International Journal of Geo-Information

MDPI

Article

Random Forest-Based Landslide Susceptibility Mapping in Coastal Regions of Artvin, Turkey

Halil Akinci [1,*], Cem Kilicoglu [2] and Sedat Dogan [3]

1 Department of Geomatics Engineering, Artvin Çoruh University, 08100 Artvin, Turkey
2 Kavak Vocational School, Samsun University, Kavak, 55850 Samsun, Turkey; cem.kilicoglu@samsun.edu.tr
3 Department of Geomatics Engineering, Ondokuz Mayis University, 55139 Samsun, Turkey; sedatdo@omu.edu.tr
* Correspondence: halil.akinci@artvin.edu.tr; Tel.: +90-(466)-215-10-00 (ext. 4626)

Received: 8 August 2020; Accepted: 13 September 2020; Published: 15 September 2020

Abstract: Natural disasters such as landslides often occur in the Eastern Black Sea region of Turkey owing to its geological, topographical, and climatic characteristics. Landslide events occur nearly every year in the Arhavi, Hopa, and Kemalpaşa districts located on the Black Sea coast in the Artvin province. In this study, the landslide susceptibility map of the Arhavi, Hopa, and Kemalpaşa districts was produced using the random forest (RF) model, which is widely used in the literature and yields more accurate results compared with other machine learning techniques. A total of 10 landslide-conditioning factors were considered for the susceptibility analysis, i.e., lithology, land cover, slope, aspect, elevation, curvature, topographic wetness index, and distances from faults, drainage networks, and roads. Furthermore, 70% of the landslides on the landslide inventory map were used for training, and the remaining 30% were used for validation. The RF-based model was validated using the area under the receiver operating characteristic (ROC) curve. Evaluation results indicated that the success and prediction rates of the model were 98.3% and 97.7%, respectively. Moreover, it was determined that incorrect land-use decisions, such as transforming forest areas into tea and hazelnut cultivation areas, induce the occurrence of landslides.

Keywords: landslides; landslide susceptibility; machine learning; random forest; Artvin

1. Introduction

Landslide is a natural disaster that causes economic damage and fatalities, both in Turkey and many different regions in the world. In Turkey, the region most affected by landslides in terms of frequency of occurrence and resultant damage is the Eastern Black Sea region [1–3]. The main factors that predispose landslide triggering in the Eastern Black Sea region are rugged terrain structure, climatic conditions with heavy rainfall, weathering, deforestation, dense tea and hazelnut cultivation, dispersed rural settlements, low-quality transportation network owing to dispersed settlement, and uncontrolled surface drainage systems created by the transportation network [4].

One of the essential procedures to reduce damage and fatalities caused by landslides is landslide susceptibility mapping (LSM) [5–9]. Trigila et al. [10] reported that LSM is the first step in estimating landslide hazard and risk. Landslide susceptibility maps show the regions where future landslides can occur in the study area [11]. Generally, LSM involves the following steps: collecting spatial data regarding the factors affecting the occurrence of landslides in the selected study area, determining landslide susceptibility using the association between existing landslides on the inventory map and conditioning factors, and validating the results [12]. One of the important steps in this process is the determination of the method or model to be used for producing the landslide susceptibility map.

Zhao et al. [5] reported that employing modeling approaches has a significant impact on the quality of landslide susceptibility maps.

The models used for LSM can be broadly divided into three categories: physical, heuristic, and statistical models. Physically-based models use mechanical laws to analyze the slope stability in terms of physical processes leading to landslide events. The main advantages of these models are that they do not need long-term landslide inventory data and are more useful in areas where landslide inventories are incomplete [13]. The heuristic approach is a qualitative analysis based on expert knowledge and previous work experience [14]. In the heuristic approach, investigators rank and weigh conditioning factors according to their assumed or expected importance in causing landslides [15]. Statistical approaches are based on the analysis of functional relationships between known or predicted instability factors and past and present distribution of landslides [15]. These methods require landslide inventory maps to create functional relationships with conditioning factors [14]. Reichenbach et al. [15] further classified the statistic-based models into six groups, such as "classical statistics, index-based, machine learning, neural networks, multi-criteria decision analysis, and other statistics." In particular, logistic regression [16–23] and the frequency ratio method [24–30] have been the most widely used classical statistical models around the world. In recent years, there has been a growing interest in machine learning techniques in LSM studies. Machine learning techniques, such as artificial neural networks (ANNs) [31,32], Naïve Bayes [33,34], support vector machines (SVMs) [35–37], decision trees (DTs) [38,39], and random forest (RF) [40–42], have been widely used by researchers. Sevgen et al. [8] stated that the application of random forest (RF) method in LSM is relatively new; thus, there is limited research on this method. Merghadi et al. [43] reported that tree-based ensemble algorithms yield excellent results compared with other machine learning algorithms and that the RF model offers a powerful performance for accurate LSM. Similarly, numerous studies in the literature have reported that the RF model provides more accurate results compared with other machine learning models [8,40,44,45]. On the other hand, deep learning has become the new trend in LSM with developments in neural networks techniques. Deep learning algorithms are defined as a specialized version of machine learning algorithms [37]. Deep learning is a self-teaching system that learns as it goes by filtering information through multiple hidden layers, in a similar way to humans [46]. Deep learning techniques, such as convolutional neural network (CNN) and recurrent neural network (RNN), have been implemented recently in LSM studies [47–51].

The study region covers the Arhavi, Hopa, and Kemalpaşa districts located on the Black Sea coast in the Artvin province in the Eastern Black Sea region in Turkey. The present research aims to produce a landslide susceptibility map of this study region using the RF model. This study region was selected because Artvin is one of the cities in Turkey in which natural disasters most commonly occur and landslides account for the majority of natural disasters [2,52,53]. Natural disasters, particularly those of meteorological nature, such as floods and landslides, induced by excessive rainfall, are frequently observed throughout the province [54,55]. Although all districts of Artvin are affected by landslides to a certain extent, the districts where landslides most commonly occur are Arhavi, Hopa, and Kemalpaşa. On August 25, 2015, floods and landslides occurred in the Arhavi, Hopa (and Kemalpaşa), and Borçka districts of Artvin owing to heavy rainfall, resulting in 11 deaths (of which six died because of landslide) and 19 injuries. The factors considered in this study include lithology, land cover, slope, aspect, elevation, curvature, topographic wetness index (TWI), distances from faults, drainage networks, and roads.

2. Characteristics of the Study Area

This study was conducted in a region with a total surface area of 621 km^2, covering the Arhavi, Hopa, and Kemalpaşa districts in the Artvin province. The study area is located between the northern latitudes of 41° 05′ 4.99″–41° 31′ 12.93″ and the eastern longitudes of 41° 14′ 39.29″–41° 37′ 14.98″ (Figure 1). The elevation in the study area ranges from 0 to 3370 m and the slope from 0° to 76.95°.

The average elevation is 1098.18 m, and the average slope is 28.38°. The slope is in the range of 0°–5° for 2.8% of the study area, 5°–10° for 2.35%, 10°–20° for 15.26%, and more than 20° for 79.58%.

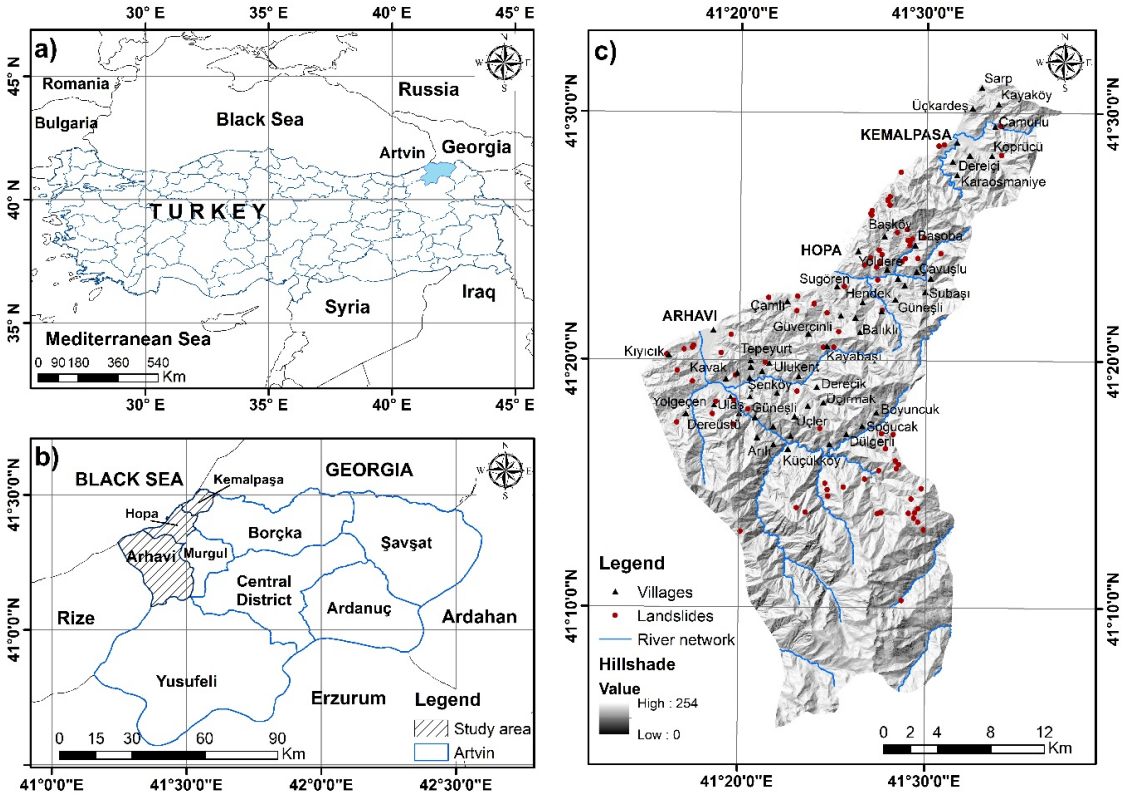

Figure 1. Study area maps (**a**) location map of the study area in Turkey, (**b**) location map of the Artvin region, (**c**) location map showing settlements and landslides in the study area.

