	(iv)	East	1.04		−1.457	0.040		8
	(v)	Southeast	1.02		−1.062	0.026		7
	(vi)	South	0.96	1.00	−1.254	−0.043	0.049	6
	(vii)	Southwest	1.31		−0.631	0.313		10
	(viii)	West	1.12		−0.915	0.129		9
	(ix)	Northwest	0.94		−0.257	−0.067		5
	(x)	North	1.38		−0.425	0.344		3
Profile curvature	(i)	Concave (−1)	1.54			0.567		2
	(ii)	Flat (0)	0.6	1.99	−0.038	−0.730	0.065	1
	(iii)	Convex (+1)	1.11			0.190		3
Distance to river (m)	(i)	0–1000	5.92			2.001		5
	(ii)	1000–2000	3.74			1.430		4
	(iii)	2000–3000	2.22	2.19	0.031	0.844	0.286	2
	(iv)	3000–4000	2.23			0.847		1
	(v)	4000–5000	2.54			0.985		3
	(vi)	>5000	0.47			4.833		6
Distance to roads (m)	(i)	0–1000	8.26			2.208		5
	(ii)	1000–2000	5.01			2.663		2
	(iii)	2000–3000	4.86	1.70	−0.005	1.630	0.374	1
	(iv)	3000–4000	6.33			1.912		4
	(v)	4000–5000	5.53			1.767		3
	(vi)	>5000	0.68			−2.177		6
LC	(i)	Natural vegetation/Shrub land	1.18		2.878	0.663		5
	(ii)	Forest	1.88	3.59	1.644	0.690	0.240	4
	(iii)	Bare land	0.42		2.256	−0.982		3
	(iv)	Water bodies	0		0	−0.004		0
	(v)	Ice/glaciers	0.13		−17.846	−2.140		2

Table 4. Cont.

Factor	Class	Values	FR	PR	LR Coefficients	W_hC	Wj (%)	Ranks
Geology	(i)	Jtr	0.93		21.561	−0.077		13
	(ii)	Ks	1.44		21.616	0.367		14
	(iii)	Mzpzi	2.20		0.926	0.984		21
	(iv)	Cs	0		0	−0.003		0
	(v)	Pz	0.13		1.368	−2.201		17
	(vi)	Ig	11.27		0.643	2.501		19
	(vii)	Tkim	0.89		2.339	−0.102		16
	(viii)	Pc	0.05		2.152	−2.968		12
	(ix)	Trcs	0		0	−0.005		0
	(x)	D	0.81		−0.694	−0.215		18
	(xi)	Pzu	6.41		0.082	2.082		20
	(xii)	Tks	0	2.89	0	−0.019	0.651	0
	(xiii)	Ts	0		0	−0.006		0
	(xiv)	Cmsm	0		0	−0.050		0
	(xv)	Trms	0		0	−0.001		0
	(xvi)	Pg	0		−21.675	−0.004		0
	(xvii)	Q	2.47		−19.821	1.124		22
	(xviii)	Ns	0		−20.141	−0.058		0
	(xix)	S	0		−18.611	−0.107		0
	(xx)	Gr	0		0.995	0.023		0
	(xxi)	Ptz	0		0	−0.042		0
	(xxii)	Mz	0.19		−19.569	−1.738		15
SPI	(i)	<0	0.80			−0.333		4
	(ii)	0–1	0			−7.318		0
	(iii)	1–2	0	4.10	−0.120	−6.054	0.223	2
	(iv)	2–3	0.02			−0.868		3
	(v)	> 3	1.22			0.559		5

2.4.6. Model Validation

In this study, derived maps were validated using geohazard inventory points, which were split into two categories: training data (70%) and validation data (30%). The Area Under the Receiving Operative Characteristic (AUROC) approach was adopted to evaluate the performance of the standalone models for geohazard susceptibility assessment [80,81].

3. Results and Discussion

3.1. Application of Logistic Regression

The LR analysis revealed the most effective variables influencing the distribution of geohazards. With *sig* (*p*) values less than 0.05, the most effective variables were found to be distance to faults, slope, elevation, geology, LC, and rainfall (Figure 4). The remaining conditioning factors were insignificant, with higher *sig* (*p*) (>0.05) values [26,82]. Furthermore, the positive and negative coefficients of LR indicate their contribution to the occurrence of geohazards. Based on the obtained coefficients, the probability of geohazards was calculated using Equation (16), and the GHSM was generated accordingly.

$$\begin{aligned}
 z = & -1.054 + (\text{NDVI} \times 0.158) + (\text{TWI} \times -0.494) + LC_a + (\text{Rainfall} \times -0.388) \\
 & + (\text{Elevation} \times -0.508) + (\text{Faults} \times -0.120) + (\text{River} \times 0.031) \\
 & + (\text{Profile curvature} \times -0.038) + \text{Geology}_a + (\text{Road} \times -0.005) \\
 & + \text{Aspect}_a + (\text{SPI} \times -0.120) + (\text{Slope} \times 1.243)
 \end{aligned} \quad (16)$$

LC_a ; Geology_a ; Aspect_a are logistic regression coefficient values listed in Table 4.

Ultimately, the GHSI for LR was calculated, using Equation (16). The probability value varies from 0 to 1 [25]. The final SI values were reclassified into five classes: very-low, low,

moderate, high, and very-high (Figure 5), using the natural breaks method [83]. In this model the percentage of very-high, high, moderate, low and very low susceptibility areas are 11.19%, 9.24%, 10.18%, 39.14%, and 30.25%, respectively, as shown in Figure 6.

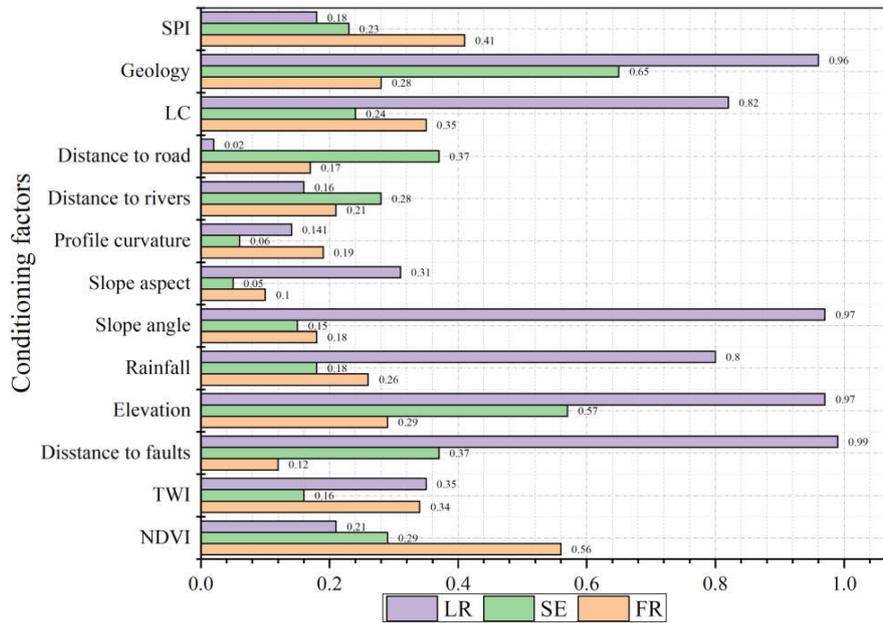


Figure 4. Factors controlling spatial distribution of geohazards obtained via LR, SE, and FR models.

