

## Article

# Risk Assessment of Resources Exposed to Rainfall Induced Landslide with the Development of GIS and RS Based Ensemble Metaheuristic Machine Learning Algorithms

Javed Mallick <sup>1,\*</sup>, Saeed Alqadhi <sup>1</sup>, Swapan Talukdar <sup>2</sup>, Majed AlSubih <sup>1</sup>, Mohd. Ahmed <sup>1</sup>,  
Roohul Abad Khan <sup>1</sup>, Nabil Ben Kahla <sup>1</sup> and Saud M. Abutayeh <sup>3</sup>

<sup>1</sup> Department of Civil Engineering, College of Engineering, King Khalid University, Abha 61411, Saudi Arabia; sdaalqadi@kku.edu.sa (S.A.); malsubih@kku.edu.sa (M.A.); mall@kku.edu.sa (M.A.); rakhan@kku.edu.sa (R.A.K.); nbohlah@kku.edu.sa (N.B.K.)

<sup>2</sup> Department of Geography, University of Gour Banga, Malda 732101, India; swapantalukdar65@gmail.com

<sup>3</sup> Geotechnical Engineer, Ministry of Transportation, Riyadh 11178, Saudi Arabia; saud.tayeh@idsge.com

\* Correspondence: jmallick@kku.edu.sa; Tel.: +966-172-428-439; Fax: +966-172-418-152

**Abstract:** Disastrous natural hazards, such as landslides, floods, and forest fires cause a serious threat to natural resources, assets and human lives. Consequently, landslide risk assessment has become requisite for managing the resources in future. This study was designed to develop four ensemble metaheuristic machine learning algorithms, such as grey wolf optimized based artificial neural network (GW-ANN), grey wolf optimized based random forest (GW-RF), particle swarm optimization optimized based ANN (PSO-ANN), and PSO optimized based RF for modeling rainfall-induced landslide susceptibility (LS) in Aqabat Al-Sulbat, Asir region, Saudi Arabia, which observes landslide frequently. To obtain very high precision and robust prediction from machine learning algorithms, the grey wolf and PSO optimization algorithms were integrated to develop new ensemble machine learning techniques. Subsequently, LS maps produced by training dataset were validated using the receiver operating characteristics (ROC) curve based on the testing dataset. Based on the area under curve (AUC) value of ROC curve, the best method for LS modeling was selected. We developed ROC curve-based sensitivity analysis to investigate the influence of the parameters for LS modeling. The Gumble extreme value distribution was employed to estimate the rainfall at 2, 5, 10, 20, 50, and 100 year return periods. Then, the landslide hazard maps were prepared at different return periods by integrating the best LS model and estimated rainfall at different return periods. The theory of danger pixels was employed to prepare a final risk assessment of the resources, which have been exposed to the landslide. The results showed that 27–42 and 6–15 km<sup>2</sup> were predicted as the very high and high LS zones using four ensemble metaheuristic machine learning algorithms. Based on the area under curve (AUC) of ROC, GR-ANN (AUC-0.905) appeared as the best model for LS modeling. The areas under high and very high landslide hazard were gradually increased over the progression of time (26 km<sup>2</sup> at the 2 year return period and 40 km<sup>2</sup> at the 100 year return period for the high landslide hazard zone, and 6 km<sup>2</sup> at the 2 year return period and 20 km<sup>2</sup> at the 100 year return period for the very high landslide hazard zone). Similarly, the areas of danger pixel also increased gradually from the 2 to 100 year return periods (37 km<sup>2</sup> to 62 km<sup>2</sup>). Various natural resources, such as scrubland, built up, and sparse vegetation, were identified under risk zone due to landslide hazards. In addition, these resources would be exposed extensively to landslides over the advancement of return periods. Therefore, the outcome of the present study will help planners and scientists to propose high precision management plans for protecting natural resources, which have been exposed to landslides.

**Keywords:** rainfall-induced landslide susceptibility; sensitivity analysis; metaheuristic optimization; risk assessment; return periods; Gumble extreme value distribution



**Citation:** Mallick, J.; Alqadhi, S.; Talukdar, S.; AlSubih, M.; Ahmed, M.; Khan, R.A.; Kahla, N.B.; Abutayeh, S.M. Risk Assessment of Resources Exposed to Rainfall Induced Landslide with the Development of GIS and RS Based Ensemble Metaheuristic Machine Learning Algorithms. *Sustainability* **2021**, *13*, 457. <https://doi.org/10.3390/su13020457>

Received: 12 November 2020

Accepted: 30 December 2020

Published: 6 January 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/>).

## 1. Introduction

Landslide has been considered as the most frequent geo-hazard in the mountainous regions across the world, which causes a serious threat to the natural resources, assets and lives [1]. It frequently occurs when the mountainous regions observe an intense rainfall [2]. Since the landslide occurrences comprise 9% of all-natural hazards in the last few decades, it is expected that the trend of landslide occurrences will be increased in future. The main reasons are the acceleration of developmental processes in mountainous regions, such as the urbanization and expansion of agricultural land, along with intense rainfall for a very shorter period of time due to climate change, which can increase the frequency of landslides [2–5]. Generally, landslides are a natural phenomenon and fairly uncertain in nature. Consequently, its occurrences cannot be controlled [2]. The damages can be minimized by implementing proper management plans, therefore, to minimize the damages caused by landslides, risk assessment and susceptibility mapping of landslides need to be investigated. Researchers have attempted to construct the landslide susceptibility map, which can identify the future probable landslide regions in spatial scale [3]. As a consequence, the construction of landslide susceptibility modeling has become a vital task to geological engineers and earth scientists to propose mitigation plans. Moreover, it is considered as the most basic and valuable step to prepare a comprehensive landslide hazard modeling and risk assessment, which can increase the performance of the disaster management plans [5].

As landslide susceptibility mapping is essential for preparing the risk assessment, it is required to have very high precision and accurate landslide maps. Very high-quality landslide susceptible models generally depend on a combination of several geospatial landslides triggering factors, such as elevation, slope, geology, land use land cover, aspect and others. Since the last decades, the utilization of advanced remote sensing technologies has allowed multi-resolution high-dimensional datasets to be conveniently accessible to researchers. The features of the earth surface have been extracted through these advanced remote sensing datasets. High-resolution satellite images provide highly accurate features of the earth surface, such as sentinel-2, advanced land observing satellite-phased array type L-band synthetic aperture radar (ALOS PALSAR) digital elevation model (DEM), QuickBird etc. [5]. The landslide susceptible models have been constructed by the integration of landslide triggering factors [6]. Researchers have developed and adopted a variety of approaches for integrating the triggering factors to generate the final landslide susceptibility models [7]. They are physical-based, heuristic, and statistical methods [7–9]. The physical process-based methods utilize the law of mechanics to model the slope instability to evaluate the landslide susceptibility. These are mainly applicable for a small area as they are required to have very detailed topography data [7]. Heuristic-based models predict the probability of the landslide occurrences based on expert views. These methods are strongly dependent on the expert's view; therefore, they are sometimes biased and produced moderately accurate output [8].

In comparison with the mentioned methods, statistical techniques have been extensively employed for landslide susceptibility modeling. However, the statistical models only require the normal distribution of the landslide triggering factors, and then it can generate good quality landslide susceptible models. However, it is very difficult to obtain normally distributed triggering factors [8]. With the advancement of remote sensing databases and machine learning algorithms, researchers have started to utilize the machine learning algorithms to overcome mentioned drawbacks and generate good quality landslide susceptibility models. Widely popular machine learning algorithms which have been extensively used for landslide susceptibility modeling are ANN [9,10], support vector machine (SVM) [11–13], RF [14–16], and decision tree [17,18]. However, these conventional machine learning algorithms also have several drawbacks, such as local optimum, over-fitting, low training speed etc. [7].

Researchers have developed and utilized the ensemble machine learning algorithms to overcome the mentioned drawbacks and generate very high-quality landslide susceptibility

maps [7]. Many ensemble machine learning algorithms have reported high performances for generating landslide susceptibility maps, such as random subspace [19,20], Reptree [21], bagging [22–24], naive Bayes [9,25], ANFIS [26–29], AdaBoost [22], and Rotation forest [8,30]. The outstanding results produced by the ensemble machine learning algorithms inspire researchers to develop, utilize, and test new ensemble machine learning algorithms. Although no general agreement has been found for the selection of the best method for landslide susceptibility modeling [31]. Researchers recommend the development and testing of new models for obtaining very high precision landslide susceptibility maps [32].

Many ensemble machine learning algorithms have the drawbacks of over-fitting [31]. Eshtay et al. [33] reported that the weights of the input layers are randomly produced, and the models do not update or optimize the weights during the training phase, which causes an unstable performance. To solve these drawbacks, metaheuristic optimized algorithms have been proposed, which searches for the best parameters and optimizes the weights of the input layer. Taormina and Chau [34] used the extreme learning machine (ELM) and its ensemble with particle swarm optimization (PSO) for streamflow prediction and reported that the ensemble model outperforms the standalone ELM model. Niu et al. [35] revealed that PSO could substantially improve the prediction reliability of the base model. Recently, some hybrid models have been developed in landslide susceptibility by combining machine learning algorithms and optimization algorithms, such as PSO and grey wolf (GWO) [35]. Very rare studies on landslide susceptibility modeling using the optimized ensemble machine learning techniques in Saudi Arabia have yet been documented. To fill the research gaps, the present study designed to develop four optimized machine learning algorithms by combining PSO and GWO with ANN and RF. They are GW-ANN, GW-RF, PSO-ANN, and PSO-RF. In addition, the area under the receiver operating characteristics curve (AUC-ROC) based sensitivity analysis was proposed for investigating the performance of landslide triggering factors for LS modeling, which has not yet been explored in the existing literature. The map removal technique and ROC curve were integrated to develop this technique. The ROC based sensitivity analysis is actually the modified version of map removal technique. First, LS models were constructed based on map removal technique. The map removal technique is the removal of individual parameters at one time during the modeling of landslide susceptibility maps and then performed same process repeatedly for all parameters [36]. In this way, the effect of removing any parameters among all selected parameters for modeling the output have been measured [36]. In addition, ROC curve was employed to validate the map removal based LS models. Consequently, the quantification of the variation of the LS models produced by map removal techniques can be possible, which were the major drawbacks of the map removal technique. Generally, several statistical methods, such as sensitivity, specificity, true predicted positive and negative, and Youden index J can be used for validation. However, numerous researchers have preferred to use the ROC curve for validation as it is considered as unbiased and statistically robust than the mentioned method [35]. For this reason, the ROC curve was employed to validate the map removal based LS models.

When the landslide susceptibility has been integrated with the temporal and magnitude information, the landslide hazard model has been generated [2]. Coro minas et al. [37] suggested several methods to construct the landslide hazard map, such as judgmental method, empirical method, rational method, magnitude-frequency relations, and indirect approach. However, the most common approach is the magnitude-frequency approach, which combines the landslide susceptibility with the temporal landslide triggering factors, such as rainfall or earthquake [38]. The generated landslide hazard model has been further combined with the socio-economic framework to produce the potential risk of different resources, such as forest, buildings, structures etc. [2]. The risk assessment has been considered as the ultimate scope for landslide research, which facilitates to propose disaster and land management plans [2,39].

Many studies on the landslide risk assessment have already been carried out using a variety of methods [40–44]. They are qualitative, semi-quantitative, and quantitative meth-

ods [2]. Qualitative methods evaluate the landslide risk based on expert knowledge [45]. Semi-quantitative methods assess the landslide risk by assigning the weights to the slope instability triggering factors [46]. Quantitative methods model the landslide risk by combining landslide hazard, vulnerability, and elements at risk, such as infrastructures (dam, barrage, buildings, road etc.), structure, human lives, transportation, public services, and mining [47]. However, the data scarcity is the main hindrance to construct an informative landslide risk model; therefore, semi-quantitative methods are suitable. In the present study, the semi-quantitative approach was adopted because of the data scarcity in terms of the historical landslide, consequences, details of resources at risk and their conditions, and values of damages in price. Many researchers successfully utilized the theory of danger pixel as the semi-quantitative approach for assessing the landslide risk in data-scarce regions [48].

Previous literature reported that numerous ensemble machine learning algorithms have already been developed for predicting rainfall induced landslide susceptibility. However, it is very rare that studies on the development of hybrid ensemble machine learning algorithms by integration of meta-heuristic algorithms and machine learning algorithms for predicting landslide susceptibility have been carried out. In addition, no studies have been carried out to construct spatial sensitivity analysis to investigate the spatial differences, which have been observed in the landslide susceptibility maps if any landslide conditioning parameters are removed from the modeling process. For spatial sensitivity analysis, landslide susceptibility models, similar to landslide conditioning factors, would be constructed. On the other hand, very rare studies on the spatiotemporal landslide susceptibility modeling have been carried out. Additionally, no studies have been carried out on the spatiotemporal risk assessment due to rainfall induced landslide. Therefore, these research gaps should be addressed to propose comprehensive and highly accurate landslide management plans.

The present study area, Ababa Al-Sulbat in Asir region, has observed frequent landslide events. Consequently, the risk assessment has become essential to minimize the damages of different resources. The main scope of the work is to develop four ensemble metaheuristic machine learning algorithms and ROC curve-based sensitivity analysis at a spatial scale. The ROC based sensitivity analysis was employed to investigate the influence of the conditioning parameters for LS modeling. The spatiotemporal landslide risk assessment using the theory of danger pixel is another scope of the study. This work is the first comprehensive attempt to landslide risk assessment in Saudi Arabia. The main novelties can be summarized as follows:

- General: The work contributes to the robustness of knowledge by developing and utilizing methods to an unstudied area on the landslide susceptibility and spatiotemporal risk assessment.
- Regional: Increased knowledge of landslide susceptibility and spatiotemporal risk assessment in Aqabat Al-Sulbat, Asir region, Saudi Arabia. The outcome of this work would be a valuable basis for the earth scientists, government authorities, and stakeholders to improve land management and disaster management.
- Methodical: Proposed ensemble metaheuristic machine learning algorithms, such as PSO-ANN, PSO-RF, GW-ANN, and GW-RF for LS modeling. Developed ROC based sensitivity model at spatial scale first time. Constructed long-term landslide risk assessment using danger pixel.

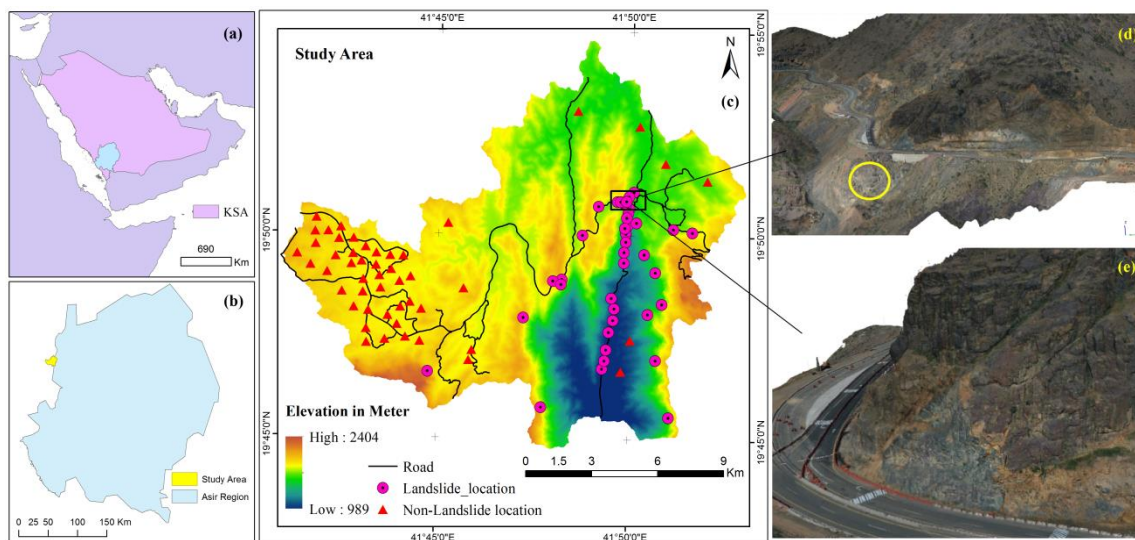
## 2. Materials and Methodology

In the present study, a variety of materials for preparing the landslide conditioning parameters and landslide hazard maps were collected from different sources. The ALOS PALSAR DEM was downloaded from Earth Science Data Systems of national aeronautics and space administration (NASA). The sentinel-2 multispectral instrument (MSI) was obtained from the earth explorer of the United States Geological Survey (<https://earthexplorer.usgs.gov/>). The geological map was extracted from the map of the Saudi Geological Survey

at the scale of 1:100,000. The soil texture data was collected from the field survey. The drainage map and road were digitized and prepared from Google earth imagery and ArcGIS 10.5 software.