According to the CORINE 2018 land cover data, 62.74% of the study area is covered with forest, 16.08% with agricultural lands, and 8.15% with natural grasslands. The main agricultural products include tea and hazelnut. Tea and hazelnut cultivation is conducted in 70% of agricultural lands in the Arhavi district and 53% of agricultural lands in the Hopa and Kemalpaşa districts.

According to the Turkish Statistical Institute's 2019 population data, Arhavi's population is 20,926, Hopa's population is 26,958, and Kemalpaşa's population is 9224 [56]. The total population of the study area is 57,108.

According to the data of the General Directorate of Meteorology (2009–2020), the total monthly precipitation in the study area is 2215.16 mm. In this region, the minimum average monthly precipitation of 94.25 mm is observed in May and the maximum average monthly precipitation of 345.2 mm is observed in September. According to the data of the General Directorate of Meteorology (2000–2020), the maximum average monthly temperature of 27.71 °C is observed in August and the average monthly minimum temperature of 3.63 °C is observed in January in the study area.

The study area has a rather steep and rough topography. In the region, there are sharp hillsides from coastal areas to interiors. Deep valleys stand out all across the study area. In accordance with the morphological structure of the region, there are many streams with regular and irregular flow regimes, forming a dendritic drainage network. The streams have a high base of slope, which increases the flow rate, causing more material to be transported [57,58]. Because of the morphology of the region, the materials carried by streams are deposited along with the flat areas near the Black Sea. Furthermore, it is seen that the settlement areas are also established in these wide and flat morphologic units.

In addition to this, the geologic formation in the region is formed from rocks belonging to various geological times. The study area is located in the Northern Zone of East Pontides within the Pontides

Tectonic Assembly. This section of the Eastern Pontides is separated into two zones. In the north, the Adjara-Trialeti belt coming from the east continues along with the Black Sea coast. On the other hand, the Somcheti-Kafan (Karabakh) belt, constituting the south section of Transcaucasia, corresponds to the Pondites Tectonic Assembly. The study area is rich in volcanic, volcano-sedimentary and intrusive rocks containing magmatism products that have developed in periods in the Early Jurassic and Late Cretaceous Eocene age range. During periods when the magmatism was not active, sedimentary stacks were accumulated [59].

3. Materials and Methods

3.1. Data Used

The first stage of landslide susceptibility assessment is data collection and creation of a spatial database in which thematic maps of the factors associated with landslides are produced [60]. In this study, 1:100,000 scale digital geology and 1:25,000 scale digital landslide inventory maps were obtained from the General Directorate of Mining Research and Exploration (GDMRE), 1:25,000 scale topographic maps from the General Directorate of Mapping, 1:100,000 scale CORINE 2018 land cover data from the European Union Copernicus Land Monitoring Service (https://land.copernicus.eu/pan-european/corine-land-cover), and 1:25,000 scale digital road network data from the Artvin Regional Directorate of Forestry. Thematic maps of the factors associated with landslides were prepared using ESRI ArcGIS 10.5 software and converted in raster format (ESRI GRID) with a spatial resolution of 10 m. The modeling operations were conducted in R 3.6.3 software.

3.2. Landslide Inventory

It is necessary to use an accurate and complete landslide inventory for the training and validation of landslide models [10]. Landslide inventory maps providing basic information to assess landslide hazards on a regional scale are the most crucial data required to examine the relationship between landslide occurrence and factors affecting it [61]. However, in Yanar et al. [62], it is clearly expressed that dense buildings and constructions change the topography, and these areas dramatically prevent the visibility of the landslides in the urban areas. Therefore, in the residential areas, it is very difficult to prepare landslide inventory maps. On the other side, Dağ and Bulut [58] stated that in the tea cultivation areas, such as our study area, the vegetation covers the relatively small-scale landslides (mainly flow debris and rotational slides); thus, making it very difficult to produce the inventory maps in these tea covered areas. Furthermore, the 1:25,000 scaled regional landslide inventory maps cannot represent the small-scale landslides as polygons since the small-scale landslide areas are too small to be represented on this scale. In this study, 1:25,000 scale landslide inventory maps produced by the GDMRE and updated by the Artvin Provincial Directorate of Disaster and Emergency were used. This inventory map, which we used in this paper, is the most accurate and the most confident landslide inventory map for this region. Landslide inventory maps used in Turkey were prepared within the scope of the "Turkey Landslide Inventory Project," which was initiated by GDMRE in 1997 on a national scale and completed in 2009. The landslides detected based on the field studies and aerial images (1:10,000–1:35,000 scale) were processed on 1:25,000 scale basic topographic maps [63].

The landslides were classified into three types based on their movements: flowing, sliding, and complex. Fall and tipping type landslides were not considered, owing to scale limitations. Furthermore, the landslides were classified into two groups based on their relative depths. Accordingly, landslides with sliding surface of less than 5-m depth were classified as shallow and those with sliding surface greater than 5-m depth were classified as deep. The landslides were further classified by dividing them into two groups based on their activities: active and inactive [63,64].

According to the inventory map, there are a total of 85 mapped landslide polygons in the study area. The landslides in the study area typically fall into the shallow landslide group. The landslides cover 0.94% of the study area. Moreover, 30 of these landslides belong to the inactive landslide class

and 55 to the active (45 sliding and 11 flowing types) landslide class. The smallest and largest landslides on the inventory map have areas of 0.002 and 0.81 km^2, respectively. The average area of the landslide polygons is 0.07 km^2. The widths of the landslides in the study region range from 45 to 981 m, while their lengths range from 60 to 1947 m.

However, models used to produce landslide susceptibility maps require both "landslide (or positive)" and "non-landslide (or negative)" samples of the study area. Although there is no accepted rule in the creation of these two subsets, most researchers use a ratio of 70:30 in LSM studies, particularly in the selection of "landslide" samples. In this approach, 70% of the randomly selected landslides on the inventory map are used for model training and the remaining 30% are used for model validation [31,32,44,65]. Huang and Zhao [36] emphasized that the positive and negative samples in the training and test datasets should be the same, i.e., a ratio of 1:1. Thus, researchers use the same number of negative and positive samples in their studies [39,66–69]. The landslide inventory map was converted to a raster format with a cell size of 10 m × 10 m, and the total number of positive samples in the study area, i.e., the number of pixels with landslides, was determined to be 58,537. The same number of "non-landslide" pixels was randomly selected in the study area. The positive samples were provided a value of "1," and the negative samples were provided a value of "0." Moreover, 70% of the dataset comprising a total of 117,074 pixels was used for model training and the remaining 30% was used for model validation.

3.3. Landslide-Conditioning Factors

One of the important stages of LSM is the determination of topographical, geological, environmental, and anthropogenic factors affecting the occurrence of landslides in the study area [70]. Vakhshoori et al. [71] reported that there is no standard rule on the selection of the conditioning factors, and the scale of the study area, geoenvironmental conditions, landslide occurrence mechanism in the study area, and data availability influence the selection of factors. Considering these criteria, a total of 10 landslide-conditioning factors, such as lithology, land cover, slope, aspect, elevation, curvature, TWI, and distances from faults, drainage networks, and roads, were used in the study. The factor maps prepared in ESRI ArcGIS 10.5 were converted to raster format with a spatial resolution of 10 m.

In this study, a digital elevation model (DEM) with a resolution of 10 m × 10 m was derived from 1:25,000 scale topographic maps with a 10-m contour interval obtained from the General Directorate of Mapping. The elevation, slope angle, aspect, curvature, TWI, and distance from drainage network parameters were derived from this DEM.

The elevation is often used in LSM studies [23,34,39,72]. The elevation of the study ranges from 0 to 3370 m (Figure 2a). The elevation is divided into 10 equal parts at 337-m intervals (Table 1).

As reported by Zhang et al. [72], the slope angle is a key factor for slope stability and is one of the most commonly used parameters in LSM studies [73–75]. The average slope of the study area, where the slope ranges from 0° to 76.95°, is 28.38°. The slope is divided into 10 classes at 5° intervals (Figure 2b).

The aspect at a point on the land surface is the direction that faces the tangent plane passing through that point and is expressed in angles (in degrees). Similar to elevation and slope, aspect is also widely used in LSM studies [42,45,48]. Bragagnolo et al. [74] reported that aspect has an indirect effect on slope stability depending on rainfall, wind, and solar radiation. The aspect map of the study area was derived from DEM and divided into nine classes (Figure 2c).

Curvature, one of the factors associated with the occurrence of landslides, represents slope shape and terrain morphology [22,76]. Negative curvatures represent concave surfaces, zero curvatures represent flat surfaces, and positive curvatures represent convex surfaces. In this study, the curvature was derived from DEM using ArcGIS 10.5 software and divided into three subclasses, i.e., concave, flat, and convex (Figure 2d).

TWI is defined as a theoretical measure of flow accumulation at any point in a basin, and hence, soil wetness [74]. In other words, TWI measures the degree of water accumulation in one region [77].

In the study area, the TWI ranges from 2.12 to 27.92 (Table 1). TWI is divided into nine subclasses using the natural breaks classification method and used in landslide susceptibility analysis (Figure 3a).