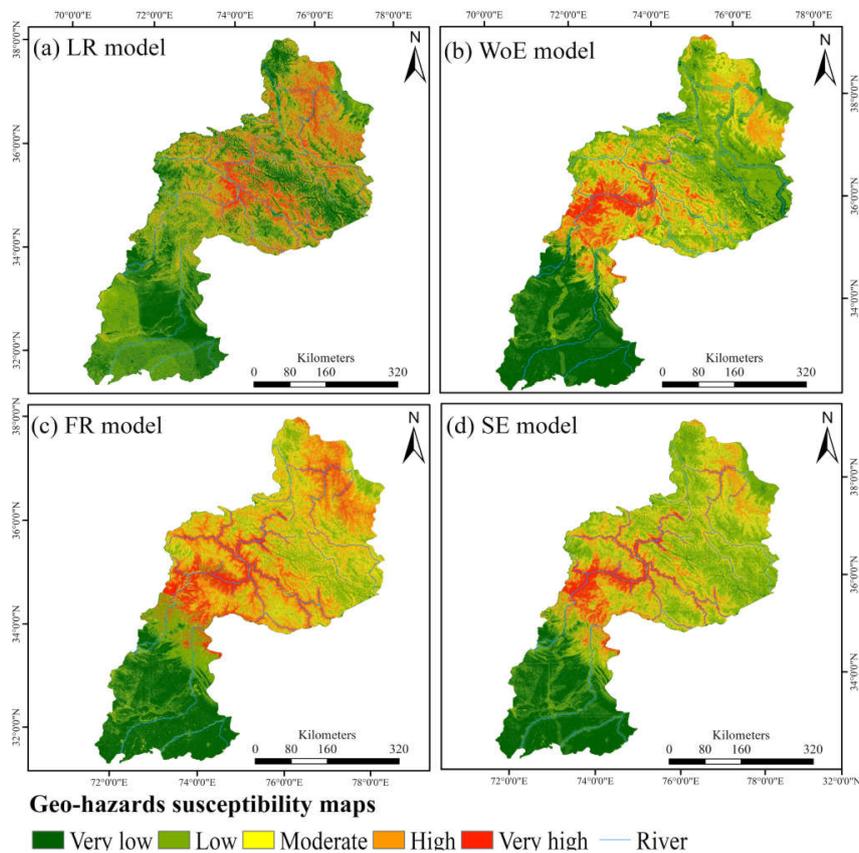


Figure 5. Geohazard susceptibility maps developed using: (a) Logistic Regression, (b) Weight-of-Evidence, (c) Frequency Ratio, and (d) Shannon Entropy models.

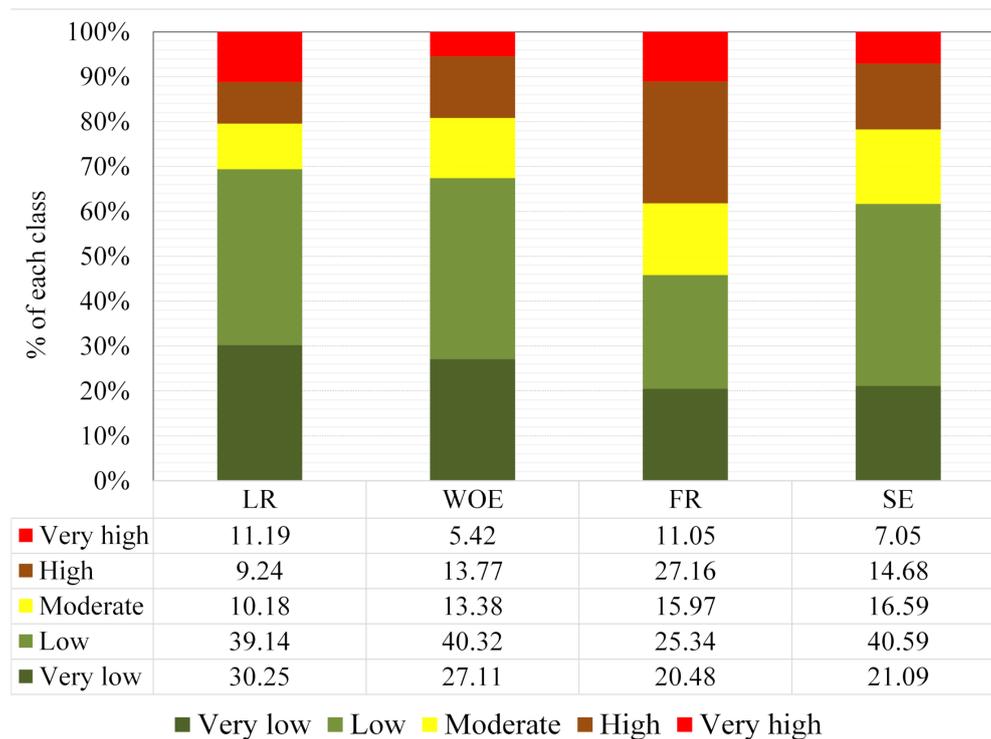


Figure 6. Percentage of geohazard susceptibility classes from the Logistic Regression, Weight-of-Evidence, Frequency Ratio, and Shannon Entropy models.

3.2. Application of Weights-of-Evidence

The weight and contrast were calculated for each class of conditioning variables using the WoE model. The output of total weight indicated the significance of each factor as shown in Table 4. In this model, distance to road, distance to faults, geology, elevation, and slope showed higher magnitudes of W_hC (Figure 7). The class of distance to road with a distance interval of 0 to 1000 m showed high susceptibility to geohazards, followed by the class with a distance interval of 3000 to 4000 m, with W_hC values of 2.663 and 1.912, respectively. The presence of the road can lead to frequent geohazards mainly due to changes in slope stability and shear stress resulting from the excavation and undercutting of fragile slopes [84,85]. In the case of distance to faults, distance intervals from 3000 to 6000 m, 9000 to 12,000 m, and 12,000 to 15,000 m showed weights (W_hC) of 0.831, 0.647 and 0.722, respectively, showing a higher probability of geohazards. Elevation intervals of 1000–2000, 2000–3000, and 3000–4000 m showed high (W_hC) values of 1.618, 1.892, and 0.468, respectively. The investigation of lithological units revealed that the Ig and Pzu groups (composed of slate, phyllite, and shale) have higher W_hC values (2.501 and 2.082), followed by the group Q, Mzpz, and Ks groups (1.124, 0.984, and 0.367). These findings suggest that rock units with the lowest shear strength (gneisses, phyllites, slate, and shale) are prone to mass wasting. The derived W_hC weights (−2.252 to 1.130) showed that slope factor has a significant impact on the occurrence of geohazards, which is in line with the findings of a previous study [70]. Slope classes of 30–45°, and 45–60° showed higher W_hC weights, with values of 1.130 and 0.796, respectively. Gentle slopes are generally believed to experience a low frequency of geohazards because of their lower shear stresses and low gradients [86]. Distance to rivers with buffer zones of 0–1000 m and >6000 m showed higher susceptibility to geohazards, with W_hC values of 2.001 and 4.833, respectively. In terms of profile curvature, concave surfaces (negative curvature) with high W_hC values (0.567) are more prone to geohazards, compared to convex surfaces (positive curvature). Concave surfaces have higher W_hC values than convex surfaces because surfaces with negative curvature retain more water for longer periods after heavy rainfall than surfaces

with a positive curvature [29]. As a result, concave surfaces have increased soil moisture content with longer saturation periods, promoting erosion and decreasing soil stability [59]. In this model, the weights of LC, rainfall, slope aspect, SPI, TWI, and NDVI were of less importance because of very low W_hC values (0 or less than 0) (Table 4). The final susceptibility map was divided into five classes from very low to very high (Figure 5); the percentages of these classes are shown in Figure 6.

3.3. Application of Frequency Ratio

The Predictive Ratio (PR) was calculated from the frequency ratio and relative frequency, respectively. NDVI, SPI, LC, TWI, and elevation exhibited relatively high values of PR, followed by geology, distance to faults, and rainfall (Figure 4). Distance to rivers, slope, profile curvature, and distance to road have a moderate impact on the prediction of geohazard occurrence. In this model, the slope aspect was the least significant factor with a low PR value. The PR values ranged from 1.00 to 5.62, and they were directly proportional to susceptibility, with higher values indicating higher hazard susceptibility. Regarding NDVI, barren surfaces showed a much high PR value of 5.62 (Table 4). Furthermore, these areas are extremely erodible, and thus exhibit higher SPI values. Compared with natural vegetation and shrublands, forest areas showed lower susceptibility to geohazards. TWI reflects the relationship between the accumulation of water at any location and the gravitational force acting on an inclined surface toward a downslope stream [56]. In this model, TWI showed an inverse relationship with geohazards; TWI values less than five indicate a high frequency of geohazards in this region. Decreasing distance from rivers and roads was proportional to susceptibility, with shorter distances showing higher PR values. River channels have significant effects on fluvial processes that influence bedrock incision in seismically active mountain ranges [87]. In this study, we observed that most of the geohazards in this region occurred along the main Indus River and associated tributaries, which are in line with the study of Ahmed, Rogers and Ismail [34]. The PR values calculated for geology and distance to faults were high. Exposed outcrops in the study area were intensely sheared, folded, and faulted due to seismicity, resulting in high susceptibility to geohazards. Geohazards have most frequently been observed in the Misgar Slates and exposed outcrop of the Jijal complex along the river. Similarly, the elevation and slope play a key role in the occurrence of geohazards. Elevations of 2000–3000 m showed higher susceptibility. Slope angles of 45–60° showed the highest PR value (2.03), followed by slope angles of 30–45° (2.00). After determining the PR values of all conditional factors, a GHSM was prepared using Equation (10). The final susceptibility map was reclassified into five classes (very-low, 20.18%; low, 25.34%; moderate, 15.97%; high, 27.16%; very-high, 11.04%) using an expert-based classification method [26], as shown in Figures 5 and 6.