### 2.1. Study Area

Aqabat Al-Sulbat is situated along Abha-Bahah Road at the northern part of Balqarn Area, in the Asir Region of Saudi Arabia, covering 199 km<sup>2</sup> area (Figure 1). The geographical location of study area extends between 19°45′4.407″ N and 19°54′42.055″ N and 41°40′52.31″ E to 41°53′6.169″ E. The study area characterizes by 989–2404 m. elevation with 1768 m. of average elevation from mean sea level. According to the geological map, the site belongs to Ablah Group, which is structurally-controlled by the Farwah Shear Zone and located between the Al Lith-Bidah and Shwas-Tayyah structural belts. Three formations occur in this group, such as Jerub Formation, Rafa Formation, and Thurat Formation, together composing a succession of epiclastic and volcanic rocks. A broad range of fractures exist, many of which split the rock into cubic or quadrangular blocks. All joints are exposed to weathering and potential soil erosion. Slope failure occurred due to the progression of a mass of rocks along a single discontinuity has been depicted as plane failure as such a mass on such a plane translates from its origin to its resting place. This is the case of the totally isolated rock block sitting on a discontinuity, and a situation seldom met in reality Sliding failure from a long continuous rock slope is often contained by release surfaces which may allow the mass to move down-dip along the basal plane. The wedge failure of rock slope may be the most common type of failure in the Aqabat Al-Sulbat area. The climatic condition of the study area is characterized by cold and semi-arid weather. The long-term average rainfall is 218 mm, out of which 75% of rainfall happens between February and June. The average maximum and minimum temperature of the study area are 29.5 and 16.8 °C.

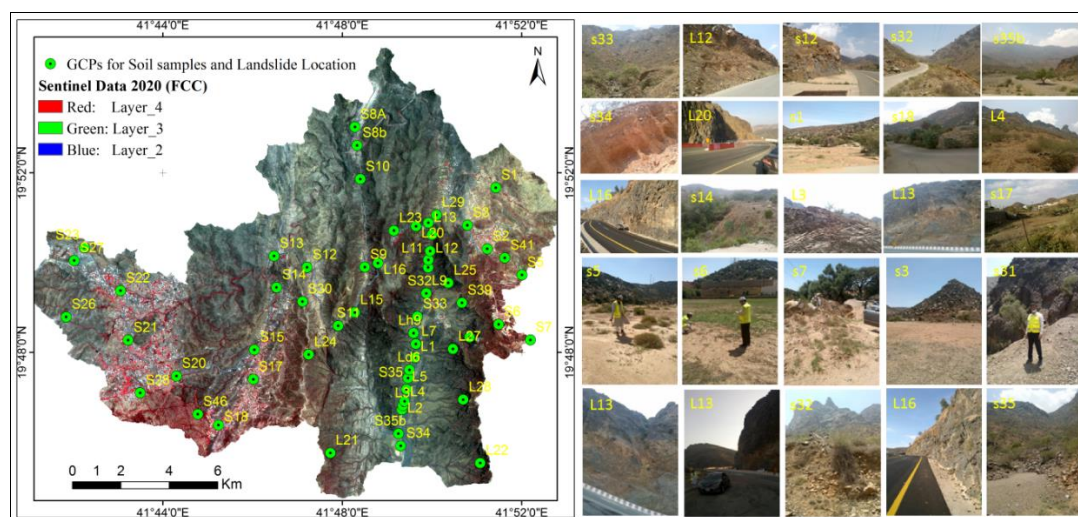


**Figure 1.** Showing the study area and landslide locations: (a) country boundary of Saudi Arabia with Asir region; (b) Asir region boundary with the study area location; (c) the study area, landslide locations and non-landslide location and (d,e) 3D model of landslide site (yellow circle showing landslide-prone area).

### 2.2. Landslide Inventories

The creation of landslide inventories is a vital task to train and test the LS model. In the present study, the landslide occurrences locations were examined and identified from the field survey and government report (Figure 2). Fifty landslide locations were collected using the global positioning system (GPS) and Google Earth. Out of which, 80% (40 points) of the total points were randomly selected for constructing the training datasets, while the rest of the points (10 points) were used for generating the validation datasets. The

classification-based LS modeling was adopted for the present study; therefore, it needs to incorporate binary data for the modeling. As a consequence, we were required to have non-landslide locations. Tang et al. (2020) [49] suggested using similar numbers of negative points for obtaining better results. Although no concrete literature has been found about the selection of the numbers of negative points. Therefore, 50 non-landslide locations were randomly selected based on the past or historical landslide records, local people's perception, and Google Earth image. The locations had not observed landslide yet, and were considered as non-landslide points or locations. The selected non-landslide locations were cross-checked with the expert geotechnical engineers and government reports, that showed that the non-landslide locations are authentic and can be used for preparing training and testing datasets. However, many researchers have followed the identical way to collect negative samples for landslide, flood and other predictive modeling [50]. Some researchers have used different instruments to test the sites in terms of soil texture, rock hardness, permeability, and other aspects and based on the testing results, they selected the sites as negative or positive samples [50]. On the other hand, in the data scarce and technologically developing regions, researchers use historical remote sensing data archive (Landsat images), government report, and local people perception to analyze the study area to pick the negative samples. Then, selected samples have been judged in respect to the characteristics of the landslide conditioning parameters and expert opinions. Therefore, the selection of negative or non-landslide location has been done very scientifically. However, similar to landslide data, the non-landslide data was partitioned by the 80–20 ratio for constructing training and testing datasets. The training dataset was created by combining 80% of both landslide and non-landslide locations. We generated training datasets in binary form by assigning 0 and 1 value for landslide and non-landslide locations. Similarly, the validation dataset was generated in binary form. The data was extracted from the landslide triggering factors based on the training dataset in 'spatial analyst' toolbox of ArcGIS 10.5 software.



**Figure 2.** Showing some field photographs with corresponding point locations for soil sample collection sites and landslide locations in the study area.

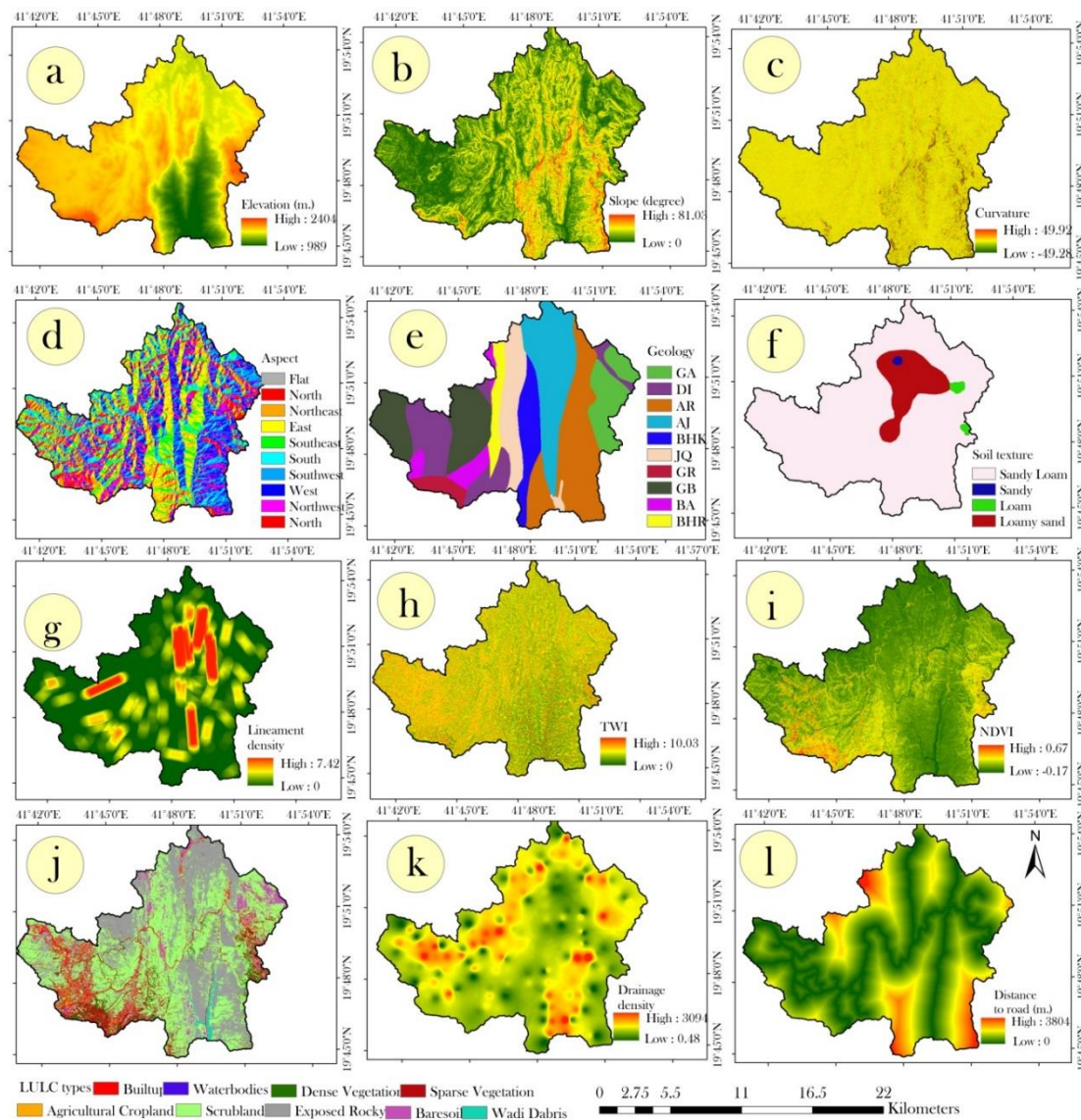
### 2.3. Landslide Conditioning Parameters

The landslide susceptibility modeling has been constructed based on the complex relationship between landslide triggering factors and past landslide [51]. The selection of suitable parameters is a challenging work. Expert knowledge on the study area and previous literature can provide the way for selecting the landslide triggering factors [51]. Twelve landslide triggering factors were selected for modeling landslide susceptibility, such

as elevation, slope, aspect, curvature, geology, soil texture, land use/land cover, drainage density, lineament, distance to road, NDVI, and topographic wetness index.

### 2.3.1. Elevation

One of the significant factors of landslide events is elevation, which controls vegetation, the direction of runoff, rate of drainage density, gravitational energy of landslides and human activities [9]. The slope failure and durability have been caused by elevation [20]. The elevation map was generated from a DEM. The elevation in the study area ranges from 2404 to 989 m (Figure 3a).



**Figure 3.** Showing the landslide triggering factors, such as (a) elevation, (b) slope, (c) curvature, (d) aspect, (e) geology, (f) soil texture, (g) lineament density, (h) topographic wetness index (TWI), (i) normalized differentiation vegetation index (NDVI), (j) land use land cover (LULC) types, (k) drainage density, (l) distance to road.

### 2.3.2. Slope

The most prevalent factor of landslide occurrences is slope [24,52,53]. The occurrence of landslides is directly controlled by the slope angle of an area [54]. The slope map was extracted from ALOS PALSAR DEM. The slope of the study area ranges from  $81.03^\circ$  to  $0^\circ$  (Figure 3b).

### 2.3.3. Curvature

Curvature has been considered as one of the influential landslide conditioning topographical parameters. The main function of curvature is the generation of surface runoff. It also influences the infiltration. Generally, the curvature is of three categories, such as the concave (negative curvature), flat (zero curvature) and convex (positive curvature). The runoff generation capability is much more for convex slope than concave slope [54]. Therefore, the runoff in mountainous regions depends upon the types of curvature (Figure 3c).

### 2.3.4. Aspect

The aspect has been described as the direction of the slope of an area. The aspect map has been extracted from the DEM of the study area. The generated aspect map was, then, reclassified into eight classes, such as flat, north, northeast, east, southeast, south, southwest, west, and northwest (Figure 3d).

### 2.3.5. Geology

Rock characteristics and its types regulate the physical features of rocks, which directly regulate the slope instability or failure. The geology map of the study area was digitized from the Saudi Geological survey at the scale of 1:100000. It was converted into raster format in 'spatial analyst' toolbox of ArcGIS. However, the geology map of the study area has ten geological classes, such as (1) quartz syenite-stock, sheets, and ring dikes (GA), (2) hornblende diorite, mafic diorite and mafictonalite (DI), (3) quartzite, metasilstone, conglomerate, feldspathic greywacke, black marble and white marble (AR), (4) andesitic and dacitic pyroclastic rocks with minor flows (AJ), (5) feldspathic and lithic greywacke (BHK), (6) andesitic, dacitic, and basaltic flow rock, lithic and crystal tuff and volcanic breccia (JQ), (7) monzogranite (GR), (8) Olivine gabbro, metagabbro, metadiorite, and anorthosite (GB), (9) amphibole and biotite schist (BA), and (10) feldspathic greywacke, carbonaceous chert, argillite, and slate (BHR) (Figure 3e).

### 2.3.6. Soil Texture

Soil samples were collected from various in-situ locations in the study area. A total of 32 soil samples of approximately 1 kg of aggregate stability (0–30 cm depth) was collected from the study area. The locations of the soil sampling sites were recorded using GPS. The stratified composite sampling technique was adopted. The study area was sub-grouped into various elevation zones, LULC and soil moisture. The site was then considered separately and two replicates, two to three meters apart, are collected at each survey site. Each sample was carefully weighed and sieved by 2 mm, and then, the soil texture and organic matter were extracted by following the standard protocol. Using the hydrometer method (Stokes law), the size of soil grains (texture analysis) were determined. Figure 3f illustrated the distribution of soil types in the study area. Four soil texture categories were recognized, such as sandy loam, loam, loamy sand, and sandy loam.

### 2.3.7. Lineament Density

The lineament refers to the linear tectonic breaks in any area, which generally declines rock potency. Generally, the lineament can be different types, such as fault, fracture, and shearing. The lineament is the type of discontinuity and weaker section in a geological formation. The weaker part of geology or lineament has been considered as the most essential factors for landslides [9,20]. In the present study area, we used sentinel-2 satellite image for extracting the lineament in ENVI software. The line density tool was employed for constructing the lineament density map (Figure 3g). Higher the lineament density, higher the chances of landslide occurrences

### 2.3.8. Topographic Wetness Index (TWI)

The TWI is an influential hydrologic parameter that regulates landslides. It measures the magnitude of water accumulation in the basin or site. The impact on soil moisture



on soil materials causes pore water pressure and reduces the soil strength, which directly controls the slope failure, especially, landslide. Higher the TWI values, higher the chances of observing landslides. In the present study, TWI value ranges from 0–10.03 (Figure 3h).