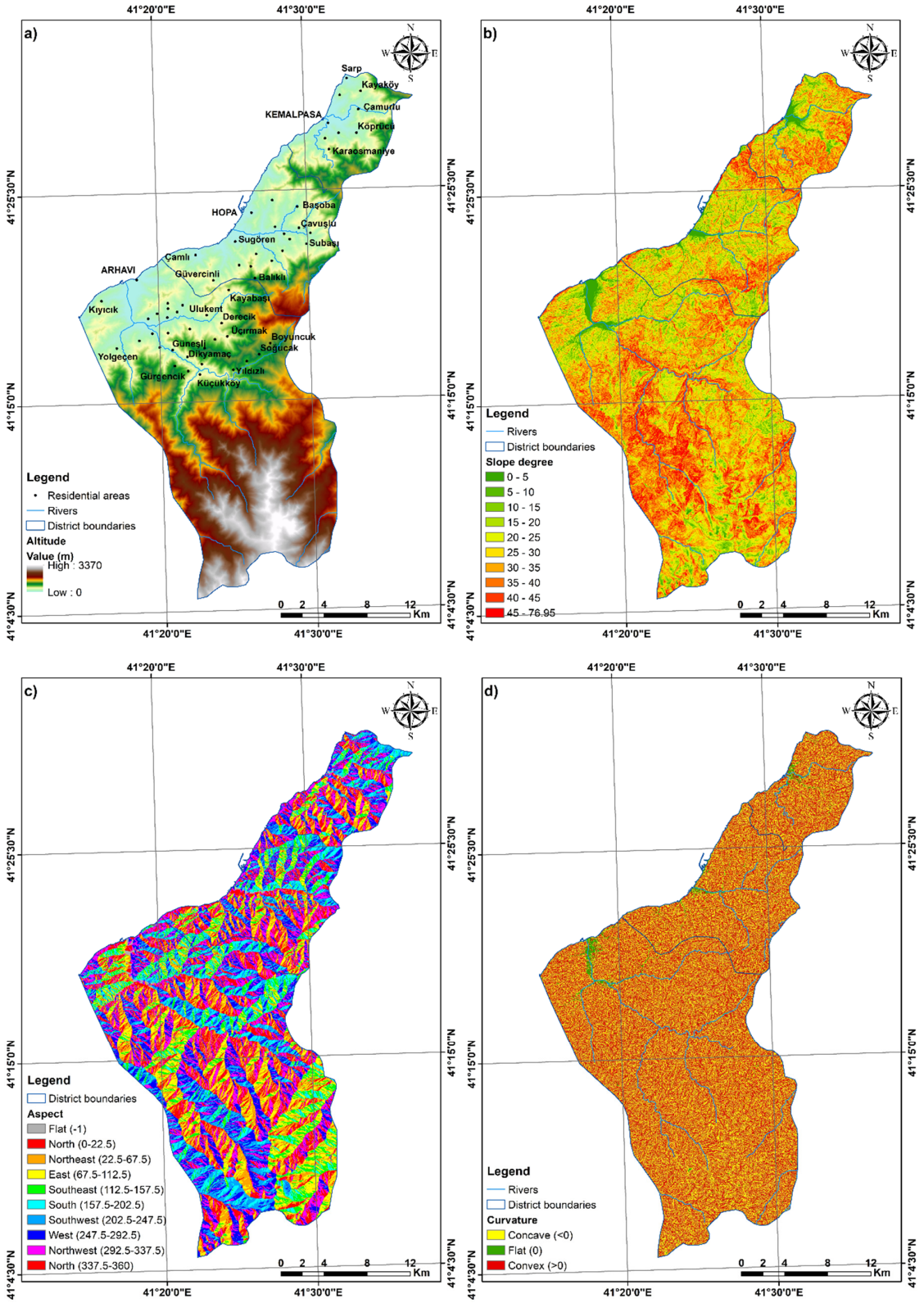

Figure 2. Topographic factor maps (**a**) elevation; (**b**) slope; (**c**) aspect; and (**d**) curvature.

Table 1. Spatial relationship between landslide-conditioning factors and landslides.

Factors	Subclasses	Pixels in Domain	Pixels with Landslide	Percentage of Landslides (%)	Percentage of Domain (%)	Frequency Ratio (FR)
Elevation (m)	0–337	1,443,794	23,096	39.46	23.24	1.6976
	337–674	1,150,034	8308	14.19	18.51	0.7666
	674–1011	810,518	1167	1.99	13.05	0.1528
	1011–1348	611,505	6893	11.78	9.84	1.1962
	1348–1685	591,042	12,484	21.33	9.51	2.2415
	1685–2022	442,710	6314	10.79	7.13	1.5135
	2022–2359	381,321	275	0.47	6.14	0.0765
	2359–2696	433,915	0	0.00	6.99	0.0000
	2696–3033	299,255	0	0.00	4.82	0.0000
	3033–3370	47,952	0	0.00	0.77	0.0000
Slope (Degree)	0–5	173,997	1078	1.84	2.80	0.6575
	5–10	146,037	5516	9.42	2.35	4.0083
	10–15	344,395	12,692	21.68	5.54	3.9109
	15–20	603,868	13,550	23.15	9.72	2.3812
	20–25	876,332	10,711	18.30	14.11	1.2971
	25–30	1,111,771	7335	12.53	17.90	0.7001
	30–35	1,201,986	4382	7.49	19.35	0.3869
	35–40	965,692	1915	3.27	15.55	0.2104
	40–45	515,204	625	1.07	8.29	0.1287
	45–76.95	272,764	733	1.25	4.39	0.2852
Aspect	Flat	57,225	226	0.39	0.92	0.4191
	North	988,912	16,769	28.65	15.92	1.7995
	Northeast	792,420	9126	15.59	12.76	1.2222
	East	684,258	5022	8.58	11.02	0.7789
	Southeast	541,325	3041	5.20	8.71	0.5962
	South	523,134	1971	3.37	8.42	0.3998
	Southwest	658,060	2553	4.36	10.59	0.4117
	West	927,276	5926	10.12	14.93	0.6782
	Northwest	1,039,436	13,903	23.75	16.73	1.4194
TWI	2.12–4.85	1,059,034	4897	8.37	17.05	0.4907
	4.85–5.96	1,907,594	14,839	25.35	30.71	0.8255
	5.96–6.97	1,688,106	17,125	29.26	27.17	1.0766
	6.97–8.09	864,051	11,582	19.79	13.91	1.4225
	8.09–9.50	365,213	5841	9.98	5.88	1.6972
	9.50–11.33	177,124	2930	5.01	2.85	1.7555
	11.33–13.65	99,210	1170	2.00	1.60	1.2515
	13.65–16.99	38,150	113	0.19	0.61	0.3143
	16.99–27.92	13,564	40	0.07	0.22	0.3130
Curvature	<0	2,962,010	30,380	51.90	47.68	1.0884
	0	124,224	1049	1.79	2.00	0.8961
	>0	3,125,812	27,108	46.31	50.32	0.9203
Lithology	Qal	144,679	613	1.05	2.33	0.4496
	Tet	188,759	0	0.00	3.04	0.0000
	Tk2	48,058	0	0.00	0.77	0.0000
	Tkd	22,815	0	0.00	0.37	0.0000
	Tek	1,097,641	13,738	23.47	17.67	1.3282
	Tekba	3375	58	0.10	0.05	1.8237
	Tee	65,515	0	0.00	1.05	0.0000
	Tpeb	52,499	1056	1.80	0.85	2.1346
	Kk1	195,127	0	0.00	3.14	0.0000
	Kk1gd	31,231	0	0.00	0.50	0.0000
	Kk1kd	225,767	152	0.26	3.63	0.0714
	Kk1a	455,881	0	0.00	7.34	0.0000
	KTc	569,041	8753	14.95	9.16	1.6324
	KTct3	270,181	2830	4.83	4.35	1.1116
	KTck	108,527	1089	1.86	1.75	1.0649

Table 1. *Cont.*

Factors	Subclasses	Pixels in Domain	Pixels with Landslide	Percentage of Landslides (%)	Percentage of Domain (%)	Frequency Ratio (FR)
	KTct2	16,917	45	0.08	0.27	0.2823
	Kçbh	89,658	412	0.70	1.44	0.4877
	Kça	1,914,832	26,387	45.08	30.82	1.4624
	Kçt	994	0	0.00	0.02	0.0000
	Kçat1	90,683	1068	1.82	1.46	1.2498
	Kk	349,483	2184	3.73	5.63	0.6632
	Kkdt	7494	0	0.00	0.12	0.0000
	Kç	250,947	152	0.26	4.04	0.0643
	Jh	10,743	0	0.00	0.17	0.0000
Distance from drainage network (m)	0–100	3,451,459	38,738	66.18	55.56	1.1911
	100–200	1,947,340	15,236	26.03	31.35	0.8303
	200–300	718,748	4212	7.20	11.57	0.6219
	300–400	86,479	302	0.52	1.39	0.3706
	400–500	7071	48	0.08	0.11	0.7204
	500–600	412	1	0.00	0.01	0.2576
	600–700	187	0	0.00	0.00	0.0000
	700–800	209	0	0.00	0.00	0.0000
	800–926.55	141	0	0.00	0.00	0.0000
Distance from roads (m)	0–200	4,005,718	40,707	69.54	64.48	1.0784
	200–400	1,234,833	11,524	19.69	19.88	0.9904
	400–600	480,588	4477	7.65	7.74	0.9886
	600–800	232,012	1566	2.68	3.73	0.7163
	800–1000	116,327	263	0.45	1.87	0.2399
	1000–1200	70,106	0	0.00	1.13	0.0000
	1200–1400	35,956	0	0.00	0.58	0.0000
	1400–1600	15,301	0	0.00	0.25	0.0000
	1600–1800	7652	0	0.00	0.12	0.0000
	1800–2868.48	13,553	0	0.00	0.22	0.0000
CORINE 2018	112	37,568	0	0.00	0.60	0.0000
	121	5670	73	0.12	0.09	1.3663
	122	42,234	1056	1.80	0.68	2.6534
	123	4738	0	0.00	0.08	0.0000
	131	21,588	872	1.49	0.35	4.2866
	133	2768	0	0.00	0.04	0.0000
	222	821,867	14,706	25.12	13.23	1.8989
	242	17,710	0	0.00	0.29	0.0000
	243	159,487	3789	6.47	2.57	2.5212
	311	1,143,634	13,626	23.28	18.41	1.2644
	312	45,538	0	0.00	0.73	0.0000
	313	2,708,333	22,087	37.73	43.60	0.8654
	321	506,223	2328	3.98	8.15	0.4880
	324	81,469	0	0.00	1.31	0.0000
	332	101,735	0	0.00	1.64	0.0000
	333	496,238	0	0.00	7.99	0.0000
	511	15,246	0	0.00	0.25	0.0000
Distance from faults (m)	0–1000	1,663,388	27,557	47.08	26.78	1.7581
	1000–2000	1,376,852	17,711	30.26	22.16	1.3651
	2000–3000	961,994	2938	5.02	15.49	0.3241
	3000–4000	634,006	6345	10.84	10.21	1.0620
	4000–5000	423,053	3323	5.68	6.81	0.8336
	5000–6000	354,957	361	0.62	5.71	0.1079
	6000–7000	289,889	302	0.52	4.67	0.1106
	7000–8000	207,933	0	0.00	3.35	0.0000
	8000–9000	150,900	0	0.00	2.43	0.0000
	9000–13102.31	149,074	0	0.00	2.40	0.0000