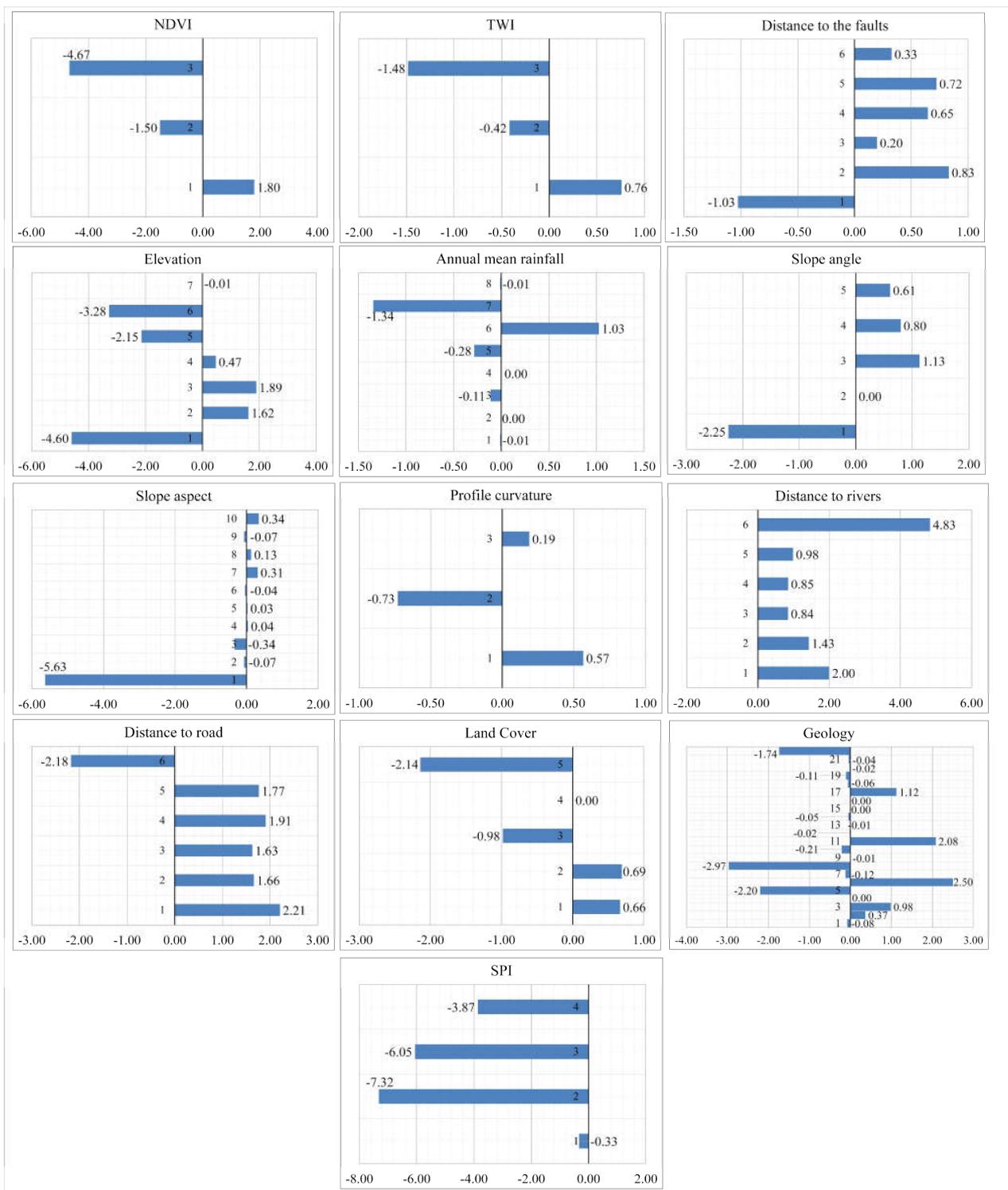


Figure 7. Factors controlling spatial distribution of geohazards obtained via WoE model.

3.4. Application of Shannon Entropy

We also used the SE model to minimize the imbalance between factors and obtain a realistic status of their effect on the susceptibility of geohazards. The relationship between the conditioning parameters and geohazards was obtained, and the most significant factors affecting the occurrence of hazards were also identified. The summarized results are listed in Table 4. According to the SE results, geology, elevation, distance to road, distance to faults, NDVI, and distance to rivers are the most significant factors controlling the occurrence and distribution of geohazards in this region (Figure 4). Therefore, these conditioning factors can better indicate the occurrence and distribution of geohazards in the study area. A GHSM was developed using the following equation [88]:

$$\begin{aligned}
 W_j = & (NDVI \times 0.298) + (TWI \times 0.164) + (LC \times 0.240) + (Rainfall \times 0.186) \\
 & + (Elevation \times 0.571) + (Faults \times 0.372) \\
 & + (Rivers \times 0.286) + (Profile\ curvature \times 0.065) \\
 & + (Geology \times 0.651) + (Road \times 0.374) + (Aspect \times 0.049) \\
 & + (SPI \times 0.223) + (Slope \times 0.154)
 \end{aligned} \tag{17}$$

Based on the SE model (Equation (11)), each class of conditional factors was divided into different ranks (Table 4). The SE-GHSM was divided into five classes from very-low to very-high susceptibility classes as shown in Figures 5 and 6.

Figure 8 shows a comparison of various results from field observations, aerial imageries, and the LR model. A, B, C, and D areas show very-high susceptibility to geohazards. In the Hunza valley, an area indicated ground subsidence and shallow landslides, which is attributable to underlying and overlying unconsolidated sediments (slate formations). The historical landslide in Attabad–Hunza blocked river flow and submerged 19 km of the KKH, developing a natural dam. Moreover, the undercutting of slopes for the reconstruction of KKH has increased the susceptibility of the region to geohazards. The entire city of Gilgit is located on a large debris fan. This section is characterized by hot and dry valleys with low precipitation rate, and some old and large debris fans were observed during the field visit. The Nanga Parbat Haramosh massif (western syntaxes) is the most susceptible area of this region, with high tectonic influence and an uplift rate of 1 cm/year [40,89,90]. The dominant tectonic features are MMT and the Raikot fault, which are responsible for the high seismicity and brittle deformation of bedrocks, ultimately leading to mass movement. The Dasu section along the KKH is strongly influenced by rainfall compared to the elevated areas.

3.5. Validation of Geohazards Susceptibility Maps

The AUROC curve was used for validating the obtained results. Success and prediction rate curves were plotted to compare the results of GHSMs for existing hazard sites. The performance of the models was evaluated by plotting AUROC curves for success rate (Figure 9a). The LR, WoE, FR, and SE models showed AUROC values of 85.30%, 76.00%, 74.60%, and 71.40%, respectively (Figure 9a). In terms of prediction rate, the AUROC values of LR, WoE, FR, and SE were 83.10%, 75.00%, 73.50%, and 70.10%, respectively (Figure 9b). The findings revealed that the GHSM prepared using the LR model had better predictive capability than those from the other three models. Other studies have also reported that the LR model has higher performance and is more suitable for susceptibility analysis compared to FR and WoE [68,70]. Our findings are in line with those of Rasyid et al. [70] in that a standalone LR model has high precision and is appropriate for hazard susceptibility assessment. Moreover, Polykretis and Chalkias [68] carried out landslide susceptibility by comparing WoE, LR, and ANN models and found that success and predication rate of LR model performed better (LR = 0.928 and 0.836, ANN = 0.912 and 0.780, WoE = 0.893 and 0.847). In another study, Benchelha et al. [91] found that the success and predication rate of LR is higher (91.8% and 91.1%) than ANN model (88.6% and 87.7%). Therefore, it is concluded that the LR model is the best option for mapping hazard potential under similar settings.