#### 2.3.9. Normalized Differentiation Vegetation Index (NDVI)

The NDVI or vegetation cover plays a significant role for landslide occurrences. The vegetation covers extensively modify the soil hydrology by amplifying the precipitation interception, evapotranspiration and infiltration, which decrease the quantity of water that reaches the soil. This can be important for long-term rainfall. In addition, it increases the soil strength by root reinforcement. Higher the presence of vegetation, lower the chances of landslide and vice versa. In the present study, to extract vegetation, we used band 8 and band 4 of sentinel-2 data. The NDVI value ranges from 0.67 to −0.17. Higher vegetation concentration in the study area can be found in the western part of the study area (Figure 3i).

#### 2.3.10. Land Use Land Cover (LULC) Mapping

The LULC affects the slope durability by changing land use/land cover and irritating the slope durability situation [55]. In the present study, we prepared a LULC map from sentinel-2 data. For LULC mapping, we employed maximum likelihood classification classifiers. We classified LULC types in the present study area into nine types, such as built up, waterbodies, dense vegetation, sparse vegetation, agricultural cropland, scrubland, exposed rocky, bare soil and wadi debris (Figure 3j).

#### 2.3.11. Drainage Density

The drainage density is the total length of all streams and river in a particular river basin divided by the total area of that basin. The drainage density is adversely correlated with the landslide. If the drainage density increases, the chances of landslide susceptibility increases and vice versa. The drainage map was prepared from Google Earth. Then the digitized drainage map was imported to ArcGIS software for creating the drainage density. The line density tool of ArcGIS software was used for mapping drainage density. In the present study, higher drainage density has been found in the eastern part of the study area (Figure 3k).

#### 2.3.12. Distance to Road

The construction of roads in the mountainous area by cutting the slope, which disturbs natural typology and affects slope stability. Therefore, road construction causes a slope failure. The stability of slope changes from stable to unstable during the construction of road and movement of vehicles, which may cause cracks. These cracks absorb huge water, which loosens the slope materials. Thus, it triggers landslides. The intense rainfall can accelerate the process of slope failure, which leads to landslide occurrences. Lower the distance or near the road in the mountain area, the probability of landslide occurrences is very high. The road map was digitized from Google earth, and subsequently, we used the Euclidean distance tool in ArcGIS software for preparing distance to the road map. The southeastern part of the study area covers a higher concentration of road (Figure 3l).

### 2.4. Methods for Multicollinearity Analysis

Twelve landslide triggering factors were selected in this study for landslide susceptibility and risk assessment. Therefore, to model landslide susceptibility, it is necessary to carefully optimize and check the collinearity of the selected parameters because the statistical and machine learning algorithms may be sensitive to collinearities. In addition, the collinearity can cause disturbance during the modeling process and, as a consequence, it reduces accuracy to predict the landslide susceptibility [56]. This analysis helps in choosing the appropriate parameters by excluding the redundancy parameters to obtain good results [57]. A variety of methods are available for analyzing the multicollinearity, such as information gain ratio, chi-square, relief-f, linear support vector machine, tolerance (TOL)

and variance inflation factors (VIF) [57]. In the present study, tolerance and VIF were used to test the multicollinearity among the selected factors. Many previous studies obtained very good findings by applying TOL and VIF. Note that, at the last stage of analysis, the factors having the collinearity problems should be excluded to achieve high precision results. Higher the VIF indicates higher the collinearity. To do the multicollinearity validation, a linear regression analysis was performed in which the landslide location was considered as a dependent variable, and the other triggering parameters are considered independent variables and the  $R^2$  value was calculated. This value was used to calculate the tolerance and variance inflation factor (VIF) of that input variable using Equations (1) and (2), respectively [57].

$$T = 1 - R^2 \quad (1)$$

$$\text{VIF} = \frac{1}{T} \quad (2)$$

## 2.5. Methods for Landslide Susceptibility Model

### 2.5.1. Optimization Algorithms

To obtain high precision and accurate prediction, the machine learning algorithms, such as ANN and RF were integrated with GWO and PSO as an ‘add-in’ optimizer algorithms. Thus, four ensemble optimization machine learning algorithms were generated, such as GW-ANN, GW-RF, PSO-ANN, and PSO-RF.

#### Grey Wolf Optimization (GWO)

At present, GWO [58] is an advanced optimization algorithm, which can replicate the hunting behaviors and leadership in the group of grey wolves [59]. Generally, a pack of grey wolves comprises four groups of leadership, such as alpha, beta, delta, and omega. The alpha wolves are considered as the dominant and strongest wolves, who guide the pack and take a decision for hunting, migration, nesting and so on [60]. The beta wolves are regarded as secondary wolves, which help the alpha for taking a decision. Furthermore, the delta wolves consist of elders, sentinels, hunters, scouts, and caretakers of the pack. The omega wolves are considered as the lowest ranking members in the pack, who act as the babysitter/scapegoat [61]. Albeit delta wolves dominate the omega wolves, but they obey the alphas and betas. To address optimization problems, alphas are considered as the fittest solution, followed by beta, gamma, and lastly omega. The finding procedure (i.e., finding the location of a prey), which starts with creating an indiscriminate population of postulant accomplishment (i.e., grey wolves) is statistically designed with the object of reproducing the hunting attitude of a grey wolves pack. However, the hunting behavior of the wolves includes four stages, such as prey encircling, hunting, attacking prey, and exploration. The prey encircling depicts that the grey wolves harass and encircle the prey during hunting. After the finishing of prey encircling, the behavior of hunting is driven by alpha, beta, and gamma as they know the detailed information about the location of prey. When the prey stops their movement, the hunting process ends. At this moment, the wolves assault the prey. Again, grey wolves start tracking and chasing the prey. The chasing of prey is known as an exploration in GWO algorithm.

#### Particle Swarm Optimization (PSO)

The PSO is a powerful meta-heuristic robust evolutionary algorithm for optimization based on the population behavior and was first proposed by Eberhart and Kennedy [62]. The PSO theory was motivated by the social behavior of the fish and birds in groups for optimizing the shortest route to find the food [63]. Recently, the PSO algorithm has been successfully and extensively applied to resolve the non-linear problems in several fields, like Geology [64,65], flood susceptibility modeling [66,67], landslide susceptibility modeling [68,69], forest fire mapping [70,71] because of the higher learning speed and it takes less memory than the other optimization algorithm like genetic algorithm [72,73]. The swarm of particles in the PSO algorithm always tries to find the potential answer to the

problem, which can be best positioned as per the best solution. The particles in PSO travel randomly along with the search space. Based on its own and neighbor's knowledge, a swarm of particles dislocates in search space. The particles become skilled from each other within the group and travel in the direction of their best neighbor based on the obtained knowledge. In a nutshell, the PSO is based on the concept that each swarm of particles changes its location in the search space in order to get the best position or location that it has ever been and the best location nearer its neighbor.

### 2.5.2. Machine Learning Algorithms

#### Artificial Neural Network (ANN)

The artificial neural network-based multilayer perceptron (MLP) model has been widely applied for natural hazard prediction because it can solve non-linear mathematical environmental problems based on the optimization of neurons in each hidden layer [67–69,72–74]. The MLP model is one of the unique elements of ANN, which is characterized by a parallel information system having numerous neurons. These are connected with input, hidden, and output layers [74,75]. During the training process, it can learn tasks without prior knowledge about the problems and identify the similarities among the patterns easily [76]. The modeling process involves several steps, such as (a) transfer function, which depicts a function utilized to the weighted input of neurons to produce the outcome, (b) architecture of the trained network, which describes the configuration of the networks, and (c) common learning law, which includes mathematical algorithms employed to the neural network's connection weights on how to change right after each learning step. In the present study, the MLP architecture was used. Then, the back-propagation algorithm was used to train the MLP architecture. The advantage of the back-propagation algorithm is that it reduces the global error between actual and predicted data during the training process. The sigmoid and linear activation functions have been utilized in the output and hidden layer. Many previous works have been successfully utilized for landslide susceptibility modeling [71–76].

#### Random Forest (RF)

An RF is considered as the ensemble machine learning technique, which creates numerous decision trees to explore the spatial relationship between existing landslide occurrences and landslide triggering factors, which finally constructs a classification [77]. To forecast or predict the output, a set of features have been selected and assigned weights by the outcome of the voting. The majority of the vote, on the basis of the outcome of the evaluated decision trees, has been ensemble and created a single decision tree to generate final classification [78,79]. To avoid the uncertainty problems, the application of a single decision tree generates highly good predictions. The RF model corrects the over-fitting problem of the decision tree during the training process. The vital task in RF classification is to obtain a high variance from several decision trees. Significant components of the RF model during training are to set maximum numbers of trees and variables, which are needed to perform a split search and sampling process [80]. In the present study, 500 trees were set to train the model.

### 2.5.3. Procedure for Optimization

In the present study, the grey wolf and PSO optimization were applied to obtain the best structural parameters of the mentioned machine learning algorithms. The ensemble procedure for the proposed GW-ANN, GW-RF could be as follows: parameters initialization of GW optimization → initialize the random position of “n” grey wolf in “d” dimension → find the fitness value of wolf → stopping criteria met? if it does not meet, starting again from the second stage → update the position of wolf → calculate the fitness of the wolf → find the value of alpha, beta, and delta's position → stopping criteria met? If met, then starts with the training of machine learning algorithms → evaluate the accuracy → meeting stopping criteria? If yes, then the optimal model is obtained. If not meet the stopping criteria,

then starts again with the second steps of GW optimization. The ensemble procedure for the proposed PSO-ANN, PSO-RF could be as follows: parameters initialization of PSO algorithm → training and testing of machine learning algorithm with the initialized parameters → calculation of fitness function → fitness value of each swarm of particle in reference to local and global best values → updating the velocity and position of each swarm of particle accordingly → reaching a maximum number of iteration? If not reached, starting again from the second stage → if it reached the maximum number of iteration, that would be the optimal parameters for the machine learning algorithms.

### 2.6. Validation of the Landslide Susceptibility Models

Validation of the model is the most important task to evaluate the accuracy of results. For that, both validation datasets and the training data should be checked for forecasting [81]. The area under the receiver operating characteristic curve (AUC), was applied to evaluate the model's performance with the ground reality.

The ROC curve is a common and popular technique worldwide for evaluating the landslide susceptibility maps [68–75,82]. In ROC, the false-positive rates and true-positive rates are presented on the X and Y axis, respectively [83]. It distinguishes the trade-offs between two rates. In ROC, the value of area under the curve ranges between 0 and 1.0. Higher the AUC value, higher the accuracy of the model or higher the agreement between the predicted models and ground reality. Talukdar and Pal [84] pointed out that the AUC value higher than 0.70 reflects a good agreement between the predicted model and the ground reality.

### 2.7. Sensitivity Analysis of the Models

#### Parameters Removal and ROC Based Sensitivity Analysis

This method depicts the effect of removing single parameters at one time on the landslide susceptibility modeling. One parameter can be removed at one time, and a new landslide susceptibility map would be constructed based on the rest of the parameters using the best method (selected by the AUC value of ROC curve). In this way, twelve new landslide susceptibility maps were constructed for the twelve parameters by removing parameters one by one. Subsequently, the model's data was extracted based on the testing points. Then, an ROC curve was employed to check the performance of the generated 12 landslide susceptibility maps. Based on the AUC value of the ROC curve, the influence of each parameter for generating the landslide susceptibility models could be obtained. If the AUC decreases by removing particular parameters, which indicates that, the parameter has the power or sensitivity to control or influence the landslide susceptibility models. If the AUC value of the ROC curve achieves higher value, then it indicates that the influence of the parameter on the landslide susceptibility is low.

### 2.8. Spatiotemporal Landslide Hazards Mapping

The landslide hazard modeling is considered as the probability of landslide occurrences within the particular space and time, which damages various resources, lives, and properties. However, the landslide hazard modeling needs to have two important factors, such as landslide susceptibility map and temporal landslide conditioning factors, such as rainfall, snowfall, and earthquake [85]. The present study area observes extreme and average rainfall-induced landslides, therefore, abnormal or extreme rainfall is considered as the major landslide triggering factors. Thus, to construct a landslide hazard model, the extreme or abnormal rainfall was estimated using Gumbel extreme value distribution method [86]. A variety of statistical techniques are available for estimating the rainfall at return periods, such as weibull, log-pearson type-III. Researchers reported that Gumbel extreme value distribution provides a better performance for regions with medium to low rainfall [86]. The study area observes medium to low rainfall. Therefore, the Gumbel method is suitable for estimating rainfall. The rainfall at 2, 5, 10, 20, 50, and 100 years return period were estimated by analyzing the long-term rainfall (1973–2019). The study area has

evenly distributed four rain gauge stations. The study area observes quite similar amounts of rainfall in all gauge stations. Therefore, the inverse distance weighting (IDW) technique can perform good results like kriging under these circumstances. Thus, in the present study, the IDW was employed for preparing interpolated rainfall layers. Spatiotemporal landslide hazard maps were constructed by combining the rainfall layers at different return periods and landslide susceptibility maps following Equation (3).

$$H = P_S \times P_T \quad (3)$$

where, the hazard probability is denoted by  $H$ ,  $P_S$  indicates the landslide susceptibility map generated by the best ensemble metaheuristic machine learning algorithm, and  $P_T$  narrates rainfall layer at different return periods.

### 2.9. Spatiotemporal Landslide Risk Assessment (LRA) Using the Concept of Danger Pixel

In order to construct spatiotemporal LRA in data scarce regions, the concept of the danger pixel and resource map were employed as a semi-quantitative approach. The resource map was generated from the LULC map. Generally, the bare soil, exposed rocks do not provide goods and services; therefore, these classes are considered as non-resource from a monetary point of view. The resource map includes water bodies, sparse vegetation, dense vegetation, built-up area, spruce land. Then, the danger pixel was identified for further research. The pixels having the very high and high landslide susceptibility zones and landslide hazards are considered as the danger pixel because these pixels have the highest chances to have landslide hazard in future, while in the past, these pixels had observed landslides. Therefore, in the present study, we utilized the danger pixel for LRA by following Kanungo et al. [87]. To construct spatiotemporal danger pixel maps, the landslide hazard maps at different return periods were considered. Subsequently, spatiotemporal landslide hazard maps were reclassified into two classes, such as danger pixel and non-danger pixel. For mapping the danger pixel, very high and high landslide hazard zones were merged into one class, i.e., danger pixel. While, rest of the classes, such as very low, low, and moderate landslide hazard maps were merged into one class, i.e., non-danger pixel. In this way, we prepared danger pixel maps for 2, 5, 10, 20, 50, and 100 year return periods.

For evaluating landslide risk, the resource map was combined with danger pixel maps at the different recurrent intervals. The idea behind the integration of resource map and danger pixel was the resource, which sits over the danger pixel, can have higher chances to be damaged. Thus, spatiotemporal landslide risk assessment maps were constructed.

## 3. Results

### 3.1. Landslide Susceptibility Modeling

In order to develop and utilize the ensemble metaheuristic machine learning algorithms-based landslide susceptibility modeling, several steps were followed, such as multicollinearity analysis, building and validation of the landslide susceptibility models, and finally the sensitivity analysis.

#### 3.1.1. Multicollinearity Analysis

Before proceeding for LS modeling, the landslide triggering factors were tested for checking the multicollinearity using TOL and VIF. The presence of collinearity among the parameters causes lower performance of the LS modeling. In the present study, the parameters having TOL < 0.84 and VIF < 4.6 for modeling were considered to be used in the landslide susceptibility modeling (Table 1). Results showed that aspect achieves lowest VIF (1.19), followed by NDVI (VIF = 1.3), LULC (VIF = 1.42), and geology (VIF = 1.45). While the highest VIF value was detected in slope (VIF = 4.6), followed by soil texture (VIF = 2.9), elevation (VIF = 2.19), and curvature (VIF = 2) (Table 1). The selected twelve parameters were not affected by collinearity, and could be utilized for landslide susceptibility modeling.