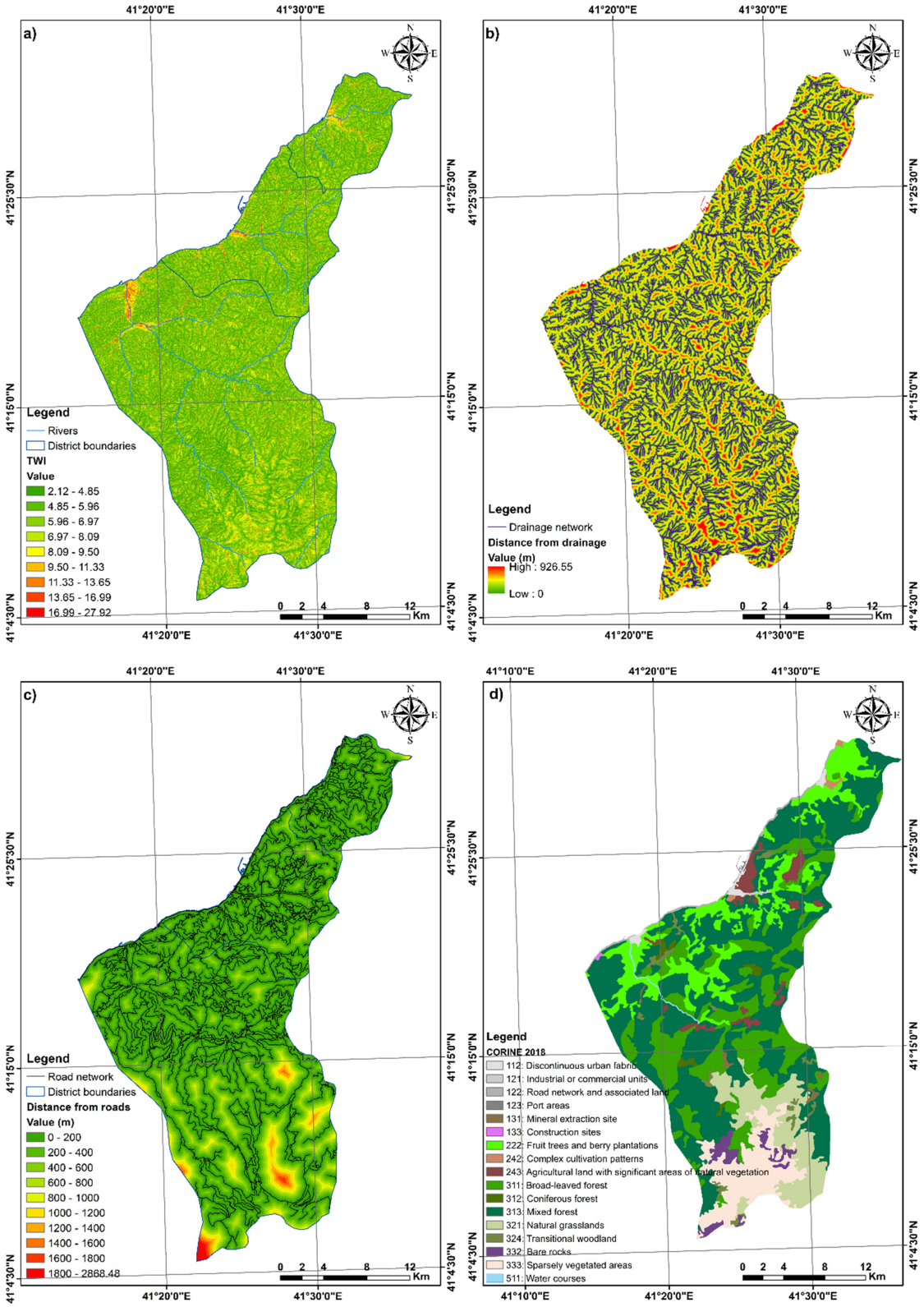

Figure 3. Landslide-conditioning factor maps: (**a**) topographic wetness index (TWI); (**b**) distance from drainage; (**c**) distance from roads; (**d**) land cover.

The proximity of the slopes to a drainage network is another important factor in terms of stability. This is because rivers disrupt the stability of slopes by eroding the toe of the slopes or saturating the section of the material constituting the slope under the stream level [78]. The drainage network of the

study area was derived from DEM using the hydrological analysis module in ArcGIS 10.5 software. Drainage buffers were calculated using the Euclidean distance function in ArcGIS 10.5 and reclassified at 100-m intervals and divided into nine subclasses (Table 1). The maximum distance to the drainage network was calculated to be 926.55 m (Figure 3b).

In regions with rugged topography, landslides occur on the slopes near the road. Road constructions, particularly on hillsides with a high slope, may reduce the load on the slope toe, form stress cracks behind the slope, and lead to the occurrence of landslides owing to the disrupted slope structure [21]. Therefore, in this study, the distance from roads was considered as a conditioning factor. The road network data of the study area were digitally provided by the Forest Information System of Artvin Regional Directorate of Forestry. This 1:25,000 scale dataset includes highways, village roads, and forest roads. In ArcGIS 10.5 software, the road buffer map was produced using the Euclidean distance function. The road buffers were divided into 10 categories at 200-m intervals to obtain a map of the proximity from roads (Figure 3c).

Land cover is an indirect indicator of slope stability. Areas covered with barren and sparse vegetation show more instability than forests. Moderately sloping agricultural lands are susceptible to landslides because of repetitive irrigation processes [79]. In this study, CORINE land cover (CLC 2018) data, which were provided by Copernicus Land Monitoring Service (CLMS), a European Union Earth Observation Programme service, updated in 2018, were used. According to this dataset, the study area includes 17 different classes of land cover (Figure 3d). Moreover, 62.74% of the study area is covered with forests, 16.08% with agricultural lands, 8.15% with natural grasslands, and 7.99% with sparsely vegetated areas (Table 1).

One of the most important parameters affecting the stability of the slopes is lithology. This is because lithological and structural variations often cause a difference in the strength and permeability of rocks and soils [80]. The lithological units and faults used in this study were obtained from the 1:100,000 scaled geology map obtained from the GDMRE [59]. According to the geological map, there are 24 different lithological units in the study area (Figure 4). The Liyas-aged Hamurkesen formation (Jh), which is considered the oldest unit in the study area, consists of basaltic, andesitic, dacitic lava and pyroclastics and sandstone, marn, limestone, and shale. Among volcanic units, Late Cretaceous-aged (Turonian-Conasian) Çatak formation (Kç) consists of basaltic, andesitic lava and pyroclastics, and intermediate levels of clay limestone, marn, sandstone, siltstone, and claystone. The stack of dacitic lava-tuffs in the study area is discriminated by the symbol "Kkdt." The Santonian-aged Kızılkaya formation (Kk) consists of dacitic, rhyodacitic lava and pyroclastics. The Campanian-Maastrichtian aged Çağlayan formation (Kça), the main lithological unit in the study area, comprises basaltic and andesitic lava and pyroclastics along with mudstone, sandstone, argillaceous limestone, marn, and tuff interlayers. In the study area, the red limestones with sand-clay-Globotruncana, which are observed in the areas where the Çağlayan formation outcrops, is shown by the symbol "Kçt," and the stack of clay tuff, tuff sandstone, sandy limestone and breccia is shown by the symbol "Kçat1." In the unit that outcrops in the study area and is called Hematite dacite (Kçbh), the dacite is brown, since it contains hematite. The unit was effectively clayed. These units include the Maastrichtian-Danian (Early Paleocene) aged Cankurtaran formation (KTc) consisting of sandy limestone, micritic limestone, tuff, marn, and volcanic sandstone. The intercalation of tuff, marn, and fine-grained agglomerate is discriminated by the symbol KTct2; the grey and red limestone is discriminated by the symbol KTck; and the intercalation of tuff, marn, limestone, volcanic sandstone, and agglomerate is discriminated by the symbol KTct3. Accordingly, volcanic intercalations are seen to be present within the unit. The Bakırköy formation (Tpeb), consisting of siltstone, claystone, sandstone, conglomerate, clayey limestone, and marn, is Paleocene-Early Eocene aged [59].