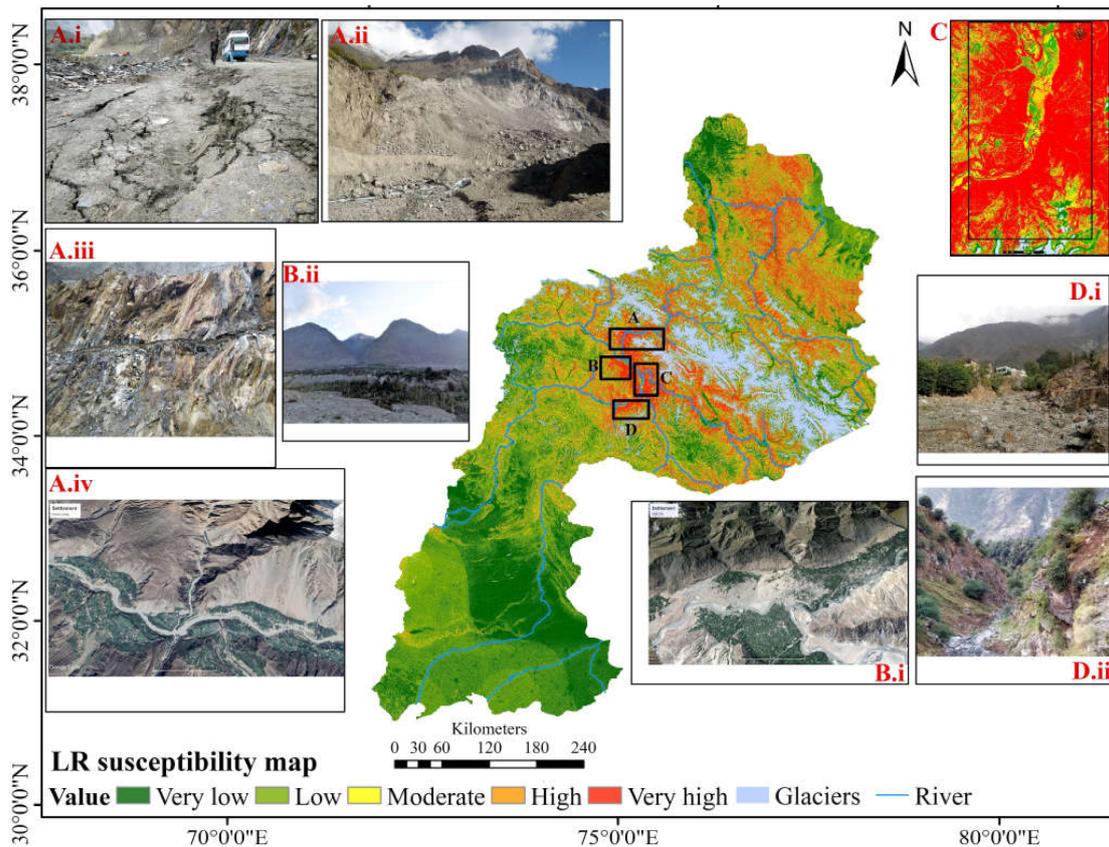


Figure 8. Comparison of ground truth data and field observation with Logistic Regression model results. (A) Overview of Hunza valley: (A.i) ground subsidence and shallow landslides, (A.ii) Attabad–Hunza, (A.iii) undercutting of slopes, (A.iv) aerial view of Hunza valley. (B) Gilgit section: (B.i) aerial view (Google earth imagery), (B.ii) debris fan. (C) Nanga Parbat Haramosh massif (western syntaxes), and (D) Dasu section.

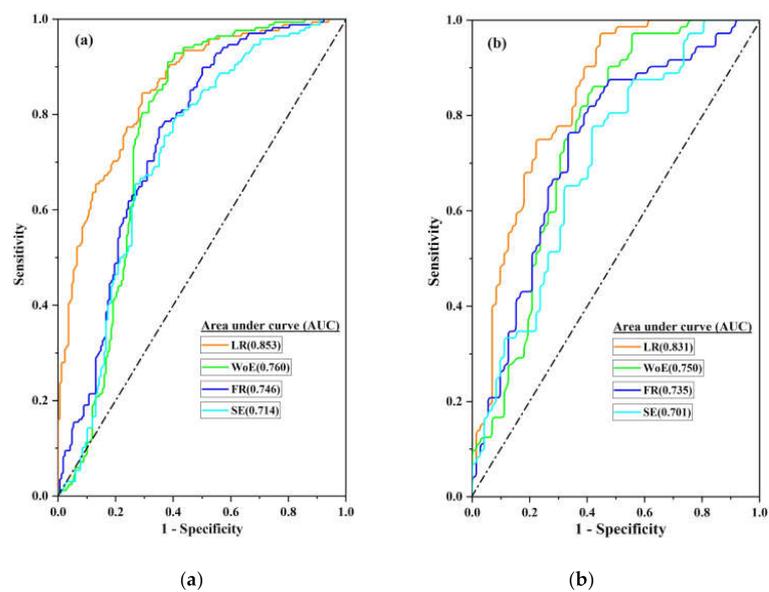


Figure 9. Comparison of area under the receiver operating curves of geohazard susceptibility maps. (a) Models training and (b) testing AUROC curve.

3.6. Model's Results Comparison

The model results were compared to assess variations between the results of each GHSM. The parameters can be primarily interpreted based on the absolute value of the correlation coefficient as follows: very weak similarities 0–20%, weak to low similarities 20–40%, moderate similarities 40–70%, strong similarities 70–90%, and very strong similarities 90–100% [92]. The similarity coefficient of LR vs. WoE was 57.70%, and those of LR vs. FR and LR vs. SE maps were 47.40%. This indicates moderate similarities of LR with WoE, and low similarities of LR with FR and SE maps (Table 5). The value of 0.898 shows a strong correlation between FR and SE (89.80%). This correlation suggests that the maps developed using the LR and WoE methods have moderate similarities, and those using the FR and SE models have a strong similarity.

Table 5. Correlation coefficients derived from the Logistic Regression, Weight-of-Evidence, Frequency Ratio and Shannon Entropy models for geohazard susceptibility maps.

Susceptibility Maps	LR	WoE	FR	SE
LR	1	57.70%	47.40%	47.40%
WoE	-	1	75.00%	81.90%
FR	-	-	1	89.80%
SE	-	-	-	1

3.7. Pixel-Wise Spatial Distribution of Susceptibility Class Comparison

To compare geohazard susceptibility maps, classified maps were ranked into five groups, i.e., very low, low, moderate, high, and very high. Then, they were compared with each of the constructed maps using a two-dimensional multiplication matrix [93,94]. The study found that the matrix's diagonal elements represent the same spatial distribution of susceptibility classes for both maps (LR–WoE, LR–FR, LR–SE, WoE–FR, WoE–SE, and FR–SE). For example, Figure 10a indicates that 39.21% of areas have been occupied by the same hazard intensity for the two susceptibility maps, e.g., susceptibility maps developed from LR and WoE. Meanwhile, 60.79% spatial distributions show dissimilarities in which 31.03% (upper elements of a diagonal matrix) distributions show the higher categories of susceptibility by the geohazard susceptibility map constructed using the WoE model. A total of 29.76% (lower elements of a diagonal matrix) distributions show the lower categories compared to the constructed map using the LR model (Figure 10a). Further, the comparison results for other models are summarized in Figure 10b–f. These differences are mainly due to the influence of factors in controlling the spatial distribution of geohazards in standalone models (Figures 4 and 5) and the proposition of the susceptibility map class [95].