**Table 1.** Multicollinearity test for analyzing the presence of collinearity among landslide triggering factors.

Parameters	Sig.	Collinearity Statistics	
		Tolerance	VIF
TWI	0.172	0.543	10.843
Geology	0.577	0.687	10.455
Distance to road	0.253	0.543	10.840
Curvature	0.722	0.500	20.000
Aspect	0.438	0.841	10.189
Lineament density	0.258	0.549	10.823
LULC	0.402	0.704	10.420
NDVI	0.354	0.761	10.314
Drainage density	0.415	0.523	10.913
Slope	0.000	0.217	40.617
Soil texture	0.000	0.455	20.197
Elevation	0.000	0.472	20.119

### 3.1.2. Configuration of the Machine Learning Algorithms

The ANN and RF algorithms were utilized for landslide susceptibility modeling. Therefore, for modeling LS, the training database (80% of the total data), which contains twelve landslide triggering factors and dependent factor or past landslide and non-landslide locations, were utilized to train the ANN and RF models. The optimum values of the parameters of the ANN and RF models were set by trial and error process to obtain the high-quality LS maps. Since the ANN model has multilayer architecture, therefore, the configuration would also be manifold. First, twelve landslide triggering factors were defined as the input layer. The most controversial parameter, i.e., hidden parameter, has to be selected, albeit there is no standard way to design the hidden layer. However, some researchers claimed that one hidden layer could solve all classification problems [36]. Then the output layer was a two-class layer, such as landslide and non-landslide. In the present study, one hidden layer with ten neurons were selected by trial and error process. In addition, the sigmoid activation function was selected. Thus, the optimal network topology was set for ANN-based LS modeling. Similarly, the RF model was also optimized. 300 trees and 2 *mtry* (the number of variables tested at each node) were optimized for the RF model.

### 3.1.3. Metaheuristic Optimizations for Configured Machine Learning Algorithms

Two metaheuristic algorithms, such as GWO and PSO were employed to optimize the configured ANN and RF models. The iteration for all models (GW-ANN, GW-RF, PSO-ANN, and PSO-RF) were set to 1000 by trial and error process because the error during the training process showed very slight changes or no changes. Although the training process for all models was performed at 400, 500 and 800 iterations and gradual changes of errors were observed. Then, 50 and 45 grey wolves were selected for ANN and RF models. 35 particles of swarm for both ANN and RF were set. The detailed optimization of the GWO and PSO for both ANN and RF were presented in Table 2.

**Table 2.** Setting of model parameters for optimizing the machine learning algorithms by PSO and GWO.

Optimization Algorithms	Optimized Model Parameters
PSO	Swarm size—30; iteration—1000; mutation type-bit-flip, mutation probability—0.01; inertia weight—0.33; social weight—0.33; individual weight—0.34; report frequency—20; seed—5
GWO	Number of population— 50 (for ANN) and 45 (for RF); absorption coefficient of the firefly members—0.001; accelerate type-normal, iteration—1000; seed-6, escape probability—0.8, mutation probability—0.01; chaotic coefficient—4.0

### 3.1.4. Model Validation and Comparisons

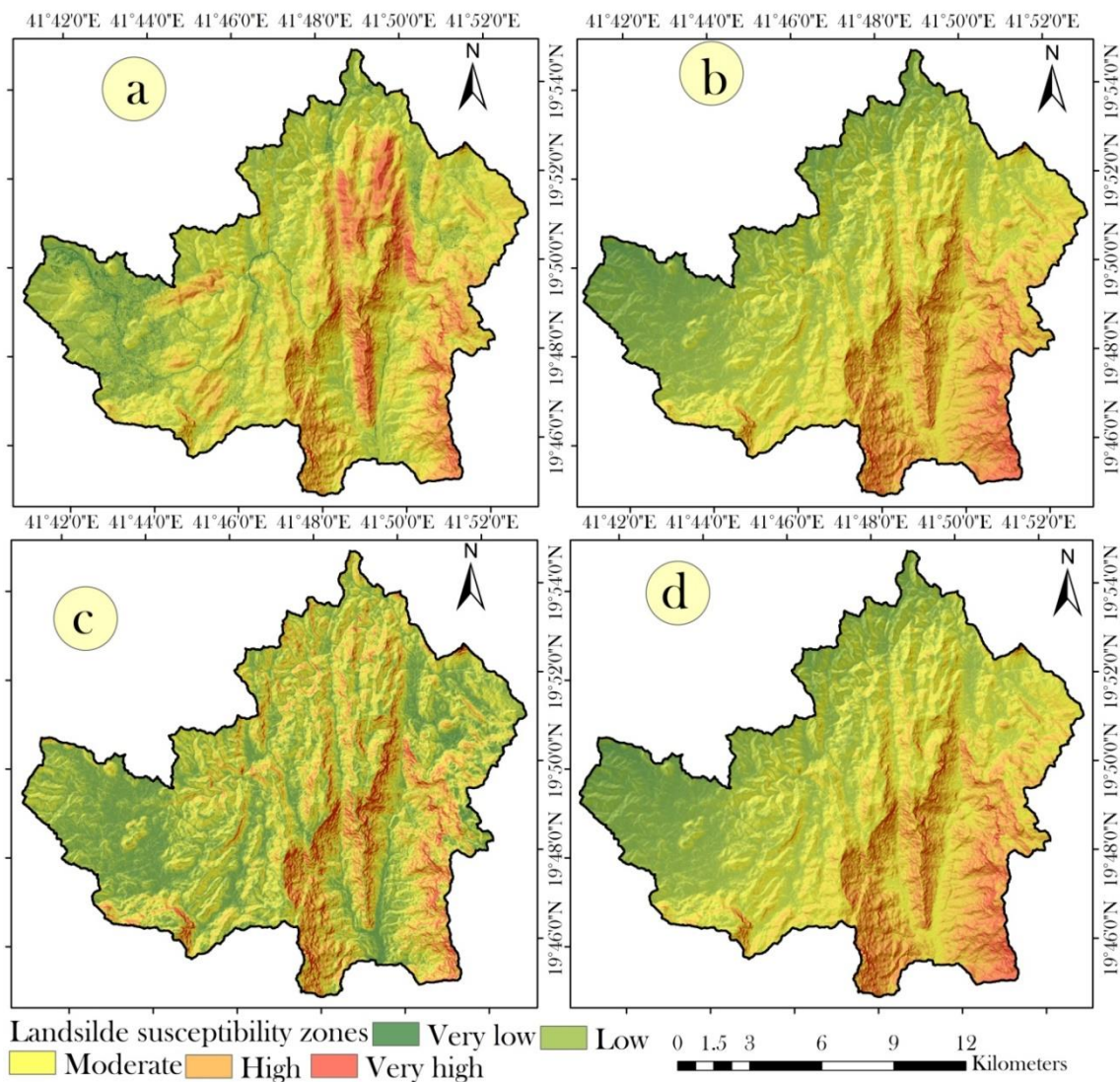
The performances of the four ensemble models were evaluated by comparing the actual data with predicted data. For evaluating the performances of the four models, several indexes, such as correlation coefficient (R), Kendall's tau (T), Spearman's rho, MAE, and RMSE were employed. The results showed that higher correlation and lower different error values are observed for four models. The GW-ANN model (R = 0.985, T = 0.79, rho = 0.897, MAE = 0.029, and RMSE = 0.067) outperformed other models, followed by GW-RF (R = 0.97, T = 0.71, rho=0.86, MAE = 0.035, and RMSE = 0.103), PSO-RF (R = 0.95, T = 0.7, rho = 0.85, MAE=0.04, and RMSE = 0.11), and PSO-ANN (R = 0.94, T = 0.68, rho = 0.82, MAE = 0.061, and RMSE = 0.12). Findings showed that all ensemble machine learning models perform well and can be used for generating the LS mapping. Based on the performance, the ensemble models could be arranged as GW-ANN > GW-RF > PSO-RF > PSO-ANN.

### 3.1.5. Generation of the Landslide Susceptibility Maps

Figure 4 showed the constructed landslide susceptibility maps based on four ensemble metaheuristic machine learning algorithms. The generated LS maps have a continuous scale or stretch format, and there is a need to separate these values into possible classes. A variety of classifiers systems are available for the conversion of the continuous maps into a classified map, such as natural break, equal interval, quantile, and standard deviation. The researcher recommended that natural break classifier is robust, efficient and consistent to predict [85]. Subsequently, landslide susceptibility maps, based on Jenks' natural break (1967) algorithm, were classified into five subclasses, such as very high, high, moderate, low, and very low. The method of Jenks optimization, also known as natural break classification, is a data segmentation process to estimate the optimal value structure of the different classes. This method intends to decrease the average deviation from the mean class value while increasing the deviation from the mean of other classes. The method thus minimizes intra-class and maximizes inter-class variance. According to GW-ANN model, 6.14 km<sup>2</sup> and 27.27 km<sup>2</sup> area were predicted as very high and high LS zones, followed by moderate (47.86 km<sup>2</sup>), low (59.97 km<sup>2</sup>), and very low (57.07 km<sup>2</sup>) LS zones. According to GW-RF, PSO-ANN, and PSO-RF models, 15.13 km<sup>2</sup>, 16.03 km<sup>2</sup>, and 15.03 km<sup>2</sup> areas were predicted as a very high LS zone (Table 3). On the other hand, based on GW-RF, PSO-ANN, and PSO-RF models, 35.06 km<sup>2</sup>, 41.32 km<sup>2</sup>, and 35.73 km<sup>2</sup> area come under very high LS zones. According to GW-RF, PSO-ANN, and PSO-RF models, 40.40 km<sup>2</sup>, 18.40 km<sup>2</sup>, and 40.31 km<sup>2</sup> area come under the very low LS zone (Figure 4).

**Table 3.** Calculation of area coverage under different landslide susceptibility categories.

Landslide Susceptibility Zones	Area (km <sup>2</sup> )			
	GW-ANN	GW-RF	PSO-ANN	PSO-RF
Very low	57.07	40.40	18.40	40.31
Low	59.97	54.25	61.33	54.76
Moderate	47.86	53.60	61.45	52.61
High	27.27	35.06	41.32	35.73
Very high	6.14	15.13	16.03	15.03

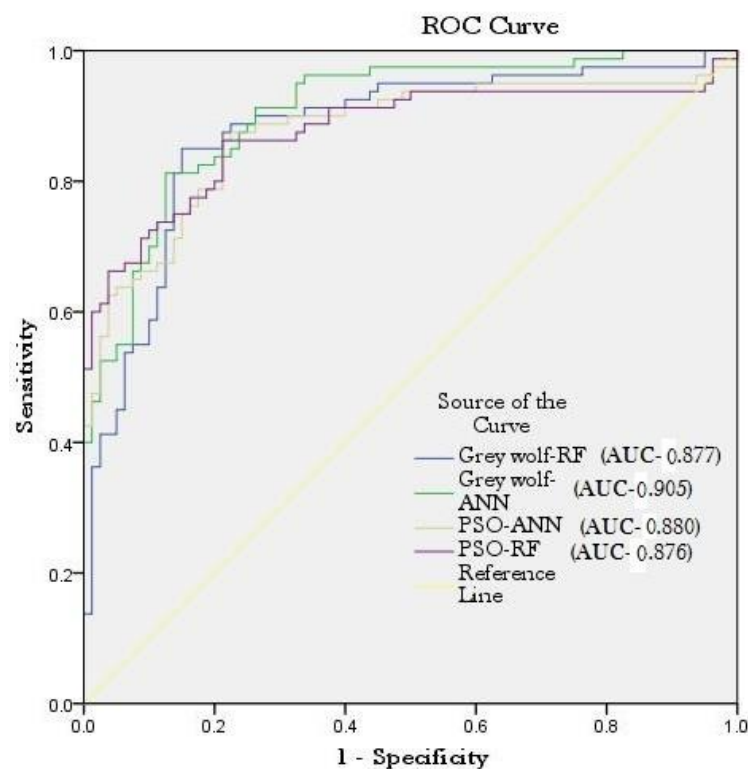


**Figure 4.** Landslide susceptibility models based on ensemble metaheuristic machine learning algorithms, such as (a) PSO-ANN, (b) PSO-RF, (c) GW-ANN, and (d) GW-RF.

### 3.1.6. Validation of LS Maps

Four landslide susceptibility models were validated using the testing datasets. The ROC curve was employed to evaluate the performance of four LS models. A higher AUC value of the ROC curve suggests a higher accuracy or agreement between the LS model and ground reality. Based on the AUC value of ROC curve, GW-ANN model appeared as the best representative model (AUC = 0.905), followed by PSO-ANN (AUC = 0.880), GW-RF (AUC = 0.877) and PSO-RF (AUC = 0.876) (Figure 5). Results showed that all ensemble machine learning models perform well for modeling landslide susceptibility. According to the AUC values, the LS models can be rearranged as GW-ANN > PSO-ANN > GW-RF > PSO-RF.