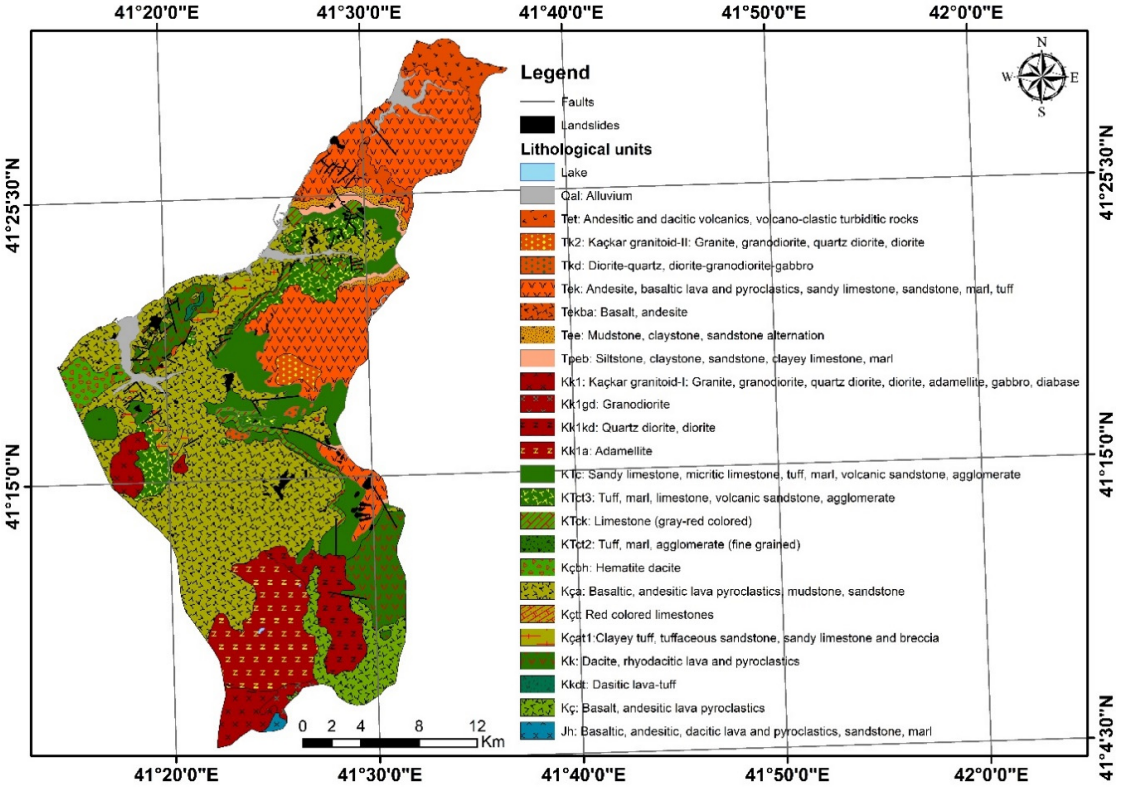

Figure 4. Geological map of the study area [59,81].

The Cankurtaran and Bakırköy formations were cut by Kaçkar granitoid-I (Kk1) consisting of acidic and basic intrusive rocks that have continued to develop in the Late Cretaceous period and completed their intrusion at the end of Paleocene. In the study area, the superficial adamellite is discriminated by (Kk1a), granodiorite by (Kk1gd), and quartz diorite, diorite by (Kk1kd). In the study area, the unit consisting of the Middle Eocene aged (lower level of the Middle Eocene) is called the Erenler formation (Tee), which consists of the mudstone, claystone, and sandstone intercalation. The volcanics, which are discriminated as basalt and andesite and rest conformably on this unit, are shown by "Tekba" (Figure 4). In this unit, there is the Kabaköy formation (Tek), a volcano-sedimentary stack consisting of conglomerate, sandy limestone, sandstone, marn, and tuff, and Middle Eocene aged andesitic, basaltic lava and pyroclastics. The first of the two rocks that cut this formation consists of Diorite-Quartz Diorite-Granodiorite-Gabbro, which is a black and dark green color, with a compact appearance, and is not fractured, is shown by the symbol "Tkd." The other is a blackish-dark green rock, which consists of diorite-quartz diorite-granodiorite-gabbro, and shown as Kaçkar Granitoid-II (Tk2). Dacitic and andesitic volcanites and volcano-clastic Taşpınar formation (Tet) rest conformably on the Eocene aged rocks. The Quaternary-aged alluvial sediments (Qal) are the youngest units in the study area [59].

Regmi et al. [82], Pourghasemi et al. [22], and Wang et al. [70] reported that landslides mainly occur along faults. In fact, areas close to the faults are highly susceptible to landslides because the rock around the faults is severely broken and the strength of the rock decreases owing to tectonic breaks [34]. The fault buffers were also produced using the Euclidean distance function. The fault buffers were reclassified by dividing them into 10 classes at equal intervals of 1000 m to obtain a map of the distance from the fault lines (Figure 5).

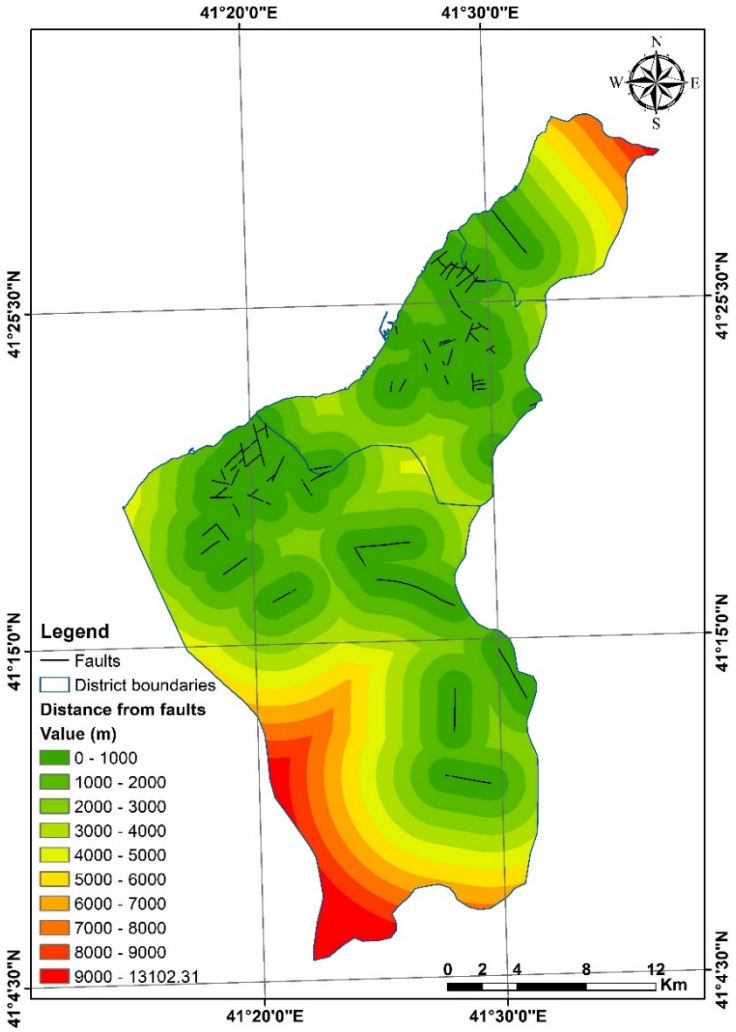

Figure 5. Map of the distance from the fault lines.

3.4. Landslide Susceptibility Assessment Using Random Forest (RF) Model

RF classification, which was originally developed by Breiman [83], is a machine learning algorithm for nonparametric multivariate classification [84]. RF is a popular ensemble learning method that is widely used for classification and regression. Generally, a single DT individually exhibits weak prediction performance because of a high variance or bias [64,85]. Therefore, RF creates numerous DTs for classification [86]. RF can also be perceived as a group of random DTs. Because the performance of an ensemble model is better than that of individual models, individually created DTs are combined to form a decision forest, i.e., an RF. Each tree in the forest has independent and identical distribution (iid), and thus, they are relatively uncorrelated with each other. This property makes the RF, which is formed by the iid DTs, far from the overtraining risk [83]. Here, the DTs are a subset that is randomly selected from the corresponding dataset. The results obtained from all DTs are combined to obtain the result of the RF.

To establish a classification model in RF, two parameters should be defined: *ntree* parameter, corresponding to the number of DTs produced by RF, and *mtry* parameter, referring to the number of factors or variables used in the node of each DT. Taalab et al. [85] reported that there is no rule regarding the number of trees that should be established in RF, and increasing the number of trees will not increase the accuracy of the model. However, Chen et al. [38] emphasized that the number of variables to be used in DTs should equal the square root of the number of causal factors. RF algorithm was explained in detail by Breiman [83], Catani et al. [84], and Taalab et al. [85]. In this study, the

"caret" package [87] is used in R 3.6.3 to employ the RF model. Herein, the number of trees (*ntree*) is set to 50 and the *mtry* parameter is set to 8 after several attempts. RF is performed using a 10-fold cross-validation approach to reduce the variability of the model results and to limit overfitting.

4. Results and Discussions

4.1. Landslide Susceptibility Mapping Results

Figure 6 shows the landslide susceptibility map produced using the RF model in the R program and divided into five classes (very low, low, moderate, high, and very high) using the natural breaks classification method.