The LR, WoE, FR, and SE models all indicated the significance of the causative factor on the occurrence of geohazards. In the LR model, distance to faults, slope angle, elevation, geology, LC, and precipitation were the most significant factors. The results of these analyses are supported by the findings of previous studies on this region. Ahmed, Rogers and Ismail [34] reported that geology, faults, slope angle, and rainfall are the most important factors for the occurrence of geohazards in the upper Indus basin. Rahim et al. [96] found that geology, faults, slope angle, and LC have the greatest impact on geohazards in this region. According to the WoE model, the distance from faults with distance intervals of 3000–6000 m, 9000–12,000 m, and 12,000–15,000 m have weights (W_iC) of 0.831, 0.647 and 0.722 respectively, indicating a positive correlation with the occurrence of geohazards. Slope classes of 30–45°, 45–60°, and >60° have W_iC weights of 1.130, 0.796, and 0.609, respectively. The geological investigation suggested that the Ig, Pzu, Q, Mzpz and ks groups, mainly comprising gneisses, phyllites, slate, and shale, are the most susceptible lithological units. Altitudes of 2000–3000 m showed high W_iC (1.892). The results are in agreement with the findings of other studies [37,97]. According to the FR model, NDVI, SPI, LC, TWI, elevation, geology, and distance to faults are important factors controlling

geohazards in this region. The significance of these factors has also been reported by previous studies [29,38]. According to the SE model, geology, elevation, distance to road, faults, NDVI, and rivers are the most important factors. These findings are in line with those of previous studies [29,34,38]. Furthermore, AUROC curves verified that the LR model has a higher performance than the other three models, suggesting that the LR model is the most suitable for regional-level analysis of geohazards under settings similar to those of the study area. Similarities in the influencing factors of geohazard occurrence among the developed maps were assessed. The LR and WoE models provided maps of moderate similarity, whereas the FR and SE models provided highly similar maps.

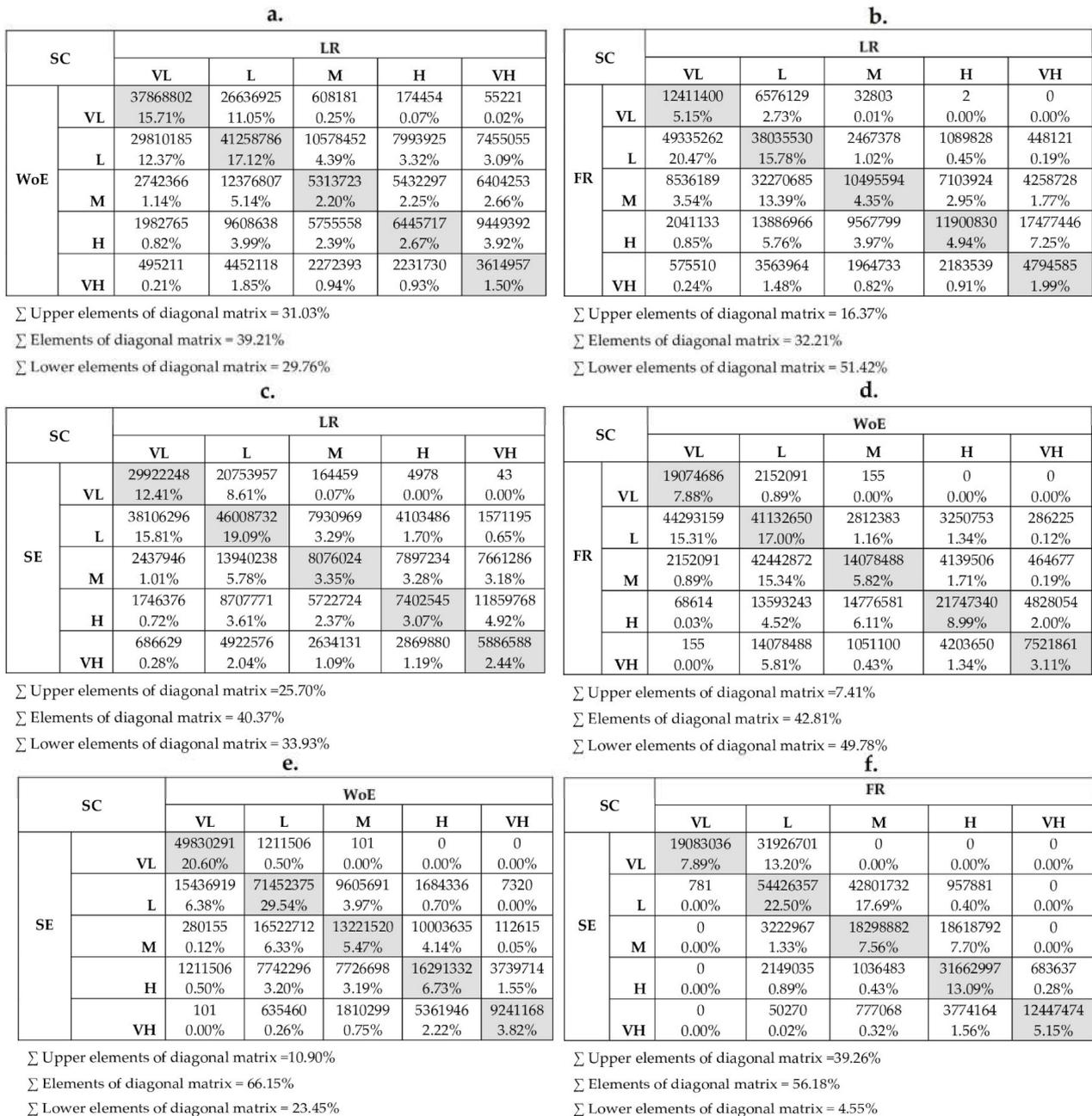


Figure 10. Comparison of susceptibility maps using a two-dimensional multiplication matrix: (a) LR vs. WoE, (b) LR vs. FR, (c) LR vs. SE, (d) WoE vs. FR, (e) WoE vs. SE, and (f) FR vs. SE. VL: very low, L: low, M: moderate, H: high, and VH: very high.

4. Conclusions

In this study, regional level geohazard susceptibility mapping from the Punjab plains (Pakistan) to the Tashkurgan Valley (China) was conducted. Thirteen conditioning factors were considered and their TOL and VIF values were calculated to verify the absence of multicollinearity for each variable. Four models, LR, WoE, FR, and SE, were used to develop geohazard susceptibility maps, and the suitability of each model was validated through AUROC curves. Furthermore, model results were analyzed to determine the statistical importance of the coefficients and the spatial association between geohazards and each class of the conditioning factors. The main conclusions of the study can be summarized as follows:

1. The LR results suggested that distance to faults, slope, elevation, geology, LC, and rainfall are the most significant parameters controlling geohazards occurrence in the study area.
2. According to the AUROC curve, the LR model showed the highest accuracy with an AUROC value of 85.30%, closely followed by the WoE model (76.00%), FR model (74.60%), and SE model (71.40%). Likewise, the prediction rate of the LR model was 83.10%. Thus, the LR model can be considered more accurate than the other three models.
3. According to the LR model, 30.25%, 39.14%, 10.18%, 9.24%, and 11.19% of the area have very low, low, moderate, high, and very high susceptibility to geohazards, respectively.

Areas susceptible to geohazards, visualized in the hazard susceptibility maps, represent important sites for geohazard assessment in the study area. The GHSM prepared using the LR model can be considered to be more realistic and suitable because of its reliability and higher accuracy. Therefore, the developed map can provide useful information for the local communities to recognize and avoid hazardous areas as well as for associated government authorities to enforce preventive measures in areas of high to very-high susceptibility.

Further, we believe that this paper will be of interest to the international readership because the developed maps and the approach will be used by relevant individuals or organizations involved in major construction projects in areas susceptible to geological disasters. One of our study's main limitations is the lack of historical geohazard information. The advanced machine-learning algorithms require a large number of training datasets. Therefore, we will consider advanced machine-learning and data-driven models for susceptibility mapping with more detailed datasets in a future study.