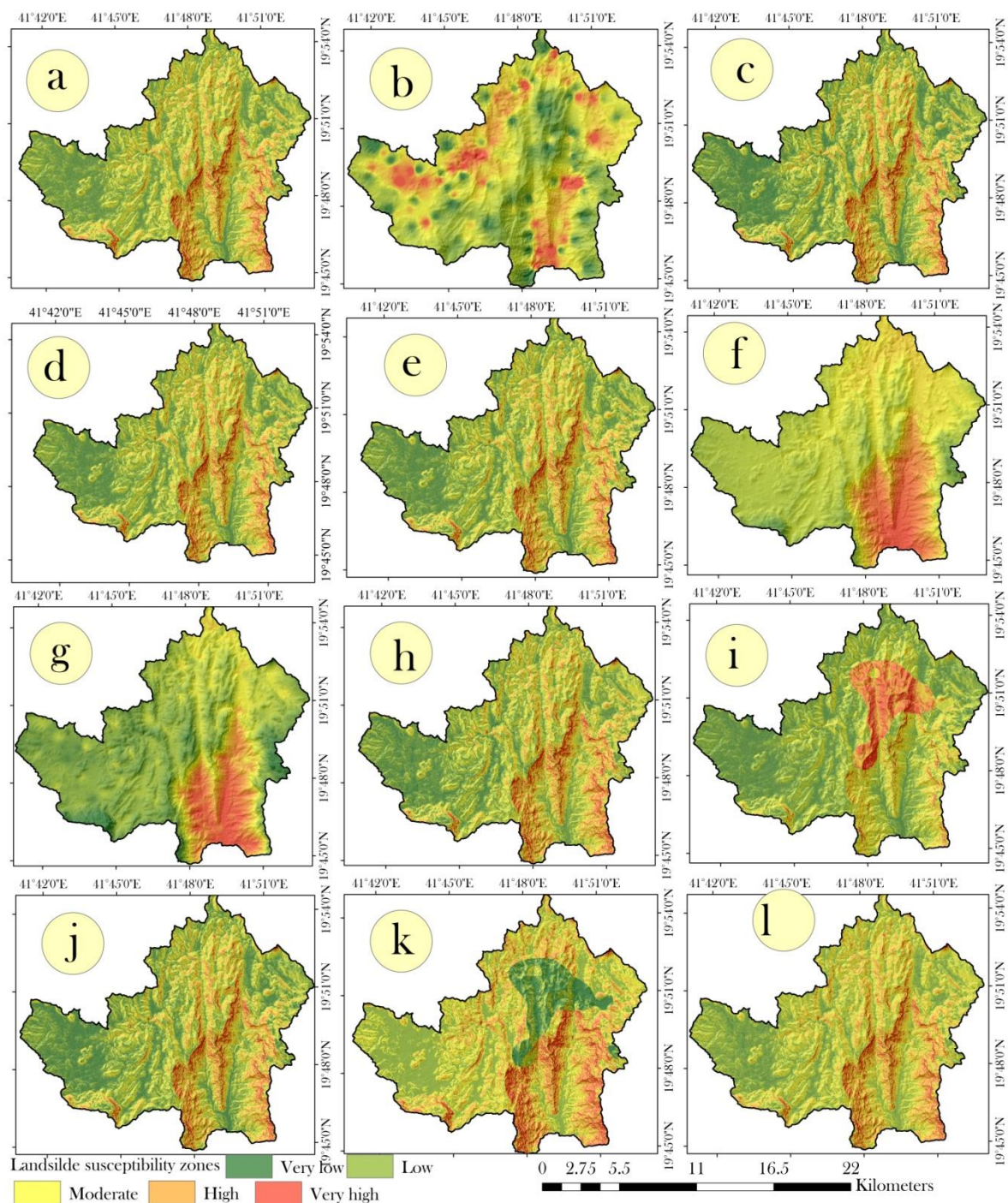




**Figure 5.** Validation of landslide susceptibility models using receiver operating characteristics (ROC) curve.

### 3.1.7. Sensitivity Analysis of LS Model

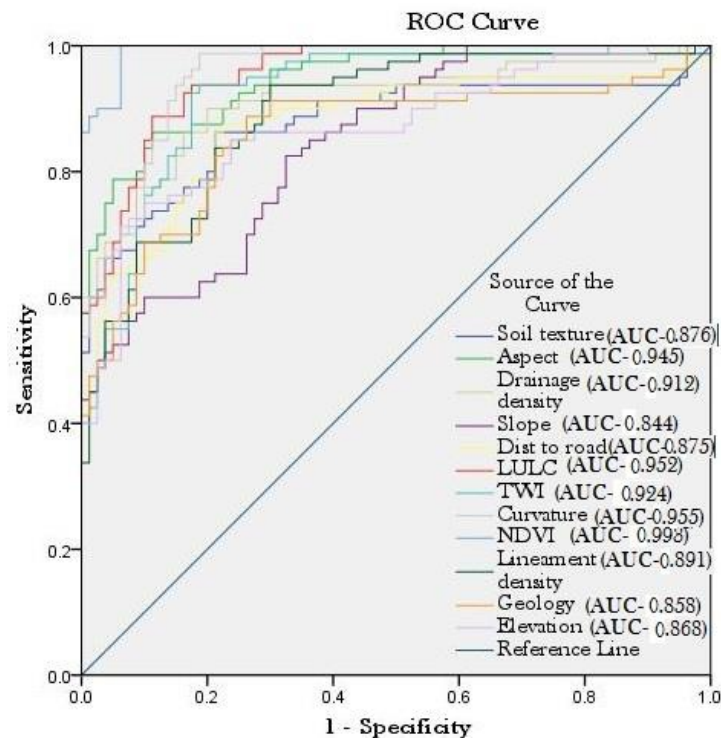
In the present study, ROC based sensitivity analysis was developed and utilized for exploring the influence of the landslide triggering factors for LS modeling. Identification of the parameters, which are mainly responsible for landslide occurrence, is a vital task for proposing disaster management plans. The ROC based sensitivity analysis is a complex process to judge the landslide triggering factors. In order to construct ROC based sensitivity analysis, several steps have to be followed. First, we used the best ensemble machine learning algorithm, i.e., GW-ANN for LS modeling by excluding individual parameters at one time. Thus, twelve LS models were constructed by repeating the same procedure (individual parameter exclusion at one time) Figure 6a–l showed the twelve LS models based on eleven parameters. For instance, Figure 6a showed the LS model based on eleven parameters, excluding elevation, similarly Figure 6b showed LS model excluding the slope parameter. Very high and high LS zones were predicted differently for twelve different models based on 12 sets of data. For example, GW-ANN predicted 6.14 km<sup>2</sup> area as very high LS zones (Figure 4a), while LS model prepared using map removal technique with the ROC based sensitivity analysis (excluding elevation, slope, geology, LULC, distance to road individually) predicted 6.28 km<sup>2</sup>, 12.08 km<sup>2</sup>, 6.52 km<sup>2</sup>, 7.09 km<sup>2</sup>, and 12.75 km<sup>2</sup> (Figure 6) respectively. Therefore, it could be stated that the exclusion of slope and distance to road increases the area of very high LS zones.



**Figure 6.** Landslide susceptibility modeling based on GW-ANN (best model) by excluding individual parameter at one time, such as (a) elevation, (b) slope, (c) curvature, (d) aspect, (e) NDVI, (f) soil texture, (g) lineament density, (h) TWI, (i) Geology, (j) LULC types, (k) distance to road, (l) drainage density.

Thus, these two parameters are the most dominant and sensitive parameters for landslides. In addition, twelve LS models were validated using the ROC curve based on testing datasets. Results showed that the LS model excluding NDVI (AUC = 0.993) outperforms other models indicating significantly less sensitivity. On the other hand, the LS model excluding curvature achieves the second best model (AUC = 0.955), followed by LULC (AUC = 0.952), TWI (AUC = 0.924) (Figure 7). The results showed that the LS models perform better if the mentioned parameters are excluded individually at one time. Therefore, it could be stated that these parameters are less sensitive to landslide susceptibility. While the

AUC of ROC curve showed that LS model excluding slope parameter achieves lowest AUC value (AUC = 0.844), followed by Geology (AUC = 0.858), elevation (AUC = 0.868), distance to road (AUC = 0.875), and soil texture (0.876) (Figure 7). These parameters are considered as highly sensitive to a landslide. During modeling, if we exclude one parameter, the performance of the LS model would be very bad. These parameters are the dominant causes for the landslide.



**Figure 7.** Validation of map removal based landslide susceptibility (LS) models using ROC curve.

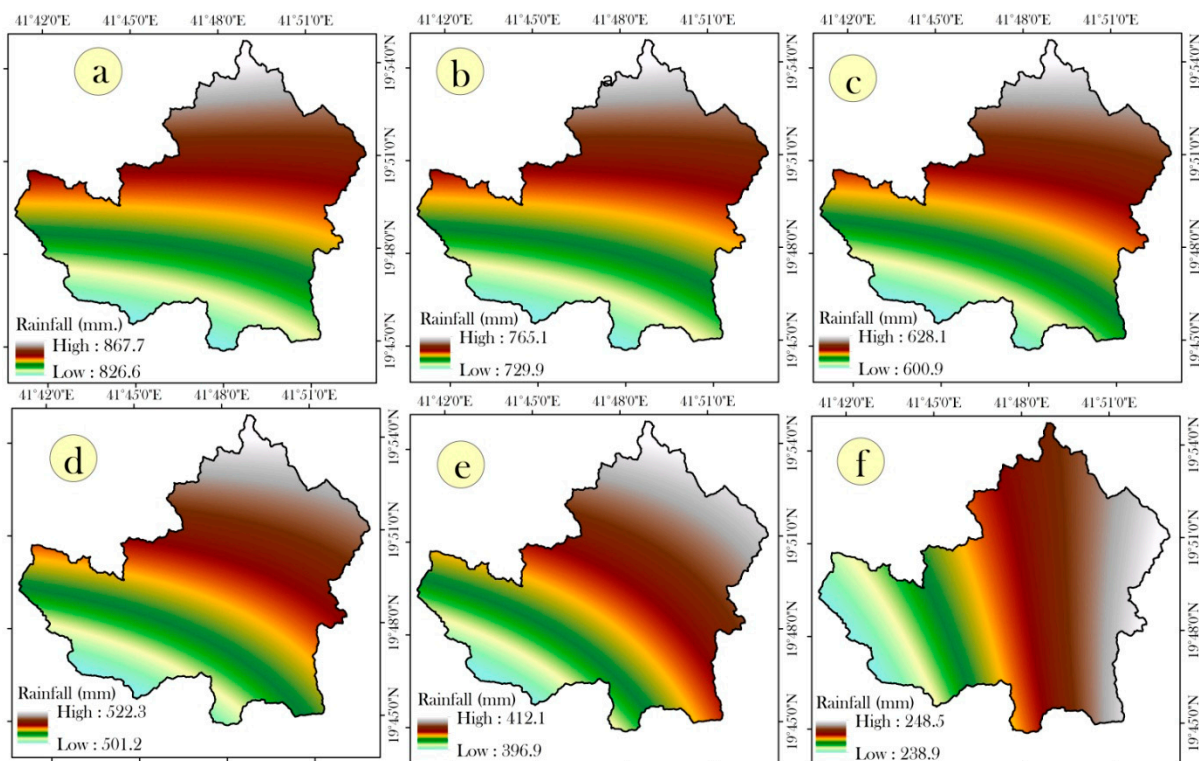
According to sensitivity analysis, the slope, elevation, geology, distance to road, and soil texture were considered as the most effective and sensitive parameters for LS modeling, while curvature, LULC, and NDVI appeared as least sensitive parameters. The improper road construction along the southern and southeastern part of the study area was the most effective reason for frequent landslides in the study area. Therefore, the LS models showed these regions as the very high landslide susceptible zones. Distance to the road excluded LS model showed very high landslide susceptible regions in the south and eastern part of the study area (Figure 6k). Eastern, Southern, and middle part of the study have a very high slope ( $67^{\circ}$ – $81^{\circ}$ ) (Figure 3) and very high–high landslide susceptible zones (Figure 4). While, the slope excluded LS model showed very high and high landslide susceptible zones in the western and northern part of the study area, which is contradictory results (Figure 6b). Very high and high landslide susceptibility regions (Eastern, Southern, and middle part) of the study area are characterized by Quartzite, metasiltstone, Conglomerate, feldspathic greywacke, black marble and white marble, where joints are frequently observed (Figure 3e). Therefore, this region is highly exposed to weathering and potential seats of erosion, which caused slope failure and landslides. While the geology excluded LS model indicated some changes in landslide susceptibility patterns (Figure 6i). Therefore, the selected sensitive parameters have indeed affected LS modeling.

### 3.2. Landslide Hazards Modeling

In order to construct landslide hazard models, several steps were followed, such as estimation of rainfall at different recurrent intervals or return periods, and integration of landslide susceptibility models with them.

### 3.2.1. Estimation of Rainfall at Different Return Periods

The rainfall data was collected from four meteorological gauge stations in the study area from 1983 to 2019. The Gumble extreme value distribution method was employed to estimate rainfall at 2, 5, 10, 20, 50, and 100 years return periods. The probability of rainfall occurrences at 2 year return period was 50%, followed by 20% at 5 year, 10% at 10 year, 4% at 20 year, 2% at 50 year, and 1% at 100 year return periods. Highest and lowest rainfalls at the 2 year return period were 248.5 mm. and 238.9 mm. (Figure 8). The highest rainfall at 5, 10, 20, 50, and 100 year return periods were 442.1 mm, 522.3 mm, 628.1 mm, 765.1 mm, and 867.7 mm, respectively (Figure 8). Similarly, the lowest rainfall at different return periods was also gradually increased, albeit the chances of occurrences had low probability.

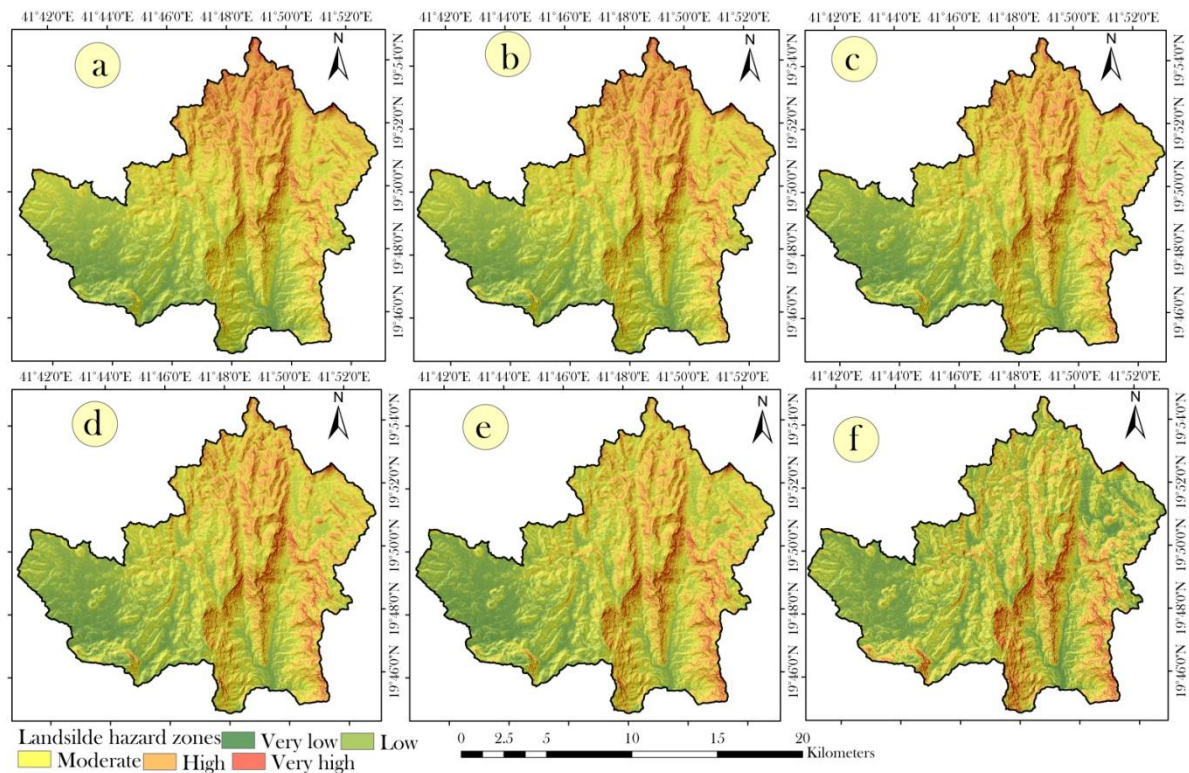


**Figure 8.** Estimation of extreme rainfall at (a) 100, (b) 50, (c) 20, (d) 10, (e) 5, and (f) 2 year return periods using Gumble extreme value distribution methods.

### 3.2.2. Generation of Landslide Hazard Models

In order to construct landslide hazard models, the landslide susceptibility map and rainfall layers were integrated. In the present study, the landslide susceptibility map (LS model constructed by excluding NDVI parameter, as it showed higher AUC value), which predicted by GW-ANN as it appeared as best performing model, was integrated (multiplied) with rainfall layers at different return periods. Thus, landslide hazard map at 2, 5, 10, 20, 50, and 100 year return periods were produced. Then, generated landslide hazard maps were classified into five sub-classes using natural break algorithm, such as very high, high, moderate, low, and very low hazard zones (Figure 9). Results showed that area under very high, and high landslide hazard zones were gradually increased at 2 to 100 year return periods. According to landslide hazard map at 2 and 100 year return period, 6.09 and 20.44 km<sup>2</sup> area come under very high landslide hazard zone, followed by high (26.28 and 40.52 km<sup>2</sup>) moderate (47.93 and 49.71 km<sup>2</sup>), low (61.01 and 50.27 km<sup>2</sup>), and very low (56.72 and 37.29 km<sup>2</sup>) hazard zones (Table 4). On the contrary, an area under very low and low landslide hazard zones have been decreased at 2–100 year return periods. For instance, area under very low landslide hazard area was 56.72, 43.23, 42.32, 41.1, 37.47, and

37.29 km<sup>2</sup> at 2, 5, 10, 20, 50, and 100 year return periods (Table 4). Based on the findings, it could be stated that if intense rainfall increases in the study area, the landslide affected area would be increased accordingly. The predicted landslide hazard-prone areas should be paid more attention to reduce damages in future.