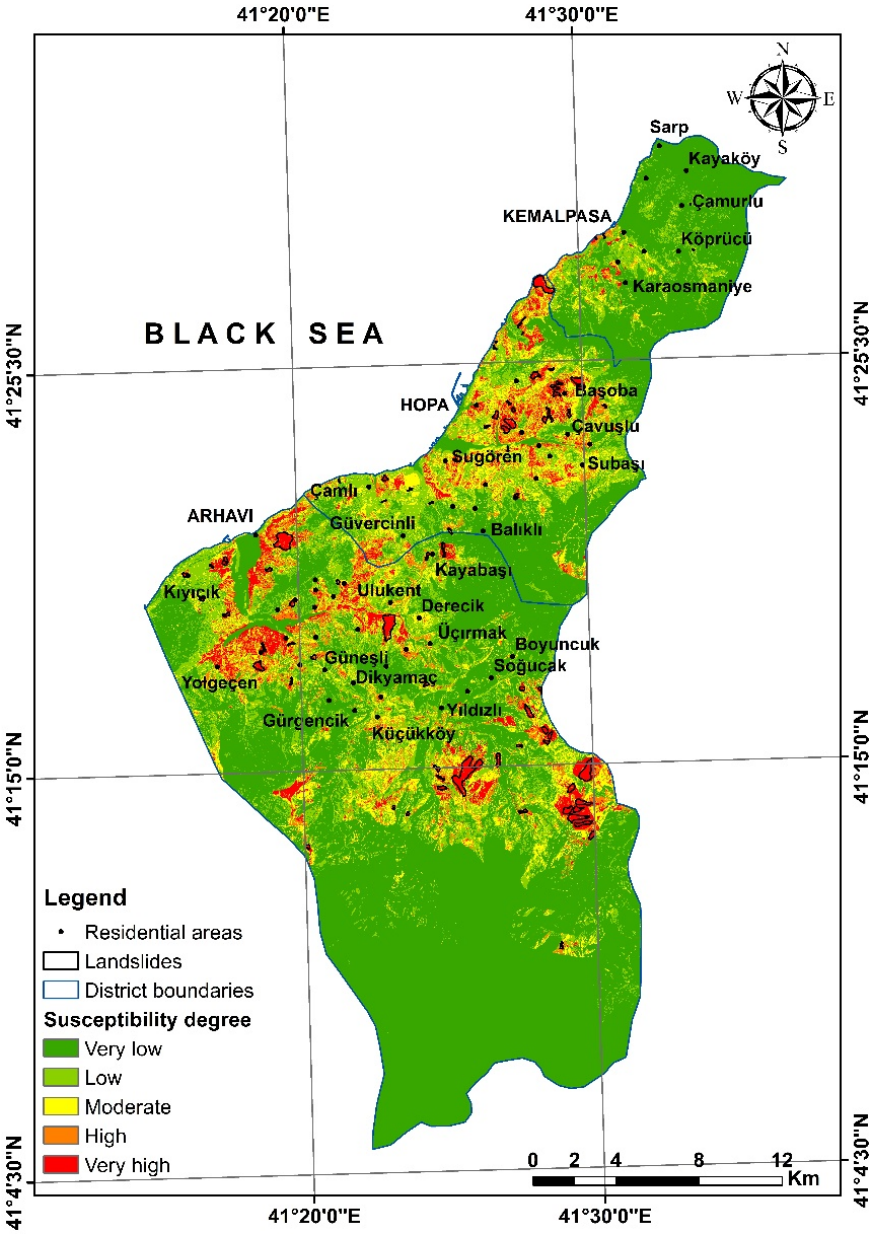

Figure 6. Landslide susceptibility map produced using the random forest model.

According to this map, 61.25%, 17.70%, 10.46%, 6.05%, and 4.54% of the study area exhibit a very low, low, moderate, high, and very high susceptibility to landslide, respectively. Although only 10.59%

of the study area is highly and very highly susceptible to landslides, approximately 96% of the existing landslides were in these two classes (Table 2).

Table 2. Percentages of different landslide susceptibility classes and existing landslide pixels.

Class	Number of Pixels	Relative Area (%)	Number of Landslide Pixels	Relative Area (%)
Very low	3,804,888	61.25	33	0.06
Low	1,099,613	17.70	476	0.81
Moderate	649,705	10.46	2000	3.42
High	375,520	6.05	7481	12.78
Very high	282,316	4.54	48,547	82.93

The importance levels of the factors used in the study are shown in Figure 7. According to Figure 7, slope, elevation, lithology, distance from faults, and land cover are the most important factors in the landslide susceptibility analysis in the study area, while TWI and curvature are the least important factors. Lin et al. [88] and Pourghasemi and Rahmati [33] reported that the relative importance of conditioning factors depends on the characteristics of the study area.

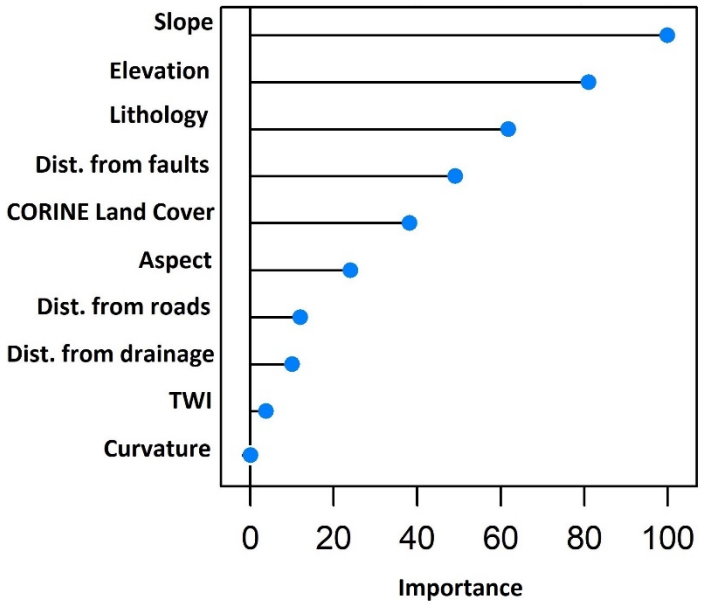

Figure 7. Importance of landslide conditioning factors.

The relative importance levels of the factors as well as the probability of landslides in each subclass of each factor were evaluated using the frequency ratio (FR) values provided in Table 1. Table 1 shows the relationship between the subclasses of conditioning factors in the study area and the occurrence of landslides. Similar to the relative importance values, the slope factor was one of the factors with the greatest FR value. In terms of the slope factor, 5°–10° (FR = 4.0083), 10°–15° (FR = 3.9109), 15°–20° (FR = 2.3812), and 20°–25° (FR = 1.2971) slope classes exhibited the greatest FR values and the probability of occurrence of landslides decreased as the slope increased above 25° (Table 1). Wang and Li [89], Vakhshoori et al. [71], and Nohani et al. [90] reported that the probability of the occurrence of landslides increased to some extent with increasing slope and then decreased.

Kavzoglu et al. [80] reported that the effect of elevation on landslide susceptibility is a controversial topic and has not yet been clarified by researchers. However, as shown in Figure 7, the elevation was the second most important factor in the landslide susceptibility analysis in the study area. In terms of elevation, the elevation classes in the range of 1348–1685 m (FR = 2.2415), 0–337 m (FR = 1.6976),

1685–2022 m (FR = 1.5135), and 1011–1348 m (FR = 1.1962) were found to be susceptible to landslide (Table 1). The results of this research revealed that the elevation class of 0–337 m was particularly susceptible to landslide because the residential and tea cultivation areas in the study site are located within this elevation class. The study conducted by Çan and Duman [91] reported that the landslides induced by heavy rainfall in Hopa district in Artvin province on August 24, 2015 showed distributions along the coastline, and 80% of the landslides occurred in zones with an elevation of less than 350 m.

Approximately 83% of the landslides in the study area occurred in three lithological units: Çağlayan formation (Kça), Kabaköy formation (Tek), and Cankurtaran formation (KTc). These three units cover approximately 58% of the study area (Table 1). Gökçe et al. [52] reported that the Cretaceous and Eocene volcanisms in the Eastern Black Sea formed the source areas for the occurrence of landslides, and the landslides were primarily flowing type in areas where the Eocene and Cretaceous volcanic rocks formed as a result of the Pontid volcanism in the Eastern Black Sea.

Approximately 77% of landslides in the study area occurred at a distance of 2000 m from faults. Areas near the fault lines have high FR values, while the probability of the occurrence of landslides decreases with increasing distance from faults (Table 1). Similarly, Althuwaynee et al. [92] stated that the probability of landslide occurrence decreases as the distance from faults increases. In terms of proximity to faults, Chen et al. [68] and Wang et al. [70] found that the <1000-m class exhibited the greatest probability of the occurrence of a landslide, which decreased as the distance from faults increased.

When Table 1 is examined, it can be seen that 31.6% of landslides in the study area occurred in agricultural areas, while 61.01% occurred in forested areas. In terms of land cover, the classes with 131 (mineral extraction sites), 122 (road and rail networks and associated land), 243 (land principally occupied by agriculture, with significant areas of natural vegetation), 222 (fruit trees and berry plantations), 121 (industrial or commercial units), and 311 (broad-leaved forest) CORINE 2018 land cover codes exhibited the greatest probability of the occurrence of a landslide. In a study in which the landslide susceptibility of Arhavi and its surroundings was evaluated by Aksoy [2], it was determined that the landslides in this region generally occurred in tea plantation areas where the slope was relatively low and the decomposition depth was not considerably high. The study by Bulut et al. [93] investigated the causes of landslides occurring in the eastern parts of the Fındıklı district in the Rize province (which is located in the west of the Arhavi district and has similar topographical, climatic, and environmental factors as Arhavi), and reported that the landslides mostly occurred at hillsides with a slope ranging from 10° to 25°. The study also stated that the fact that 77% of landslides occurred in residential areas and tea cultivation areas was important owing to the effect of altered vegetation on the occurrence of a landslide. In a study conducted by Erener et al. [64] in the Şavşat district in the Artvin province in Turkey, the landslide activity was reported to increase in areas where the original vegetation was removed or altered.