Author Contributions: Conceptualization, Hilal Ahmad, Chen Ningsheng, and Mahfuzur Rahman; methodology, Hilal Ahmad and Mahfuzur Rahman; software, Hilal Ahmad and Mahfuzur Rahman; validation, Hilal Ahmad, Mahfuzur Rahman, Md Monirul Islam, Hamid Reza Pourghasemi, and Ashraf Dewan; formal analysis, Hilal Ahmad and Mahfuzur Rahman; investigation, Hilal Ahmad, Chen Ningsheng, Mahfuzur Rahman, and Hamid Reza Pourghasemi; resources, Mahfuzur Rahman; writing—original draft preparation, Hilal Ahmad and Mahfuzur Rahman; writing—review and editing, Chen Ningsheng, Md Monirul Islam, Hamid Reza Pourghasemi, Syed Fahad Hussain, Jules Maurice Habumugisha, Enlong Liu, Han Zheng, Huayong Ni, and Ashraf Dewan; visualization, Chen Ningsheng, Mahfuzur Rahman, Md Monirul Islam, and Hamid Reza Pourghasemi; supervision, Chen Ningsheng, Mahfuzur Rahman, and Hamid Reza Pourghasemi; project administration, Chen Ningsheng; funding acquisition, Chen Ningsheng. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Grant number 41861134008); the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) of China (Grant number 2019QZKK0902); the National Key Research and Development Program of China (Project number 2018YFC1505202); and the Key R&D Projects of Sichuan Science and Technology (Grant number 18ZDYF0329).

Data Availability Statement: Data are available upon request on the corresponding author.

Acknowledgments: The authors would like to show gratitude to the University of Chinese Academy of Sciences and the Institute of Mountain Hazards and Environment, Chinese Academy of Sciences for their financial assistance. The authors would also like to acknowledge and appreciate the provision of rainfall data by the Pakistan Meteorological Department (PMD), without which this study would not have been possible.

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

References

1. Dowling, C.A.; Santi, P.M. Debris flows and their toll on human life: A global analysis of debris-flow fatalities from 1950 to 2011. *Nat. Hazards* **2014**, *71*, 203–227. [\[CrossRef\]](#)
2. Petley, D. Global patterns of loss of life from landslides. *Geology* **2012**, *40*, 927–930. [\[CrossRef\]](#)
3. Hewitt, K. Catastrophic landslides and their effects on the Upper Indus streams, Karakoram Himalaya, northern Pakistan. *Geomorphology* **1998**, *26*, 47–80. [\[CrossRef\]](#)
4. Khattak, G.A.; Owen, L.A.; Kamp, U.; Harp, E.L. Evolution of earthquake-triggered landslides in the Kashmir Himalaya, northern Pakistan. *Geomorphology* **2010**, *115*, 102–108. [\[CrossRef\]](#)
5. Clayton, A.; Stead, D.; Kinakin, D.; Wolter, A. Engineering geomorphological interpretation of the Mitchell Creek Landslide, British Columbia, Canada. *Landslides* **2017**, *14*, 1655–1675. [\[CrossRef\]](#)
6. Khanal, N.R.; Watanabe, T. Landslide and debris flow in the Himalayas: A case study of the Madi Watershed in Nepal. *Himal. J. Sci.* **2006**, *2*, 180–181. [\[CrossRef\]](#)
7. Ma, C.; Wu, X.; Li, B.; Hu, X. The susceptibility assessment of multi hazard in the Pearl River Delta Economic Zone, China. *Nat. Hazards Earth Sys. Sci. Discuss.* **2018**. [\[CrossRef\]](#)
8. Bazai, N.A.; Cui, P.; Zhou, K.J.; Abdul, S.; Cui, K.F.; Wang, H.; Zhang, G.T.; Liu, D.Z. Application of the soil conservation service model in small and medium basins of the mountainous region of Heilongjiang, China. *Int. J. Environ. Sci. Technol.* **2021**, 1–16.
9. Schuster, R.L.; Highland, L. *Socioeconomic and Environmental Impacts of Landslides in the Western Hemisphere*; US Department of the Interior, US Geological Survey: Denver, CO, USA, 2001.
10. Alimohammadlou, Y.; Najafi, A.; Yalcin, A. Landslide process and impacts: A proposed classification method. *Catena* **2013**, *104*, 219–232. [\[CrossRef\]](#)
11. Del Soldato, M.; Di Martire, D.; Bianchini, S.; Tomás, R.; De Vita, P.; Ramondini, M.; Casagli, N.; Calcaterra, D. Assessment of landslide-induced damage to structures: The Agnone landslide case study (southern Italy). *Bull. Int. Assoc. Eng. Geol.* **2019**, *78*, 2387–2408. [\[CrossRef\]](#)
12. Migoń, P.; Kasprzak, M.; Traczyk, A. How high-resolution DEM based on airborne LiDAR helped to reinterpret landforms—Examples from the Sudetes, SW Poland. *Landf. Anal.* **2013**, *22*, 89–101. [\[CrossRef\]](#)
13. Sassa, K. The 2017 Ljubljana Declaration on landslide risk reduction and the Kyoto 2020 Commitment for global promotion of understanding and reducing landslide disaster risk. *Landslides* **2017**, *14*, 1289–1296. [\[CrossRef\]](#)
14. Ayalew, L.; Yamagishi, H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* **2005**, *65*, 15–31. [\[CrossRef\]](#)
15. Akbar, T.A.; Ha, S.R. Landslide hazard zoning along Himalayan Kaghan Valley of Pakistan—by integration of GPS, GIS, and remote sensing technology. *Landslides* **2011**, *8*, 527–540. [\[CrossRef\]](#)
16. Reis, S.; Yalcin, A.; Atasoy, M.; Nisanci, R.; Bayrak, T.; Erduran, M.; Sancar, C.; Ekercin, S. Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio and analytical hierarchy methods in Rize province (NE Turkey). *Environ. Earth Sci.* **2011**, *66*, 2063–2073. [\[CrossRef\]](#)
17. Chalkias, C.; Ferentinou, M.; Polykretis, C. GIS-Based Landslide Susceptibility Mapping on the Peloponnese Peninsula, Greece. *Geosciences* **2014**, *4*, 176–190. [\[CrossRef\]](#)
18. Hervás, J.; Bobrowsky, P. Mapping: Inventories, Susceptibility, Hazard and Risk. In *Landslides—Disaster Risk Reduction*; Springer Science and Business Media LLC.: Berlin/Heidelberg, Germany, 2008; pp. 321–349.
19. Liucci, L.; Melelli, L.; Suteanu, C.; Ponziani, F. The role of topography in the scaling distribution of landslide areas: A cellular automata modeling approach. *Geomorphology* **2017**, *290*, 236–249. [\[CrossRef\]](#)
20. Baharvand, S.; Rahnamarad, J.; Soori, S.; Saadatkhah, N. Landslide susceptibility zoning in a catchment of Zagros Mountains using fuzzy logic and GIS. *Environ. Earth Sci.* **2020**, *79*, 1–10. [\[CrossRef\]](#)
21. Peng, D.; Xu, Q.; Liu, F.; He, Y.; Zhang, S.; Qi, X.; Zhao, K.; Zhang, X. Distribution and failure modes of the landslides in Heitai terrace, China. *Eng. Geol.* **2018**, *236*, 97–110. [\[CrossRef\]](#)
22. Mirdda, H.A.; Bera, S.; Siddiqui, M.A.; Singh, B. Analysis of bi-variate statistical and multi-criteria decision-making models in landslide susceptibility mapping in lower Mandakini Valley, India. *GeoJournal* **2019**, *85*, 681–701. [\[CrossRef\]](#)
23. Gariano, S.L.; Guzzetti, F. Landslides in a changing climate. *Earth Sci. Rev.* **2016**, *162*, 227–252. [\[CrossRef\]](#)
24. Saha, A.; Saha, S. Comparing the efficiency of weight of evidence, support vector machine and their ensemble approaches in landslide susceptibility modelling: A study on Kurseong region of Darjeeling Himalaya, India. *Remote. Sens. Appl. Soc. Environ.* **2020**, *19*, 100323. [\[CrossRef\]](#)
25. Nefeslioglu, H.A.; Gokceoglu, C.; Sonmez, H. An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Eng. Geol.* **2008**, *97*, 171–191. [\[CrossRef\]](#)

26. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **2013**, *11*, 425–439. [CrossRef]
27. Segoni, S.; Rosi, A.; Lagomarsino, D.; Fanti, R.; Casagli, N. Brief communication: Using averaged soil moisture estimates to improve the performances of a regional-scale landslide early warning system. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 807–812. [CrossRef]
28. Shahri, A.A.; Spross, J.; Johansson, F.; Larsson, S. Landslide susceptibility hazard map in southwest Sweden using artificial neural network. *Catena* **2019**, *183*, 104225. [CrossRef]
29. Acharya, T.D.; Lee, D.H. Landslide Susceptibility Mapping using Relative Frequency and Predictor Rate along Araniko Highway. *Ksce J. Civ. Eng.* **2018**, *23*, 763–776. [CrossRef]
30. Sharma, L.P.; Patel, N.; Ghose, M.K.; Debnath, P. Influence of Shannon’s entropy on landslide-causing parameters for vulnerability study and zonation—a case study in Sikkim, India. *Arab. J. Geosci.* **2010**, *5*, 421–431. [CrossRef]
31. Akgun, A. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: A case study at İzmir, Turkey. *Landslides* **2011**, *9*, 93–106. [CrossRef]
32. Vakhshoori, V.; Zare, M. Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods. *Geomat. Nat. Hazards Risk* **2015**, *7*, 1731–1752. [CrossRef]
33. Akgun, A.; Dag, S.; Bulut, F. Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood-frequency ratio and weighted linear combination models. *Environ. Earth Sci.* **2008**, *54*, 1127–1143. [CrossRef]
34. Ahmed, M.F.; Rogers, J.D.; Ismail, E.H. A regional level preliminary landslide susceptibility study of the upper Indus river basin. *Eur. J. Remote Sens.* **2017**, *47*, 343–373. [CrossRef]
35. Derbyshire, E.; Fort, M.; Owen, L.A. Geomorphological Hazards along the Karakoram Highway: Khunjerab Pass to the Gilgit River, Northernmost Pakistan (Geomorphologische Hazards entlang des Karakorum Highway: Khunjerab Paß bis zum Gilgit River, nördlichstes Pakistan). *Erdkunde* **2001**, 49–71. [CrossRef]
36. Kamp, U.; Growley, B.J.; Khattak, G.A.; Owen, L.A. GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. *Geomorphology* **2008**, *101*, 631–642. [CrossRef]
37. Bacha, A.S.; Shafique, M.; van der Werff, H. Landslide inventory and susceptibility modelling using geospatial tools, in Hunza-Nagar valley, northern Pakistan. *J. Mt. Sci.* **2018**, *15*, 1354–1370. [CrossRef]
38. Ali, S.; Biermanns, P.; Haider, R.; Reicherter, K. Landslide susceptibility mapping by using a geographic information system (GIS) along the China–Pakistan Economic Corridor (Karakoram Highway), Pakistan. *Nat. Hazards Earth Syst. Sci.* **2019**, *19*, 999–1022. [CrossRef]
39. Khan, H.; Shafique, M.; Khan, M.A.; Bacha, M.A.; Shah, S.U.; Calligaris, C. Landslide susceptibility assessment using Frequency Ratio, a case study of northern Pakistan. *Egypt. J. Remote Sens. Space Sci.* **2019**, *22*, 11–24. [CrossRef]
40. Zeitler, P.K. Cooling history of the NW Himalaya, Pakistan. *Tectonics* **1985**, *4*, 127–151. [CrossRef]
41. Jade, S.; Bhatt, B.; Yang, Z.; Bendick, R.; Gaur, V.; Molnar, P.; Anand, M.; Kumar, D. GPS measurements from the Ladakh Himalaya, India: Preliminary tests of plate-like or continuous deformation in Tibet. *Geol. Soc. Am. Bull.* **2004**, *116*, 1385–1391. [CrossRef]
42. Goudie, A.; Brunsden, D.; Whalley, W.; Collins, D.; Derbyshire, E. The geomorphology of the Hunza valley, Karakoram mountains, Pakistan. In *The International Karakoram Project. International Conference*; Cambridge University Press: Cambridge, UK, 1984; pp. 359–410.
43. DiPietro, J.A.; Pogue, K.R. Tectonostratigraphic subdivisions of the Himalaya: A view from the west. *Tectonics* **2004**, *23*. [CrossRef]
44. Yeats, R.S.; Lillie, R.J. Contemporary tectonics of the Himalayan frontal fault system: Folds, blind thrusts and the 1905 Kangra earthquake. *J. Struct. Geol.* **1991**, *13*, 215–225. [CrossRef]
45. Izaz, A.; Su, L.; Aamir, A.; Mehtab, A.; Liu, Z.; Dong, Z.; Hamid, F. Numerical analysis of rockfall and slope stability along the Karakorum Highway in Jijal-Pattan. *J. Civ. Environ. Eng.* **2021**, *43*, 36–47.
46. Ali, S.; Schneiderwind, S.; Reicherter, K. Structural and Climatic Control of Mass Movements Along the Karakoram Highway. In *Workshop on World Landslide Forum*; Springer: Cham, Switzerland, 2017. [CrossRef]
47. Bazai, N.A.; Cui, P.; Carling, P.A.; Wang, H.; Hassan, J.; Liu, D.; Zhang, G.; Wen, J. Increasing glacial lake outburst flood hazard in response to surge glaciers in the Karakoram. *Earth Sci. Rev.* **2020**, 103432. [CrossRef]
48. Shroder, J. Quaternary glaciation of the Karakoram and Nanga Parbat Himalaya. In *Himalaya Sea*; Routledge: London, UK, 1993.
49. Kazmi, A.H.; Jan, M.Q. *Geology and Tectonics of Pakistan*; Graphic Publishers, 1997. Available online: <https://books.google.com.sg/books?id=tImVAAAACAAJ> (accessed on 2 February 2021).
50. Hewitt, K.; Gosse, J.; Clague, J.J. Rock avalanches and the pace of late Quaternary development of river valleys in the Karakoram Himalaya. *Bulletin* **2011**, *123*, 1836–1850. [CrossRef]
51. Hepdeniz, K. Using the analytic hierarchy process and frequency ratio methods for landslide susceptibility mapping in Isparta-Antalya highway (D-685), Turkey. *Arab. J. Geosci.* **2020**, *13*. [CrossRef]
52. Kuradusenge, M.; Kumaran, S.; Zennaro, M. Rainfall-Induced Landslide Prediction Using Machine Learning Models: The Case of Ngororero District, Rwanda. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4147. [CrossRef]
53. Catani, F.; Lagomarsino, D.; Segoni, S.; Tofani, V. Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 2815–2831. [CrossRef]
54. Li, L.; Bakelants, L.; Solana, C.; Canters, F.; Kervyn, M. Dating lava flows of tropical volcanoes by means of spatial modeling of vegetation recovery. *Earth Surf. Process. Landf.* **2018**, *43*, 840–856. [CrossRef]