**Figure 9.** Landslide hazard modeling at (a) 100, (b) 50, (c) 20, (d) 10, (e) 5, and (f) 2 year return periods.

**Table 4.** Area coverage under different landslide hazard zones at different return periods.

Landslide Hazard Zones	Area (km <sup>2</sup> ) in Different Return Periods					
	2 Year	5 Year	10 Year	20 Year	50 Year	100 Year
Very low	56.72	43.23	42.32	41.10	37.47	37.29
Low	61.01	54.91	50.88	50.41	48.91	50.27
Moderate	47.93	53.83	53.23	51.67	50.51	49.71
High	26.28	37.59	40.26	40.02	42.36	40.52
Very high	6.09	8.90	11.53	15.03	18.97	20.44

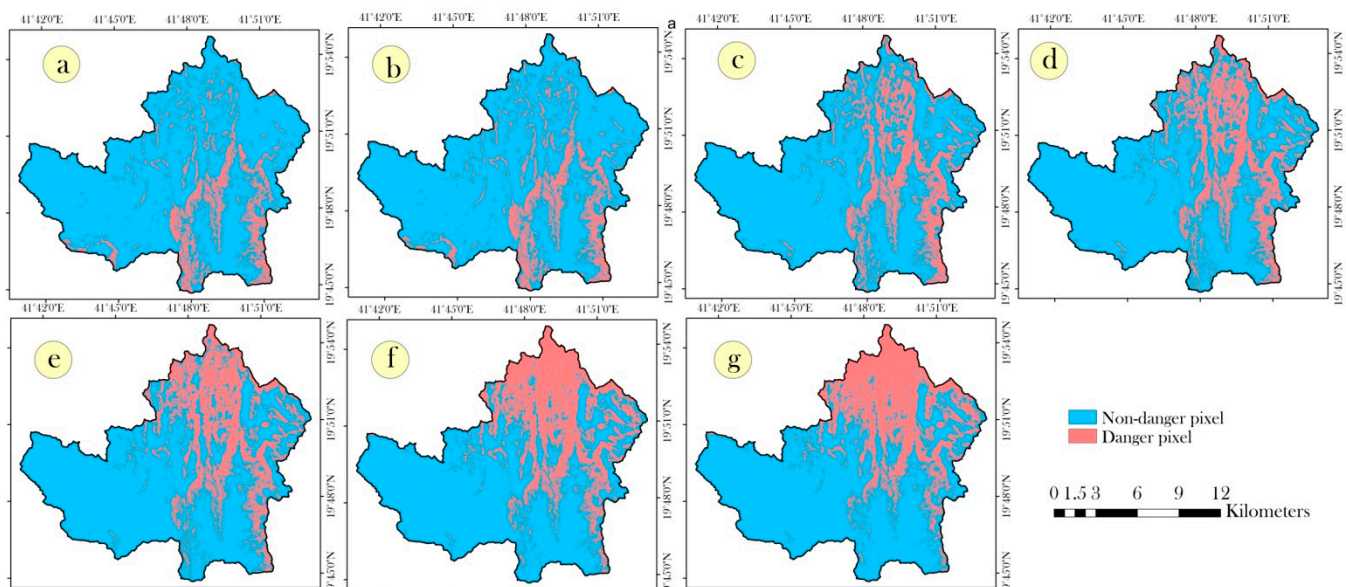
### 3.3. Landslide Risk Assessment Using the Theory of Danger Pixel

In order to assess landslide risk, a variety of socio-economic and landslide damage data are highly required. The present study area is data-scarce regions; therefore, the socio-economic and landslide damage data are not available and very difficult to collect. Considering the data scarcity, a semi-quantitative approach like the theory of danger pixel was adopted to assess the landslide risk. Following Kanungo et al. [87], the theory of danger pixel and resource map were combined to produce landslide risk assessment maps at different return periods.

#### 3.3.1. Identification of Danger Pixel

The danger pixel was identified from the landslide hazard maps at different return periods. Each pixel of landslide hazard maps contains information regarding the magnitude of potential hazards caused by landslides. As a consequence, the areas or pixels under very high and high hazard zones have a very high potentiality to be exposed to landslides

and could be observed with very high magnitudes of damages. Therefore, based on this concept, danger pixels were extracted from very high and high landslide hazard zones at 2, 5, 10, 20, 50, and 100 year return periods. Results showed that danger pixels are 37.28, 48.21, 53.55, 56.81, 63.07, and 62.57 km<sup>2</sup> at 2, 5, 10, 20, 50, and 100 return periods, respectively (Figure 10).



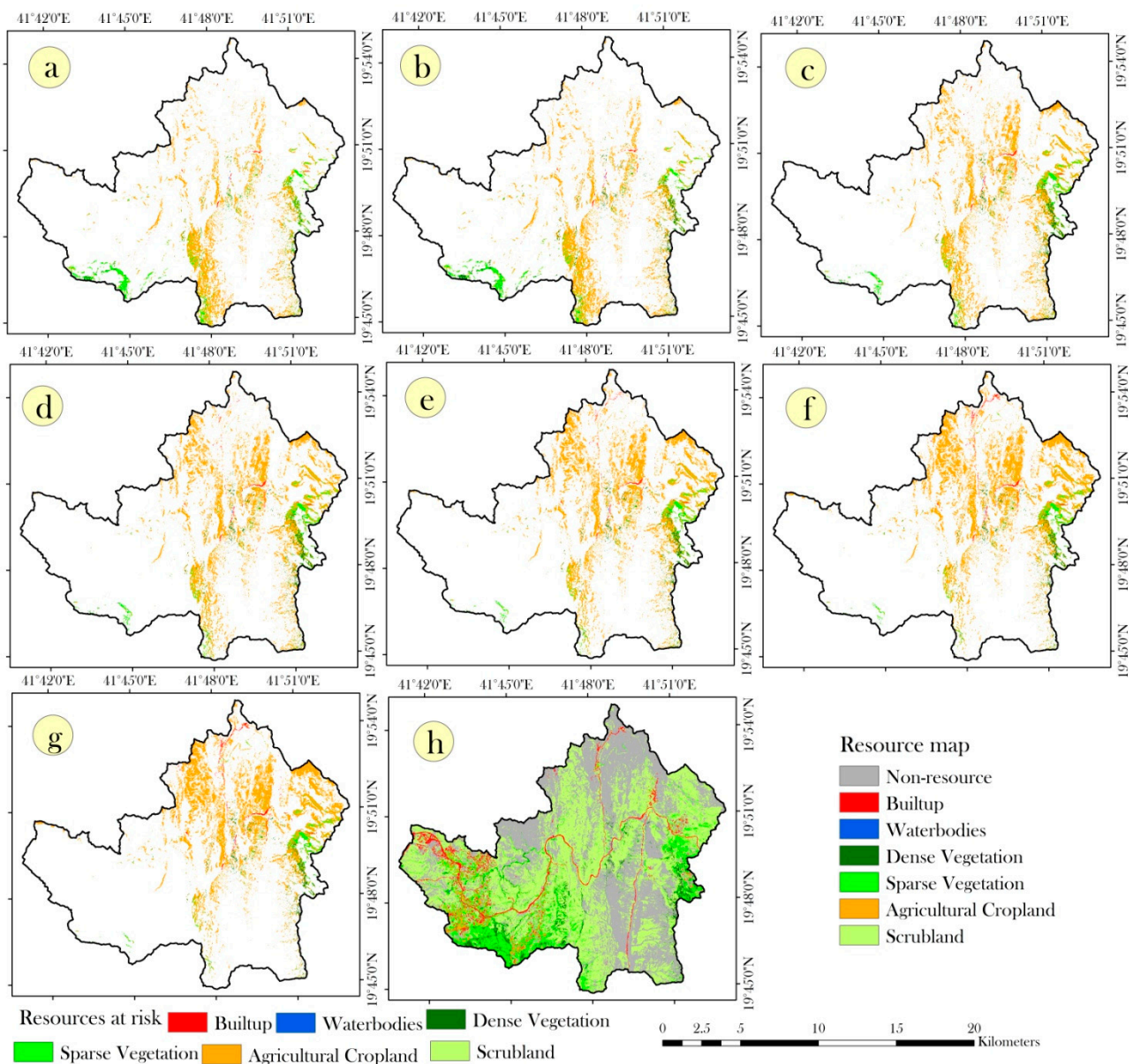
**Figure 10.** Identification of danger and non-danger pixel at (a) GW-ANN model, (b) 2, (c) 5, (d) 10, (e) 20, (f) 50, and (g) 100 year return periods.

### 3.3.2. Risk Assessment by the Integration of Danger Pixel and Resource Map

The spatiotemporal landslide risk assessment maps were constructed by integrating with the resource map and danger pixel. The identification of the resource map is the most important element for risk assessment. In the present study, natural resources and human made resources were identified from the generated LULC map. All LULC types cannot be considered as a resource because bare soil, wadis debris, and exposed rocks in the study area do not provide important goods and services. On the contrary, the built-up, water bodies, dense vegetation, sparse vegetation, and scrubland in the study area provide a variety of goods and services. Therefore, the resource map was prepared by excluding bare soil, wadis debris, and exposed rocks from the LULC map. The final resource map contains various natural and human made resources, such as built up, water bodies, dense vegetation, sparse vegetation, and scrubland. Subsequently, the resource map was integrated with the danger pixels at different return periods. The resources located in the danger pixel can have a very high potentiality to be damaged if a landslide occurs. The final risk map at different return periods showed that scrubland is the most affected resource. The area under scrubland at risk was 96, 19.02, 24.89, and 24.56 km<sup>2</sup> at 2, 5, 50, and 100 year return periods (Table 5 and Figure 11), respectively. The least affected resources were water bodies and agricultural land, albeit the area of these resources at risk also increased over time (2–100 year return periods). Therefore, based on the findings, it could be stated that intense rainfall can trigger the landslide at a large scale, which can cause huge damages. Therefore, the identified risk element and their location should be maintained very carefully to minimize the damages in future.

**Table 5.** Area coverage of different resources at risk at different return periods and best model.

Resources at Risk	Area (km <sup>2</sup> ) in Different Return Periods						
	GW-ANN	2 Year	5 Year	10 Year	20 Year	50 Year	100 Year
Built up	0.14	0.16	0.26	0.31	0.40	0.56	0.63
Waterbodies	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Dense vegetation	1.43	1.53	1.43	1.39	1.32	1.28	1.16
Sparse vegetation	2.78	3.00	2.60	2.45	2.33	2.26	2.13
Agriculture cropland	0.01	0.02	0.01	0.01	0.01	0.01	0.02
Scrubland	13.65	14.96	19.02	21.02	22.15	24.89	24.56



**Figure 11.** Landslide risk assessment for (a) GW-ANN model, (b) 2, (c) 5, (d) 10, (e) 20, (f) 50, and (g) 100 year return periods, and (h) resource map of the study area.

#### 4. Discussion

The present study was designed to assess the spatiotemporal landslide risk in Aqabat Al-Sulbat Asir region of Saudi Arabia. Four ensemble metaheuristic machine learning

algorithms, such as GW-ANN, GW-RF, PSO-ANN, and PSO-RF were developed and tested for landslide susceptibility modeling. The generated LS models were validated using the ROC curve. GW-ANN model appeared as the best representative for LS modeling, followed by PSO-ANN, GW-RF, and PSO-RF models. All models achieved AUC value  $>0.8$  indicating the good performance of all models. ROC based sensitivity analysis was developed to investigate the importance of the landslide triggering factors for landslide susceptibility modeling. The slope, LULC, geology, elevation, and texture appeared as the dominant and sensitive factors for landslide susceptibility. Landslide hazard models at 2–100 year return periods were constructed by integrating with the best landslide susceptibility model and extreme rainfall at 2–100 year return periods. The study area lacks socio-economic and damage data, therefore, a semi-quantitative approach, like the theory of danger pixel was employed to assess the landslide risk. The danger pixels at 2–100 year return periods were combined with the resource map of the study area to produce landslide risk assessment maps at 2–100 year return periods. The assessment of long-term landslide risk maps will be helpful to planners and government authorities to take disaster mitigation strategies. Some drawbacks existed in the present work, which could be investigated further.

#### 4.1. Landslide Susceptibility Modeling

One of the major aims of the study was to develop ensemble metaheuristic machine learning models to construct landslide susceptibility mapping. During the training process, GW-ANN model had a high correlation between actual and predicted data and low error. Moayedi et al. [88] reported high correlation and low error between actual and predicted landslide data using metaheuristic models. Results of the present study showed that GW-ANN model outperforms other models for LS modelling as per the AUC value during the validation process. Chen et al. [89] reported that grey wolf optimizer performs better than whale optimization algorithm for landslide susceptibility mapping. Li et al. [90] accounted that PSO-ANN outperforms other models for LS modeling in Shicheng County in China, and it was identical to our findings. Because, in the present study, PSO-ANN appeared as the second-best model for LS modeling and had very slight differences with GW-ANN in terms of performance. Xi et al. [91] also reported that PSO based ANN model performs better for LS modeling than other models. Pham et al. [92] reported that PSO-RF appears as the best model for LS modeling and observes lower error during the training process (RMSE = 0.453) and a higher correlation between actual and predicted data ( $R = 0.89$ ). This work is identical with our findings for LS modeling, although the error was less in the case of our study. Therefore, it could be stated that all ensemble metaheuristic machine learning algorithms perform better for LS modeling.

#### 4.2. Sensitivity Analysis

ROC based sensitivity analysis was introduced and employed for investigating the performance of parameters for landslide susceptibility. However, rare studies have been carried out for sensitivity analysis at a spatial scale [93]. Ilia and Tsangaratos [94] employed changing criteria weights methods for sensitivity analysis. They changed the weights of input parameters to generate the LS models and reported good results. In the present, nothing was changed during modeling, while we just excluded individual parameters at one time. Therefore, the influence of the parameters for landslide susceptibility could be identified accurately. On the other hand, no researchers have yet validated the LS models produced by excluding individual parameters at one time. Generally, researchers have compared the area coverage under different landslide susceptibility categories with each other and original models. In the present study, the generated twelve models were validated using the ROC curve. The higher AUC value of ROC curve indicates lower sensitivity for LS modeling and vice-versa. Pal and Paul [93] conducted a similar type of ROC based sensitivity analysis for wetland vulnerability studies. However, during modeling they gradually excluded each parameter one by one until the AUC values



reached to 0.7. The low AUC value of the models could automatically be produced if each parameter excludes one by one. Therefore, the influence of the input parameters cannot be identified accurately. Based on the discussion, it could be stated that ROC based sensitivity analysis can be used in different studies for conducting sensitivity analysis.

#### 4.3. Landslide Risk Assessment Analysis

Research on landslide risk assessment is not new work. Many pieces of research on landslide risk assessment have already been conducted in mountainous regions [95–100]. Researchers reported that estimation of potential risk directly associated with the landslide susceptibility maps, spatial-temporal probability and intensity of events. Many researchers have employed socio-economic data for landslide risk assessment [52,53,101–103], but in data-scarcity regions like Saudi Arabia, socio-economic, and damage data are not available and very difficult to collect. Therefore, to model risk assessment, the present study employed the theory of danger pixels. Althuwaynee and Pradhan [104] employed the theory of danger pixels for landslide risk assessment in Kuala Lumpur city, which was identical with the present study. Kanungo et al. [87] incorporated the theory of danger pixel for evaluating the landslide risk assessment. In addition to this, very rare studies on landslide risk assessment at a temporal scale have been conducted [105]. In the present study, landslide risk assessment at 2–100 year return periods were constructed to propose long-term planning for reducing the damages of resources. Based on the above discussion, it could be stated that the present study provides a comprehensive work on rainfall induced landslide risk assessment.

Prediction of long-term risk due to rainfall induced landslides can provide lots of information about the upcoming damages of natural resources and artificial resources if different magnitudes of landslide occur. If authorities take serious steps by considering the prediction of long-term risk assessment, economic loss, natural resource destruction and human lives loss would be possible to minimize. Therefore, sustainable development along the landslide affected regions could be possible. In fact, this study provides huge information, which are required for proposing landslide management plans for a long time.