Conversely, approximately 68% of the current landslides were found to have occurred at slopes with north, northwest, and northeast aspects. The slopes with these three aspects were found to be at a higher risk of landslides (Table 1). It is assumed that the general rainfall direction of the region and the general morphological structure of the study site play a significant role in the fact that landslides mostly occurred in the north aspect. In a study conducted by Dağ and Bulut [58] in the Çayeli district in the Rize province in Turkey, most landslides in the study area occurred on the slopes with a north–northeast aspect. In the study by Bahrami et al. [94] in Sarv-Abad (Iran), the frequency rates of landslides with north, northeast, and northwest aspects were found to be higher than other directions.

Approximately 66% of landslides in the study area occurred at 100 m distance from drainage networks, and 70% of them at 200 m distance from roads. As distances from drainage networks and roads increase, the likelihood of landslides decreased (Table 1). These results have shown that uncontrolled road excavations in the region and carving of slope heels in areas close to drainage networks have been effective on landslides. Similarly, Yılmaz [95], in his study on the general characteristics and causes of landslides occurring in the Eastern Black Sea Region, found that two of

the main factors affecting landslides in the region were associated with excavations made by locals for road and home constructions as well as erosion along the stream side.

The curvature results show that 51.90% of landslides occur on concave slopes and 46.31% on convex slopes (Table 1). Concave areas are seen to have the highest FR (1.0884) value. It was determined that based on TWI, the landslide occurrence probability has gradually increased up to the 9.50–11.33 subclass, reaching the maximum value in the 9.50–11.33 subclass (FR = 1.7555), which then gradually decreased (Table 1).

4.2. Model Performance and Validation

Validation is an important step in assessing the quality of the landslide susceptibility model. The quality or performance of the RF model used in the study was evaluated using the success and prediction rate curves based on the receiver operating characteristic (ROC)/area under the curve (AUC). Seventy percent of the dataset, consisting of landslide (positive) and non-landslide (negative) pixels, was used for model training, and the remaining 30% was used for model validation. The success rate curve was produced using examples from the training dataset (Figure 8a). Pourghasemi et al. [12] noted that the success rate curve may help to determine how well landslide susceptibility maps classify existing landslide areas. In this study, the AUC value for the success rate curve was found to be 98.3%. However, several researchers have stated that there is no established method to evaluate the prediction capability of the model as the success rate curve is generated from the training dataset [12,66,92]. To explain how well landslide models and landslide-conditioning factors predict landslides, the prediction rate curve produced using the verification dataset should be used [12,66,92,96,97]. The AUC value for the prediction rate curve generated using the verification dataset was found to be 97.7% (Figure 8b).

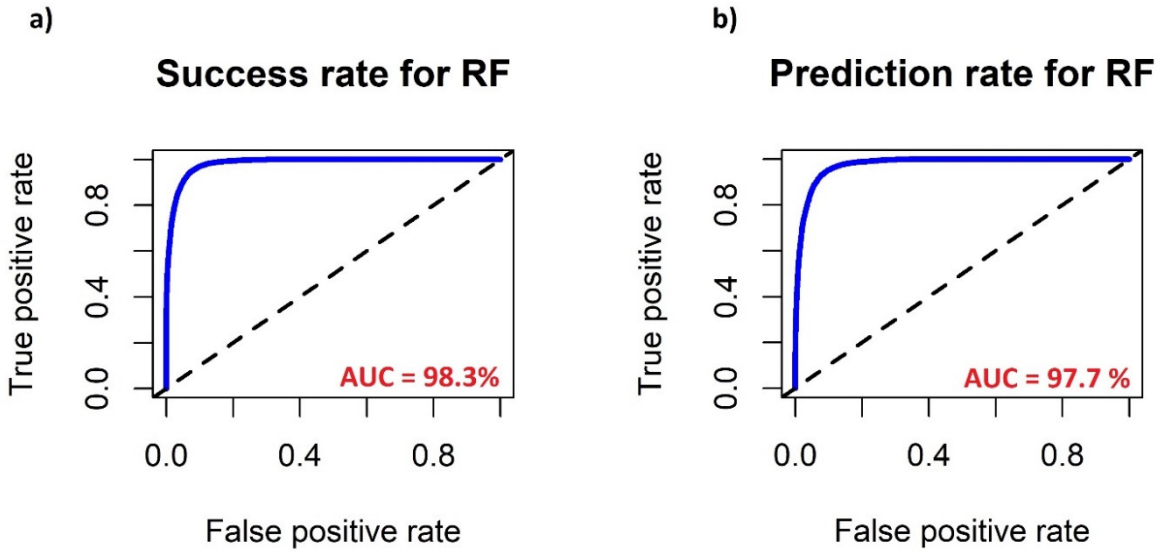

Figure 8. (**a**) Success and (**b**) prediction rate curves for random forest.

Lin et al. [88] reported that the AUC values greater than 0.7 for the ROC curves are generally considered good, whereas values greater than 0.9 indicate excellent model compatibility. According to the success and prediction rate values, the model used in this study showed excellent performance in predicting landslide susceptibility in the study area. In the LSM study conducted by Dou et al. [40], using DT and RF models in Izu-Oshima Volcanic Island of Japan, the authors found that the RF model with 95.6% AUC exhibited better performance than the DT model. In the landslide susceptibility study conducted by Achour and Pourghasemi [45] using the RF, SVM, and boosted regression tree models, researchers found that the RF model with 97.2% AUC exhibited the highest prediction accuracy compared with other models.

5. Conclusions

In this study, LSM of the Arhavi, Hopa, and Kemalpaşa districts in the Artvin province was prepared using the RF machine learning technique and various factors, such as lithology, land cover, slope, aspect, elevation, curvature, TWI, and distances from faults, drainage networks, and roads. The landslide susceptibility map was divided into five susceptibility classes: very low, low, moderate, high, and very high. The results show that approximately 96% of the landslides were observed in high and very high susceptibility classes. The landslide susceptibility map was validated using the success and prediction rate curves. The area below the prediction rate curve was found to be 97.7%. This value reveals the accuracy of the generated landslide susceptibility map. It was found that the most important factors in the occurrence of landslides in the study area were slope, elevation, lithology, distance from faults, and land cover. The majority of the tea and hazelnut cultivation areas in the region were obtained by transforming forests into agricultural lands. Therefore, incorrect land-use decisions, such as the removal of natural vegetation at hillsides with a high slope, transformation of these areas into tea and hazelnut cultivation areas, and building and road construction works performed by tea and hazelnut producers at these regions, lay the foundation for landslides, which are also triggered by heavy rainfall. Consequently, necessary measures should be taken to prevent landslide in the tea and hazelnut agricultural areas, and the establishment of new agricultural areas should not be allowed. Furthermore, in areas susceptible to landslides, forestation should be performed with deep-rooted forest trees and any human activities creating any additional burden on hillsides with a high slope should not be allowed.

Author Contributions: Conceptualization, Halil Akinci; methodology, Halil Akinci and Sedat Dogan; validation, Halil Akinci and Cem Kilicoglu; resources, Halil Akinci and Cem Kilicoglu; writing—original draft preparation, Halil Akinci; writing—review and editing, Halil Akinci and Sedat Dogan; visualization, Halil Akinci. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Scientific Research Projects Office of Artvin Çoruh University (AÇÜBAP) (Scientific Research Project No. 2016.F40.02.05).

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

References

1. Akgun, A.; Dag, S.; Bulut, F. Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood-frequency ratio and weighted linear combination models. *Environ. Geol.* **2008**, *54*, 1127–1143. [CrossRef]
2. Aksoy, G. Landslide Susceptibility Analysis of Arhavi Region (Artvin). Master's Thesis, Karadeniz echnical University, Trabzon, Turkey, 2011. (In Turkish).
3. Yalcin, A.; Reis, S.; Aydinoglu, A.C.; Yomralioglu, T. A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. *Catena* **2011**, *85*, 274–287. [CrossRef]
4. Çan, T.; Bulut, İ.; Tekin, S.; Özcan, A.K. Landslide hazard maps of the watersheds along the coastline between Gülyali (Ordu)—Bulancak (Giresun). In Proceedings of the 70th Geological Congress of Turkey, Ankara, Turkey, 10–14 April 2017; pp. 184–185.
5. Zhao, Y.; Wang, R.; Jiang, Y.; Liu, H.; Wei, Z. GIS-based logistic regression for rainfall-induced landslide susceptibility mapping under different grid sizes in Yueqing, Southeastern China. *Eng. Geol.* **2019**, *259*, 105147. [CrossRef]
6. Carabella, C.; Miccadei, E.; Paglia, G.; Sciarra, N. Post-Wildfire Landslide Hazard Assessment: The Case of the 2017 Montagna Del Morrone Fire (Central Apennines, Italy). *Geosciences* **2019**, *9*, 175. [CrossRef]
7. Marsala, V.; Galli, A.; Paglia, G.; Miccadei, E. Landslide Susceptibility Assessment of Mauritius Island (Indian Ocean). *Geosciences* **2019**, *9*, 493. [CrossRef]
8. Sevgen, E.; Kocaman, S.; Nefeslioglu, H.A.; Gokceoglu, C. A Novel Performance Assessment Approach Using Photogrammetric Techniques for Landslide Susceptibility Mapping with Logistic Regression, ANN and Random Forest. *Sensors* **2019**, *19*, 3940. [CrossRef]