55. Intarawichian, N.; Dasananda, S. Analytical hierarchy process for landslide susceptibility mapping in lower mae chaem watershed, Northern Thailand. *Suranaree J. Sci. Technol.* **2010**, *17*, 1–16.
56. Hengl, T.; Reuter, H.I. *Geomorphometry: Concepts, Software, Applications*; Newnes: Boston, MA, USA, 2008.
57. Peduzzi, P. Landslides and vegetation cover in the 2005 North Pakistan earthquake: A GIS and statistical quantitative approach. *Nat. Hazards Earth Syst. Sci.* **2010**, *10*, 623–640. [[CrossRef](#)]
58. Hearn, G.J.; Hart, A.B. Geomorphological Contributions to Landslide Risk Assessment. *Develop. Earth Surf. Process.* **2011**, *15*, 107–148.
59. Kayastha, P.; Dhital, M.; De Smedt, F. Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: A case study from the Tinau watershed, west Nepal. *Comput. Geosci.* **2013**, *52*, 398–408. [[CrossRef](#)]
60. Gao, J.; Sang, Y. Identification and estimation of landslide-debris flow disaster risk in primary and middle school campuses in a mountainous area of Southwest China. *Int. J. Disaster Risk Reduct.* **2017**, *25*, 60–71. [[CrossRef](#)]
61. Myronidis, D.; Papageorgiou, C.; Theophanous, S. Landslide susceptibility mapping based on landslide history and analytic hierarchy process (AHP). *Nat. Hazards* **2016**, *81*, 245–263. [[CrossRef](#)]
62. Dolidon, N.; Hofer, T.; Jansky, L.; Sidle, R. Watershed and Forest Management for Landslide Risk Reduction. In *Landslides—Disaster Risk Reduction*; Metzler, J.B., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 633–649.
63. Forbes, K.; Broadhead, J.; Brardinoni, A.D.; Gray, D.; Stokes, B.V. Forests and landslides: The role of trees and forests in the prevention of landslides and rehabilitation of landslide-affected areas in Asia Second edition. *Rap Publ.* **2013**, *2*, 1–61.
64. Dai, F.; Lee, C. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology* **2002**, *42*, 213–228. [[CrossRef](#)]
65. Ortiz, J.A.V.; Martínez-Graña, A.M. A neural network model applied to landslide susceptibility analysis (Capitanajo, Colombia). *Geomat. Nat. Hazards Risk* **2018**, *9*, 1106–1128. [[CrossRef](#)]
66. Toebe, M.; Filho, A.C. Multicollinearity in path analysis of maize (*Zea mays* L.). *J. Cereal Sci.* **2013**, *57*, 453–462. [[CrossRef](#)]
67. O'Brien, R.M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Qual. Quant.* **2007**, *41*, 673–690. [[CrossRef](#)]
68. Polykretis, C.; Chalkias, C. Comparison and evaluation of landslide susceptibility maps obtained from weight of evidence, logistic regression, and artificial neural network models. *Nat. Hazards* **2018**, *93*, 249–274. [[CrossRef](#)]
69. Sahin, E.K.; Colkesen, I.; Kavzoglu, T. A comparative assessment of canonical correlation forest, random forest, rotation forest and logistic regression methods for landslide susceptibility mapping. *Geocarto Int.* **2018**, *35*, 341–363. [[CrossRef](#)]
70. Sun, X.; Chen, J.; Bao, Y.; Han, X.; Zhan, J.; Peng, W. Landslide susceptibility mapping using logistic regression analysis along the Jinsha river and its tributaries close to Derong and Deqin County, southwestern China. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 438. [[CrossRef](#)]
71. Rasyid, A.R.; Bhandary, N.P.; Yatabe, R. Performance of frequency ratio and logistic regression model in creating GIS based landslides susceptibility map at Lompobattang Mountain, Indonesia. *Geoenviron. Disasters* **2016**, *3*, 19. [[CrossRef](#)]
72. Denison, D.G.; Holmes, C.C.; Mallick, B.K.; Smith, A.F. *Bayesian Methods for Nonlinear Classification and Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2002; Volume 386, pp. 1–296.
73. Spiegelhalter, D.J.; Knill-Jones, R.P. Statistical and Knowledge-Based Approaches to Clinical Decision-Support Systems, with an Application in Gastroenterology. *J. R. Stat. Soc. Ser. A Gen.* **1984**, *147*, 35. [[CrossRef](#)]
74. Bonham-Carter, G.F.; Agterberg, F.P.; Wright, D.F. Integration of geological datasets for gold exploration in Nova Scotia. *Digit. Geol. Geogr. Inf. Syst.* **1989**, *54*, 15–23. [[CrossRef](#)]
75. Bonham-Carter, G.F. Geographic information systems for geoscientists-modeling with GIS. *Comput. Methods Geosci.* **1994**, *13*, 398.
76. Van Westen, C.; Rengers, N.; Soeters, R. Use of geomorphological information in indirect landslide susceptibility assessment. *Nat. Hazards* **2003**, *30*, 399–419. [[CrossRef](#)]
77. Sharma, S.; Mahajan, A. A comparative assessment of information value, frequency ratio and analytical hierarchy process models for landslide susceptibility mapping of a Himalayan watershed, India. *Bull. Eng. Geol. Environ.* **2019**, *78*, 2431–2448. [[CrossRef](#)]
78. Yufeng, S.; Fengxiang, J. Landslide Stability Analysis Based on Generalized Information Entropy. In Proceedings of the 2009 International Conference on Environmental Science and Information Application Technology, Wuhan, China, 4–5 July 2009; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2009; Volume 2, pp. 83–85.
79. Milaghardan, A.H.; Abbaspour, R.A.; Khalesian, M. Evaluation of the effects of uncertainty on the predictions of landslide occurrences using the Shannon entropy theory and Dempster–Shafer theory. *Nat. Hazards* **2020**, *100*, 49–67. [[CrossRef](#)]
80. Ozdemir, A.; Altural, T. A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *J. Asian Earth Sci.* **2013**, *64*, 180–197. [[CrossRef](#)]
81. Hassan, J.; Chen, X.; Muhammad, S.; Bazai, N.A. Rock glacier inventory, permafrost probability distribution modeling and associated hazards in the Hunza River Basin, Western Karakoram, Pakistan. *Sci. Total Environ.* **2021**, *782*, 146833. [[CrossRef](#)]
82. Eeckhaut, M.V.D.; Vanwalleghem, T.; Poesen, J.; Govers, G.; Verstraeten, G.; Vandekerckhove, L. Prediction of landslide susceptibility using rare events logistic regression: A case-study in the Flemish Ardennes (Belgium). *Geomorphology* **2006**, *76*, 392–410. [[CrossRef](#)]
83. Wubalem, A. Landslide susceptibility mapping using statistical methods in Uatzau catchment area, northwestern Ethiopia. *Geoenviron. Disasters* **2021**, *8*, 1. [[CrossRef](#)]
84. Collins, T.K. Debris flows caused by failure of fill slopes: Early detection, warning, and loss prevention. *Landslides* **2008**, *5*, 107–120. [[CrossRef](#)]

85. Yalcin, A. A geotechnical study on the landslides in the Trabzon Province, NE, Turkey. *Appl. Clay Sci.* **2011**, *52*, 11–19. [[CrossRef](#)]
86. Hong, Y.; Adler, R.; Huffman, G. Use of satellite remote sensing data in the mapping of global landslide susceptibility. *Nat. Hazards* **2007**, *43*, 245–256. [[CrossRef](#)]
87. Egholm, D.L.; Knudsen, M.F.; Sandiford, M. Lifespan of mountain ranges scaled by feedbacks between landsliding and erosion by rivers. *Nature* **2013**, *498*, 475–478. [[CrossRef](#)]
88. Constantin, M.; Bednarik, M.; Jurchescu, M.C.; Vlaicu, M. Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). *Environ. Earth Sci.* **2011**, *63*, 397–406. [[CrossRef](#)]
89. Tahirkheli, R.K. The India-Eurasia suture zone in northern Pakistan: Synthesis and interpretation of recent data at plate scale. *Geodyn. Pak.* **1979**, 125–130.
90. DiPietro, J.A.; Hussain, A.; Ahmad, I.; Khan, M.A. The Main Mantle Thrust in Pakistan: Its character and extent. *Geol. Soc. Lond. Spec. Publ.* **2000**, *170*, 375–393. [[CrossRef](#)]
91. Benchelha, S.; Chennaoui Aoudjehane, H.; Hakdaoui, M.; El Hamdouni, R.; Mansouri, H.; Benchelha, T.; Layelmam, M.; Alaoui, M. Landslide Susceptibility Mapping in the Municipality of Oudka, Northern Morocco: A Comparison between Logistic Regression and Artificial Neural Networks Models. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-4/W12*, 41–49. [[CrossRef](#)]
92. Nijmeijer, R.; de Haas, A.; Dost, R.; Budde, P. ILWIS 3.0 Academic: User's Guide. 2001. Available online: <https://www.itc.nl/ilwis/users-guide/> (accessed on 2 February 2021).
93. Islam, M.; Sado, K. Development of flood hazard maps of Bangladesh using NOAA-AVHRR images with GIS. *Hydrol. Sci. J.* **2000**, *45*, 337–355. [[CrossRef](#)]
94. Ochi, S.; Rahman, N.; Kakiuchi, H. A Study on Flood Risk Evaluation in Bangladesh using Remote Sensing and GIS. *J. Jpn. Soc. Photogramm. Remote. Sens.* **1991**, *30*, 34–38. [[CrossRef](#)]
95. Park, S.; Choi, C.; Kim, B.; Kim, J. Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea. *Environ. Earth Sci.* **2012**, *68*, 1443–1464. [[CrossRef](#)]
96. Rahim, I.; Ali, S.M.; Aslam, M. GIS Based Landslide Susceptibility Mapping with Application of Analytical Hierarchy Process in District Ghizer, Gilgit Baltistan Pakistan. *J. Geosci. Environ. Prot.* **2018**, *6*, 34–49. [[CrossRef](#)]
97. Kanwal, S.; Atif, S.; Shafiq, M. GIS based landslide susceptibility mapping of northern areas of Pakistan, a case study of Shigar and Shyok Basins. *Geomat. Nat. Hazards Risk* **2016**, *8*, 348–366. [[CrossRef](#)]