## 5. Conclusions

The present study provides comprehensive knowledge of the development of ensemble metaheuristic machine learning algorithms, such as GW-ANN, GW-RF, PSO-ANN, and PSO-RF for LS modeling. The very high LS zone covered by the 6 km<sup>2</sup>–16 km<sup>2</sup> area, according to the four LS models. The LS models were validated using the ROC curve. The GW-ANN (AU = 0.905) model appeared as the best representative model for LS modeling, followed by PSO-ANN, and GW-RF. ROC based sensitivity analysis was proposed to explore the importance of the input parameters for LS modeling. Slope, Geology, LULC, and elevation were reported to be the most dominating and sensitive parameters for LS modeling. The landslide hazard maps were prepared at 2–100 years return periods. Very high landslide hazards zone increased gradually over the increasing of return periods (6 km<sup>2</sup> at 2 year return period to 40 km<sup>2</sup> at 100 year return periods). While the area under very low landslide hazard decreased at 2–100 year return periods. Finally, the resource risk assessment models at 2–100 year return periods were produced by an integration of the theory of danger pixels. Scrubland, dense vegetation, sparse vegetation and built-up area were at risk of a landslide. The area coverage of these resources at risk increased over the increasing return periods.

The present study provides a detailed framework of the development of four ensemble metaheuristic machine learning algorithms, which can be utilized for predicting the different natural hazards, such as flood susceptibility, gully erosion, fire susceptibility etc. The proposed ROC based sensitivity analysis can be incorporated in other studies for sensitivity analysis. Finally, landslide risk assessment maps provided the spatial map of resources which are at risk at 2–100 year return periods. Therefore, huge information regarding the identification of resources, which are at risk, can be obtained from spatiotemporal risk

maps. This work can be helpful to the planners to propose disaster management strategies so that future damages due to landslides could be under control. The present study has some limitations, such as the utilization of moderate resolution satellite images instead of high-resolution image, use of a semi-quantitative approach due to data scarcity instead of quantitative approach, rainfall layer using fewer numbers of rainfall gauge stations, and use of a small number of landslide locations for modeling instead of more data. The future study on risk assessment can be improved if the mentioned issues resolve.

**Author Contributions:** Conceptualization, J.M. and S.A.; data curation, J.M., M.A. (Mohd. Ahmed) and R.A.K.; formal analysis, J.M., S.A., S.T., M.A. (Majed AlSubih) and M.A. (Mohd. Ahmed); funding acquisition, M.A. (Majed AlSubih); investigation, S.T., M.A. (Mohd. Ahmed) and N.B.K.; methodology, J.M., S.A., S.T. and R.A.K.; project administration, J.M., S.A., M.A. (Majed AlSubih) and S.M.A.; resources, S.A., M.A. (Mohd. Ahmed), R.A.K., N.B.K. and S.M.A.; software, J.M. and S.T.; supervision, J.M. and S.A.; visualization, M.A. (Majed AlSubih); writing—original draft, J.M., S.A. and S.T.; writing—review and editing, J.M., M.A. (Mohd. Ahmed), N.B.K. and S.M.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** Funding for this research was given under award numbers R.G.P1/173/41 by the Deanship of Scientific Research King Khalid University, Ministry of Education, Saudi Arabia.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Publicly available sentinel datasets were analyzed in this study. This data can be found here: <http://earthexplorer.usgs.gov>.

**Acknowledgments:** The authors extend their appreciation to the Emirate of Aseer province- Saudi Arabia, Ministry of Transportation-Saudi Arabia, Mohammed K. Almesfer Dean of College of Engineering, King Khalid University Saudi Arabia and Deanship of Scientific Research at King Khalid University for their administrative and technical support The Sentinel-2 data have been downloaded from the earth explorer system of USGS from their website: <http://earthexplorer.usgs.gov/>. We are grateful for this service.

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

## References

1. Azimi, S.R.; Nikraz, H.; Yazdani-Chamzini, A. Landslide Risk Assessment by using a New Combination Model based on a Fuzzy Inference System Method. *KSCE J. Civ. Eng.* **2018**, *22*, 4263–4271. [[CrossRef](#)]
2. Sarkar, S. Landslide Risk Assessment in Parts of the Darjeeling Himalayas, India. In *Integrating Disaster Science and Management: Global Case Studies in Mitigation and Recovery*; Elsevier: Amsterdam, The Netherlands, 2018; ISBN 9780128120576.
3. Gheshlaghi, H.A.; Feizizadeh, B. An integrated approach of analytical network process and fuzzy based spatial decision making systems applied to landslide risk mapping. *J. Afr. Earth Sci.* **2017**, *133*, 15–24. [[CrossRef](#)]
4. Zare, M.; Pourghasemi, H.R.; Vafakhah, M.; Pradhan, B. Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: A comparison between multilayer perceptron (MLP) and radial basic function (RBF) algo-rithms. *Arab. J. Geosci.* **2013**, *6*, 2873–2888. [[CrossRef](#)]
5. Glade, T.; Nadim, F. Early warning systems for natural hazards and risks. *Nat. Hazards* **2014**, *70*, 1669–1671. [[CrossRef](#)]
6. Basu, T.; Das, A.; Pal, S. Application of geographically weighted principal component analysis and fuzzy approach for un-supervised landslide susceptibility mapping on Gish River Basin, India. *Geocarto Int.* **2020**. [[CrossRef](#)]
7. Huang, F.; Yin, K.; Huang, J.; Gui, L.; Wang, P. Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine. *Eng. Geol.* **2017**, *223*, 11–22. [[CrossRef](#)]
8. He, Q.; Xu, Z.; Li, S.; Li, R.; Zhang, S.; Wang, N.; Pham, B.T.; Chen, W. Novel entropy and rotation forest-based credal-decision tree classifier for landslide susceptibility modeling. *Entropy* **2019**, *21*, 106. [[CrossRef](#)]
9. Nhu, V.H.; Shirzadi, A.; Shahabi, H.; Chen, W.; Clague, J.J.; Geertsema, M.; Jaafari, A.; Avand, M.; Miraki, S.; Asl, D.T.; et al. Shallow Landslide Susceptibility Mapping by Random Forest Base Classifier and Its Ensembles in a Semi-Arid Region of Iran. *Forests* **2020**, *11*, 421. [[CrossRef](#)]
10. Lee, D.-H.; Kim, Y.-T.; Lee, S. Shallow Landslide Susceptibility Models Based on Artificial Neural Networks Considering the Factor Selection Method and Various Non-Linear Activation Functions. *Remote Sens.* **2020**, *12*, 1194. [[CrossRef](#)]
11. Tien Bui, D.T.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Alizadeh, M.; Chen, W.; Mohammadi, A.; Bin Ahmad, B.; Panahi, M.; Hong, H.; et al. Landslide Detection and Susceptibility Mapping by AIRSAR Data Using Support Vector Machine and Index of Entropy Models in Cameron Highlands, Malaysia. *Remote Sens.* **2018**, *10*, 1527. [[CrossRef](#)]

12. Pandey, V.K.; Pourghasemi, H.R.; Sharma, M.C. Landslide susceptibility mapping using maximum entropy and support vector machine models along the highway corridor, Garhwal Himalaya. *Geocarto Int.* **2020**, *35*, 168–187. [[CrossRef](#)]
13. Yu, C.; Chen, J. Landslide Susceptibility Mapping Using the Slope Unit for Southeastern Helong City, Jilin Province, China: A Comparison of ANN and SVM. *Symmetry* **2020**, *12*, 1047. [[CrossRef](#)]
14. Sahin, E.K. Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Appl. Sci.* **2020**, *2*, 1–17. [[CrossRef](#)]
15. Achour, Y.; Pourghasemi, H.R. How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? *Geosci. Front.* **2020**, *11*, 871–883. [[CrossRef](#)]
16. Sahin, E.K.; Colkesen, I.; Kavzoglu, T. A comparative assessment of canonical correlation forest, random forest, rotation for-est and logistic regression methods for landslide susceptibility mapping. *Geocarto Int.* **2020**, *35*, 341–363. [[CrossRef](#)]
17. Wu, Y.; Ke, Y.; Chen, Z.; Liang, S.; Zhao, H.; Hong, H. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. *Catena* **2020**, *187*, 104396. [[CrossRef](#)]
18. Nhu, V.H.; Zandi, D.; Shahabi, H.; Chapi, K.; Shirzadi, A.; Al-Ansari, N.; Singh, S.K.; Dou, J.; Nguyen, H. Comparison of support vector machine, bayesian logistic regression, and alternating decision tree algorithms for shallow landslide suscepti-bility mapping along a mountainous road in the west of Iran. *Appl. Sci.* **2020**, *10*, 5047. [[CrossRef](#)]
19. Nhu, V.H.; Ngo, P.T.T.; Pham, T.D.; Dou, J.; Song, X.; Hoang, N.-D.; Tran, D.A.; Duong, P.C.; Aydilek, I.B.; Amiri, M.; et al. A New Hybrid Firefly-PSO Optimized Random Subspace Tree Intelligence for Torrential Rainfall-Induced Flash Flood Susceptible Mapping. *Remote Sens.* **2020**, *12*, 2688. [[CrossRef](#)]
20. Pham, B.T.; Van Phong, T.; Nguyen-Thoi, T.; Parial, K.; Singh, S.K.; Ly, H.B.; Nguyen, K.T.; Ho, L.S.; Van Le, H.; Prakash, I. Ensemble modeling of landslide susceptibility using random subspace learner and different decision tree classifiers. *Geocarto Int.* **2020**. [[CrossRef](#)]
21. Pham, B.T.; Prakash, I.; Singh, S.K.; Shirzadi, A.; Shahabi, H.; Bui, D.T. Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. *Catena* **2019**, *175*, 203–218. [[CrossRef](#)]
22. Park, S.; Hamm, S.Y.; Kim, J. Performance evaluation of the gis-based data-mining techniques decision tree, random forest, and rotation forest for landslide susceptibility modeling. *Sustainability* **2019**, *11*, 5659. [[CrossRef](#)]
23. Wu, B.; Qiu, W.; Jia, J.; Liu, N. Landslide Susceptibility Modeling Using Bagging-Based Positive-Unlabeled Learning. *IEEE Geosci. Remote Sens. Lett.* **2020**. [[CrossRef](#)]
24. Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.-W.; Han, Z.; Pham, B.T. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides* **2020**, *17*, 641–658. [[CrossRef](#)]
25. Pham, B.T.; Prakash, I. A novel hybrid model of Bagging-based Naïve Bayes Trees for landslide susceptibility assessment. *Bull. Eng. Geol. Environ.* **2019**, *78*, 1911–1925. [[CrossRef](#)]
26. Chen, W.; Chen, X.; Peng, J.; Panahi, M.; Lee, S. Landslide susceptibility modeling based on ANFIS with teach-ing-learning-based optimization and Satin bowerbird optimizer. *Geosci. Front.* **2020**, *12*, 93–107. [[CrossRef](#)]
27. Chen, W.; Hong, H.; Panahi, M.; Shahabi, H.; Wang, Y.; Shirzadi, A.; Pirasteh, S.; Alesheikh, A.A.; Khosravi, K.; Panahi, S.; et al. Spatial Prediction of Landslide Susceptibility Using GIS-Based Data Mining Techniques of ANFIS with Whale Optimization Algorithm (WOA) and Grey Wolf Optimizer (GWO). *Appl. Sci.* **2019**, *9*, 3755. [[CrossRef](#)]
28. Polykretis, C.; Chalkias, C.; Ferentinou, M. Adaptive neuro-fuzzy inference system (ANFIS) modeling for landslide suscepti-bility assessment in a Mediterranean hilly area. *Bull. Eng. Geol. Environ.* **2019**, *78*, 1173–1187. [[CrossRef](#)]
29. Wang, Y.; Hong, H.; Chen, W.; Li, S.; Panahi, M.; Khosravi, K.; Shirzadi, A.; Shahabi, H.; Panahi, S.; Costache, R. Flood sus-ceptibility mapping in Dingnan County (China) using adaptive neuro-fuzzy inference system with biogeography based opti-mization and imperialistic competitive algorithm. *J. Environ. Manag.* **2019**, *247*, 712–729. [[CrossRef](#)]
30. Kim, J.C.; Lee, S.; Jung, H.S.; Lee, S. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. *Geocarto Int.* **2018**, *33*, 1000–1015. [[CrossRef](#)]
31. Talukdar, S.; Ghose, B.; Salam, R.; Mahato, S.; Pham, Q.B.; Linh, N.T.T.; Costache, R.; Avand, M. Flood suscepti-bility modeling in Teesta River basin, Bangladesh using novel ensembles of bagging algorithms. *Stoch. Environ. Res. Risk Assess* **2020**, *34*, 2277–2300. [[CrossRef](#)]
32. Islam, A.R.M.T.; Talukdar, S.; Mahato, S.; Kundu, S.; Eibek, K.U.; Pham, Q.B.; Kuriqi, A.; Linh, N.T.T. Flood susceptibility modelling using advanced ensemble machine learning models. *Geosci. Front.* **2020**. [[CrossRef](#)]
33. Eshtay, M.; Faris, H.; Obeid, N. Improving Extreme Learning Machine by Competitive Swarm Optimization and its applica-tion for medical diagnosis problems. *Expert Syst. Appl.* **2018**, *104*, 134–152. [[CrossRef](#)]
34. Taormina, R.; Chau, K.-W. Neural network river forecasting with multi-objective fully informed particle swarm optimization. *J. Hydroinform.* **2015**, *17*, 99–113. [[CrossRef](#)]
35. Niu, W.J.; Feng, Z.K.; Cheng, C.T.; Wu, X.Y. A parallel multi-objective particle swarm optimization for cascade hy-dropower reservoir operation in southwest China. *Appl. Soft Comput. J.* **2018**, *70*, 562–575. [[CrossRef](#)]
36. Tomer, T.; Katal, D.; Joshi, V. Sensitivity analysis of groundwater vulnerability using DRASTIC method: A case study of National Capital Territory, Delhi, India. *Groundw. Sustain. Dev.* **2019**, *9*, 100271. [[CrossRef](#)]

37. Corominas, J.; VanWesten, C.; Frattini, P.; Cascini, L.; Malet, J.-P.; Fotopoulou, S.D.; Catani, F.; Eeckhaut, M.V.D.; Mavrouli, O.C.; Agliardi, F.; et al. Recommendations for the quantitative analysis of landslide risk. *Bull. Eng. Geol. Environ.* **2013**, *73*, 209–263. [[CrossRef](#)]
38. Abdulwahid, W.M.; Pradhan, B. Landslide vulnerability and risk assessment for multi-hazard scenarios using airborne laser scanning data (LiDAR). *Landslides* **2017**, *14*, 1057–1076. [[CrossRef](#)]
39. Liu, C.-N.; Dong, J.; Chen, C.-J.; Lee, W.-F. Typical landslides and related mechanisms in Ali Mountain highway induced by typhoon Morakot: Perspectives from engineering geology. *Landslides* **2012**, *9*, 239–254. [[CrossRef](#)]
40. Li, L.; Wang, Y. Identification of failure slip surfaces for landslide risk assessment using smoothed particle hydrodynamics. *Georisk* **2019**, *14*, 91–111. [[CrossRef](#)]
41. Dikshit, A.; Sarkar, R.; Pradhan, B.; Acharya, S.; Alamri, A.M. Spatial Landslide Risk Assessment at Phuentsholing, Bhutan. *Geosciences* **2020**, *10*, 131. [[CrossRef](#)]
42. Lu, M.; Zhang, J.; Zhang, L.; Zhang, L. Quantitative risk assessment of vehicles hit by landslides: A case study. *Nat. Hazards Earth Syst. Sci.* **2020**, *20*, e2020-11.
43. Guo, Z.; Chen, L.; Yin, K.; Shrestha, D.P.; Zhang, L. Quantitative risk assessment of slow-moving landslides from the view-point of decision-making: A case study of the Three Gorges Reservoir in China. *Eng. Geol.* **2020**, *273*, 105667. [[CrossRef](#)]
44. Zhu, C.; Zhang, J.; Liu, Y.; Ma, D.; Li, M.; Xiang, B. Comparison of GA-BP and PSO-BP neural network models with initial BP model for rainfall-induced landslides risk assessment in regional scale: A case study in Sichuan, China. *Nat. Hazards* **2020**, *100*, 173–204. [[CrossRef](#)]
45. Cardinali, M.; Reichenbach, P.; Guzzetti, F.; Ardizzone, F.; Antonini, G.; Galli, M.; Cacciano, M.; Castellani, M.; Salvati, P. A geomorphological approach to the estimation of landslide hazards and risks in Umbria, Central Italy. *Nat. Hazards Earth Syst. Sci.* **2002**, *2*, 57–72. [[CrossRef](#)]
46. Hadmoko, D.S.; Lavigne, F.; Sartohadi, J.; Hadi, P. Landslide hazard and risk assessment and their application in risk management and land use planning in eastern flank of Menoreh Mountains, Yogyakarta Province, Indonesia. *Nat. Hazards* **2010**, *54*, 623–642. [[CrossRef](#)]
47. Ghosh, S.; van Westen, C.J.; Carranza, E.J.M.; Jetten, V.G. Integrating spatial, temporal, and magnitude probabilities for medium-scale landslide risk analysis in Darjeeling Himalayas, India. *Landslides* **2012**, *9*, 371–384. [[CrossRef](#)]
48. Kanungo, D.; Arora, M.K.; Gupta, R.P.; Sarkar, S. Landslide risk assessment using concepts of danger pixels and fuzzy set theory in Darjeeling Himalayas. *Landslides* **2008**, *5*, 407–416. [[CrossRef](#)]
49. Tang, X.; Li, J.; Liu, M.; Liu, W.; Hong, H. Flood susceptibility assessment based on a novel random Naïve Bayes method: A comparison between different factor discretization methods. *Catena* **2020**, *190*, 104536. [[CrossRef](#)]
50. Pourghasemi, H.R.; Kornejady, A.; Kerle, N.; Shabani, F. Investigating the effects of different landslide positioning techniques, landslide partitioning approaches, and presence-absence balances on landslide susceptibility mapping. *Catena* **2020**, *187*, 104364. [[CrossRef](#)]
51. Wang, L.-J.; Sawada, K.; Moriguchi, S. Landslide susceptibility analysis with logistic regression model based on FCM sampling strategy. *Comput. Geosci.* **2013**, *57*, 81–92. [[CrossRef](#)]
52. Remondo, J.; Bonachea, J.; Cendrero, A. Quantitative landslide risk assessment and mapping on the basis of recent occurrences. *Geomorphology* **2008**, *94*, 496–507. [[CrossRef](#)]
53. Leonardi, G.; Palamara, R.; Suraci, F. A fuzzy methodology to evaluate the landslide risk in road lifelines. *Transp. Res. Procedia* **2020**, *45*, 732–739. [[CrossRef](#)]
54. Lee, S.; Min, K. Statistical analysis of landslide susceptibility at Yongin, Korea. *Environ. Geol.* **2001**, *40*, 1095–1113. [[CrossRef](#)]
55. Sidle, R.; Ochiai, H. Processes, prediction, and land use. In *Water Resources Monograph*; American Geophysical Union: Washington, DC, USA, 2006.
56. Mukherjee, I.; Singh, U.K. Delineation of groundwater potential zones in a drought-prone semi-arid region of east India using GIS and analytical hierarchical process techniques. *Catena* **2020**, *194*, 104681. [[CrossRef](#)]
57. Talukdar, S.; Pal, S.; Singha, P. Proposing artificial intelligence based livelihood vulnerability index in river islands. *J. Clean. Prod.* **2020**. [[CrossRef](#)]
58. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [[CrossRef](#)]
59. Faris, H.; Aljarah, I.; Al-Betar, M.A.; Mirjalili, S. Grey wolf optimizer: A review of recent variants and applications. *Neural Comput. Appl.* **2018**, *30*, 413–435. [[CrossRef](#)]
60. Devi, R.L.; Jayashree, L.S. Grey Wolf Optimization-Based Big Data Analytics for Dengue Outbreak Prediction. In *Advances in Intelligent Systems and Computing*; Springer: New York, NY, USA, 2018.
61. Wong, L.; Sulaiman, M.; Mohamed, M.; Hong, M. Grey Wolf Optimizer for solving economic dispatch problems. In Proceedings of the 2014 IEEE International Conference on Power and Energy (PECon), Kuching, Malaysia, 1–3 December 2014.
62. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948. [[CrossRef](#)]
63. Roshanravan, B.; Aghajani, H.; Yousefi, M.; Kreuzer, O. Particle Swarm Optimization Algorithm for Neuro-Fuzzy Prospection Analysis Using Continuously Weighted Spatial Exploration Data. *Nat. Resour. Res.* **2019**, *28*, 309–325. [[CrossRef](#)]
64. Gilani, S.-O.; Sattarvand, J.; Hajihassani, M.; Abdullah, S.S. A stochastic particle swarm based model for long term production planning of open pit mines considering the geological uncertainty. *Resour. Policy* **2020**, *68*, 101738. [[CrossRef](#)]

65. Yasin, Q.; Sohail, G.M.; Ding, Y.; Ismail, A.; Du, Q. Estimation of Petrophysical Parameters from Seismic Inversion by Combining Particle Swarm Optimization and Multilayer Linear Calculator. *Nat. Resour. Res.* **2020**, *29*, 3291–3317. [[CrossRef](#)]
66. Bui, Q.; Nguyen, Q.-H.; Nguyen, X.L.; Pham, V.; Nguyen, H.D.; Pham, V. Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. *J. Hydrol.* **2020**, *581*, 124379. [[CrossRef](#)]
67. Pourghasemi, H.R.; Razavi-Termeh, S.V.; Kariminejad, N.; Hong, H.; Chen, W. An assessment of metaheuristic approaches for flood assessment. *J. Hydrol.* **2020**, *582*, 124536. [[CrossRef](#)]
68. Mehrabi, M.; Pradhan, B.; Moayedi, H.; Alamri, A. Optimizing an adaptive neuro-fuzzy inference system for spatial prediction of landslide susceptibility using four state-of-the-art metaheuristic techniques. *Sensors* **2020**, *20*, 1723. [[CrossRef](#)]
69. Sun, D.; Wen, H.; Wang, D.; Xu, J. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology* **2020**, *362*, 107201. [[CrossRef](#)]
70. Gafar, M.G.; Elhoseny, M.; Gunasekaran, M. Modeling neutrosophic variables based on particle swarm optimization and information theory measures for forest fires. *J. Supercomput.* **2020**, *76*, 2339–2356. [[CrossRef](#)]
71. Zhang, H.; Liang, Z.; Liu, H.; Wang, R.; Liu, Y. Ensemble framework by using nature inspired algorithms for the early-stage forest fire rescue—A case study of dynamic optimization problems. *Eng. Appl. Artif. Intell.* **2020**, *90*. [[CrossRef](#)]
72. Xi, W.; Li, G.; Moayedi, H.; Nguyen, H. A particle-based optimization of artificial neural network for earthquake-induced landslide assessment in Ludian county, China. *Geomat. Nat. Hazards Risk* **2019**, *10*, 1750–1771. [[CrossRef](#)]
73. Nguyen, V.V.; Pham, B.T.; Vu, B.T.; Prakash, I.; Jha, S.; Shahabi, H.; Shirzadi, A.; Ba, D.N.; Kumar, R.; Chatterjee, J.M.; et al. Hybrid Machine Learning Approaches for Landslide Susceptibility Modeling. *Forests* **2019**, *10*, 157. [[CrossRef](#)]
74. Pradhan, B.; Lee, S. Landslide susceptibility assessment and factor effect analysis: Backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environ. Model. Softw.* **2010**, *25*, 747–759. [[CrossRef](#)]
75. Friedel, M.J. Modeling hydrologic and geomorphic hazards across post-fire landscapes using a self-organizing map approach. *Environ. Model. Softw.* **2011**, *26*, 1660–1674. [[CrossRef](#)]
76. Arnone, E.; Francipane, A.; Noto, L.V.; Scarbaci, A.; La Loggia, G. Strategies investigation in using artificial neural network for landslide susceptibility mapping: Application to a Sicilian catchment. *J. Hydroinform.* **2014**, *16*, 502–515. [[CrossRef](#)]
77. Genuer, R.; Poggi, J.M.; Tuleau-Malot, C. Variable selection using random forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236. [[CrossRef](#)]
78. Beven, K.J.; Kirkby, M.J. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69. [[CrossRef](#)]
79. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
80. Stumpf, A.; Kerle, N. Object-oriented mapping of landslides using Random Forests. *Remote Sens. Environ.* **2011**, *115*, 2564–2577. [[CrossRef](#)]
81. Abedini, M.; Ghasemian, B.; Shirzadi, A.; Bui, D.T. A comparative study of support vector machine and logistic model tree classifiers for shallow landslide susceptibility modeling. *Environ. Earth Sci.* **2019**, *78*, 560. [[CrossRef](#)]
82. Pal, S.; Talukdar, S.; Ghosh, R. Damming effect on habitat quality of riparian corridor. *Ecol. Indic.* **2020**, *114*, 106300. [[CrossRef](#)]
83. Singha, P.; Das, P.; Talukdar, S.; Pal, S. Modeling livelihood vulnerability in erosion and flooding induced river island in Ganges riparian corridor, India. *Ecol. Indic.* **2020**, *119*, 106825. [[CrossRef](#)]
84. Talukdar, S.; Pal, S. Modeling flood plain wetland transformation in consequences of flow alteration in Punarbhabariver in India and Bangladesh. *J. Clean. Prod.* **2020**, *261*, 120767. [[CrossRef](#)]
85. Baeza, C.; Lantada, N.; Amorim, S. Statistical and spatial analysis of landslide susceptibility maps with different classification systems. *Environ. Earth Sci.* **2016**, *75*, 1318. [[CrossRef](#)]
86. Li, Z.; Li, Z.; Zhao, W.; Wang, Y. Probability modeling of precipitation extremes over two river basins in northwest of China. *Adv. Meteorol.* **2015**, *2015*, 374127. [[CrossRef](#)]
87. Oh, H.J.; Pradhan, B. Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput. Geosci.* **2011**, *37*, 1264–1276. [[CrossRef](#)]
88. Moayedi, H.; Nguyen, H.; Foong, L.K. Nonlinear evolutionary swarm intelligence of grasshopper optimization algorithm and gray wolf optimization for weight adjustment of neural network. *Eng. Comput.* **2019**, 1–11. [[CrossRef](#)]
89. Jaafari, A.; Panahi, M.; Pham, B.T.; Shahabi, H.; Bui, D.T.; Rezaie, F.; Lee, S. Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility. *Catena* **2019**, *175*, 430–445. [[CrossRef](#)]
90. Li, D.; Huang, F.; Yan, L.; Cao, Z.; Chen, J.; Ye, Z. Landslide susceptibility prediction using particle-swarm-optimized multi-layer perceptron: Comparisons with multilayer-perceptron-only, BP neural network, and information value models. *Appl. Sci.* **2019**, *9*, 3664. [[CrossRef](#)]
91. Moayedi, H.; Mehrabi, M.; Mosallanezhad, M.; Rashid, A.S.A.; Pradhan, B. Modification of landslide susceptibility mapping using optimized PSO-ANN technique. *Eng. Comput.* **2019**, *35*, 967–984. [[CrossRef](#)]
92. Pham, B.T.; Qi, C.; Ho, L.S.; Nguyen-Thoi, T.; Al-Ansari, N.; Nguyen, M.D.; Nguyen, H.D.; Ly, H.-B.; Van Le, H.; Prakash, I. A Novel Hybrid Soft Computing Model Using Random Forest and Particle Swarm Optimization for Estimation of Undrained Shear Strength of Soil. *Sustainability* **2020**, *12*, 2218. [[CrossRef](#)]

93. Pal, S.; Paul, S. Assessing wetland habitat vulnerability in moribund Ganges delta using bivariate models and machine learning algorithms. *Ecol. Indic.* **2020**, *119*, 106866. [[CrossRef](#)]
94. Ilia, I.; Tsangaratos, P. Applying weight of evidence method and sensitivity analysis to produce a landslide susceptibility map. *Landslides* **2016**, *13*, 379–397. [[CrossRef](#)]
95. Svalova, V.B.; Zaalishvili, V.B.; Ganapathy, G.P.; Nikolaev, A.V.; Melkov, D.A. Landslide risk in mountain areas. *Геология и геофизика Юга России* **2019**, *9*, 109–126.
96. Terzi, S.; Torresan, S.; Schneiderbauer, S.; Critto, A.; Zebisch, M.; Marcomini, A. Multi-risk assessment in mountain regions: A review of modelling approaches for climate change adaptation. *J. Environ. Manag.* **2019**, *232*, 759–771. [[CrossRef](#)] [[PubMed](#)]
97. Maes, J.; Mertens, K.; Jacobs, L.; Bwambale, B.; Vranken, L.; Dewitte, O.; Poesen, J.; Kervyn, M. Social multi-criteria evaluation to identify appropriate disaster risk reduction measures: Application to landslides in the Rwenzori Mountains, Uganda. *Landslides* **2019**, *16*, 1793–1807. [[CrossRef](#)]
98. Han, L.; Ma, Q.; Zhang, F.; Zhang, Y.; Zhang, J.; Bao, Y.; Zhao, J. Risk assessment of an earthquake-collapse-landslide disaster chain by Bayesian network and Newmark models. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3330. [[CrossRef](#)] [[PubMed](#)]
99. Qiu, H.; Cui, Y.; Hu, S.; Yang, D.; Pei, Y.; Yang, W. Temporal and spatial distributions of landslides in the Qinba Mountains, Shaanxi Province, China. *Geomat. Nat. Hazards Risk* **2019**, *10*, 599–621. [[CrossRef](#)]
100. Marin, R.J.; Mattos, Á.J. Physically-based landslide susceptibility analysis using Monte Carlo simulation in a tropical mountain basin. *Georisk* **2020**, *14*, 192–205.
101. Mineo, S. Comparing rockfall hazard and risk assessment procedures along roads for different planning purposes. *J. Mt. Sci.* **2020**, *17*, 653–669. [[CrossRef](#)]
102. Mavrouli, O.; Corominas, J.; Ibarbia, I.; Alonso, N.; Jugo, I.; Ruiz, J.; Luzuriaga, S.; Navarro, J.A. Integrated risk assessment due to slope instabilities in the roadway network of Gipuzkoa, Basque Country. *Nat. Hazards Earth Syst. Sci.* **2019**, *19*, 399–419. [[CrossRef](#)]
103. Winter, M.G.; Wong, J.C.F. The assessment of quantitative risk to road users from debris flow. *Geoenviron. Disasters* **2020**, *7*, 4. [[CrossRef](#)]
104. Althuwaynee, O.F.; Pradhan, B. Semi-quantitative landslide risk assessment using GIS-based exposure analysis in Kuala Lumpur City. *Geomat. Nat. Hazards Risk* **2017**, *8*, 706–732. [[CrossRef](#)]
105. Dikshit, A.; Sarkar, R.; Pradhan, B.; Jena, R.; Drukpa, D.; Alamri, A.M. Temporal probability assessment and its use in landslide susceptibility mapping for Eastern Bhutan. *Water* **2020**, *12*, 267. [[CrossRef](#)]