9. Karakas, G.; Can, R.; Kocaman, S.; Nefeslioglu, H.A.; Gokceoglu, C. Landslide Susceptibility Mapping with Random Forest Model for Ordu, Turkey. Int. Arch. Photogramm. *Remote. Sens. Spat. Inf. Sci.* **2020**, *43*, 1229–1236.
10. Trigila, A.; Iadanza, C.; Esposito, C.; Scarascia-Mugnozza, G. Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). *Geomorphology* **2015**, *249*, 119–136. [CrossRef]
11. Vakhshoori, V.; Zare, M. Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods. *Geomat. Nat. Hazards Risk* **2016**, *7*, 1731–1752. [CrossRef]
12. Pourghasemi, H.R.; Mohammady, M.; Pradhan, B. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. *Catena* **2012**, *97*, 71–84. [CrossRef]
13. Luo, W.; Liu, C.C. Innovative landslide susceptibility mapping supported by geomorphon and geographical detector methods. *Landslides* **2018**, *15*, 465–474. [CrossRef]
14. Wang, G.; Chen, X.; Chen, W. Spatial Prediction of Landslide Susceptibility Based on GIS and Discriminant Functions. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 144. [CrossRef]
15. Reichenbach, P.; Rossi, M.; Malamud, B.D.; Mihir, M.; Guzzetti, F. A review of statistically-based landslide susceptibility models. *Earth-Sci. Rev.* **2018**, *180*, 60–91. [CrossRef]
16. Yilmaz, I. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat—Turkey). *Comput. Geosci.* **2009**, *35*, 1125–1138. [CrossRef]
17. Bai, S.-B.; Wang, J.; Lü, G.-N.; Zhou, P.-G.; Hou, S.-S.; Xu, S.-N. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. *Geomorphology* **2010**, *115*, 23–31. [CrossRef]
18. 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]
19. Bai, S.; Lu, P.; Wang, J. Landslide susceptibility assessment of the Youfang catchment using logistic regression. *J. Mt. Sci* **2015**, *12*, 816–827. [CrossRef]
20. Chen, T.; Niu, R.; Jia, X. A comparison of information value and logistic regression models in landslide susceptibility mapping by using GIS. *Environ. Earth Sci.* **2016**, *75*, 867. [CrossRef]
21. Du, G.-L.; Zhang, Y.-S.; Iqbal, J.; Yang, Z.-H.; Yao, X. Landslide susceptibility mapping using an integrated model of information value method and logistic regression in the Bailongjiang watershed, Gansu Province, China. *J. Mt. Sci.* **2017**, *14*, 249–268. [CrossRef]
22. Pourghasemi, H.R.; Gayen, A.; Park, S.; Lee, C.-W.; Lee, S. Assessment of Landslide-Prone Areas and Their Zonation Using Logistic Regression, LogitBoost, and NaïveBayes Machine-Learning Algorithms. *Sustainability* **2018**, *10*, 3697. [CrossRef]
23. Shou, K.-J.; Lin, J.-F. Evaluation of the extreme rainfall predictions and their impact on landslide susceptibility in a sub-catchment scale. *Eng. Geol.* **2020**, *265*, 105434. [CrossRef]
24. Lee, S.; Pradhan, B. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* **2007**, *4*, 33–41. [CrossRef]
25. Ozdemir, A.; Altural, T. A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *J. Asian Earth Sci.* **2013**, *64*, 180–197. [CrossRef]
26. Shahabi, H.; Khezri, S.; Bin Ahmad, B.; Hashim, M. Landslide susceptibility mapping at central Zab basin, Iran: A comparison between analytical hierarchy process, frequency ratio and logistic regression models. *Catena* **2014**, *115*, 55–70. [CrossRef]
27. Tasoglu, I.K.; Keskin Citiroglu, H.; Mekik, C. GIS-based landslide susceptibility assessment: A case study in Kelemen Valley (Yenice—Karabuk, NW Turkey). *Environ. Earth Sci.* **2016**, *75*, 1291. [CrossRef]
28. Ding, Q.; Chen, W.; Hong, H. Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. *Geocarto Int.* **2017**, *32*, 619–639. [CrossRef]
29. Hong, H.; Chen, W.; Xu, C.; Youssef, A.M.; Pradhan, B.; Bui, D.T. Rainfall-induced landslide susceptibility assessment at the Chongren area (China) using frequency ratio, certainty factor, and index of entropy. *Geocarto Int.* **2017**, *32*, 139–154. [CrossRef]

30. Khan, H.; Shafique, M.; Khan, M.A.; Bacha, M.A.; Shah, S.U.; Calligaris, C. Landslide susceptibility assessment using Frequency Ratio, a case study of northern Pakistan. *Egypt. J. Remote. Sens. Space Sci.* **2019**, *22*, 11–24. [CrossRef]
31. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B.; Li, Y.; Catani, F.; Pourghasemi, H.R. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. *Comput. Geosci.* **2018**, *112*, 23–37. [CrossRef]
32. Soma, A.S.; Kubota, T.; Mizuno, H. Optimization of causative factors using logistic regression and artificial neural network models for landslide susceptibility assessment in Ujung Loe Watershed, South Sulawesi Indonesia. *J. Mt. Sci.* **2019**, *16*, 383–401. [CrossRef]
33. Pourghasemi, H.R.; Rahmati, O. Prediction of the landslide susceptibility: Which algorithm, which precision? *Catena* **2018**, *162*, 177–192. [CrossRef]
34. Hu, Q.; Zhou, Y.; Wang, S.; Wang, F. Machine learning and fractal theory models for landslide susceptibility mapping: Case study from the Jinsha River Basin. *Geomorphology* **2020**, *351*, 106975. [CrossRef]
35. Hong, H.; Pourghasemi, H.R.; Pourtaghi, Z.S. Landslide susceptibility assessment in Lianhua County (China): A comparison between a random forest data mining technique and bivariate and multivariate statistical models. *Geomorphology* **2016**, *259*, 105–118. [CrossRef]
36. Huang, Y.; Zhao, L. Review on landslide susceptibility mapping using support vector machines. *Catena* **2018**, *165*, 520–529. [CrossRef]
37. Bui, D.T.; Tsangaratos, P.; Nguyen, V.-T.; Liem, N.V.; Trinh, P.T. Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *Catena* **2020**, *188*, 104426. [CrossRef]
38. Chen, T.; Zhu, L.; Niu, R.-Q.; Trinder, C.J.; Peng, L.; Lei, T. Mapping landslide susceptibility at the Three Gorges Reservoir, China, using gradient boosting decision tree, random forest and information value models. *J. Mt. Sci.* **2020**, *17*, 670–685. [CrossRef]
39. 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]
40. Dou, J.; Yunus, A.P.; Tien Bui, D.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.-W.; Khosravi, K.; Yang, Y.; Pham, B.T. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Sci. Total Environ.* **2019**, *662*, 332–346. [CrossRef]
41. Hong, H.; Miao, Y.; Liu, J.; Zhu, A.-X. Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping. *Catena* **2019**, *176*, 45–64. [CrossRef]
42. 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]
43. Merghadi, A.; Yunus, A.P.; Dou, J.; Whiteley, J.; ThaiPham, B.; Tien Bui, D.; Avtar, R.; Abderrahmane, B. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Sci. Rev.* **2020**, *207*, 103225. [CrossRef]
44. Chen, W.; Xie, X.; Wang, J.; Pradhan, B.; Hong, H.; Tien Bui, D.; Duan, Z.; Ma, J. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena* **2017**, *151*, 147–160. [CrossRef]
45. 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]
46. Dou, J.; Yunus, A.P.; Merghadi, A.; Shirzadi, A.; Nguyen, H.; Hussain, Y.; Avtar, R.; Chen, Y.; Pham, B.T.; Yamagishi, H. Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning. *Sci. Total Environ.* **2020**, *720*, 137320. [CrossRef] [PubMed]
47. Wang, Y.; Fang, Z.; Hong, H. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci. Total Environ.* **2019**, *666*, 975–993. [CrossRef] [PubMed]
48. Sameen, M.I.; Pradhan, B.; Lee, S. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. *Catena* **2020**, *186*, 104249. [CrossRef]
49. Fang, Z.; Wang, Y.; Peng, L.; Hong, H. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. *Comput. Geosci.* **2020**, *139*, 104470. [CrossRef]

50. Yi, Y.; Zhang, Z.; Zhang, W.; Jia, H.; Zhang, J. Landslide susceptibility mapping using multiscale sampling strategy and convolutional neural network: A case study in Jiuzhaigou region. *Catena* **2020**, *195*, 104851. [CrossRef]
51. Wang, Y.; Fang, Z.; Wang, M.; Peng, L.; Hong, H. Comparative study of landslide susceptibility mapping with different recurrent neural networks. *Comput. Geosci.* **2020**, *138*, 104445. [CrossRef]
52. Gökçe, O.; Özden, Ş.; Demir, A. *Spatial and Statistical Distribution of Disasters in Turkey Inventory of Disaster Information*; Ministry of Public Works and Settlement General Directorate of Disaster Affairs, Department of Disaster Survey and Damage Assessment: Ankara, Turkey, 2008; p. 127. (In Turkish)
53. Akıncı, H.; Yavuz Özalp, A.; Özalp, M.; Temuçin Kılıçer, S.; Kılıçoğlu, C.; Everan, E. Production of Landslide Susceptibility Map using Bayesian Probability Model. *Int. J. 3-D Inf. Model.* **2015**, *4*, 16–33. [CrossRef]
54. Ersoy, S. *2016 Year Natural Disaster Welded Annual "Turkey and the World"*; No: 129; Chamber of Geological Engineers Publications: Ankara, Turkey, 2017. (In Turkish)
55. Erkan, M.A.; Kılıç, G.; Çamalan, G.; Güser, Y.; Çetin, S.; Odabaşi, E.; Soydam, M.; Akgündüz, A.S.; Eren, O.; Arabacı, H.; et al. *Meteorological Disasters 2018*; General Directorate of Meteorology, Meteorological Disasters Branch Office: Ankara, Turkey, 2019. (In Turkish)
56. TSI. Population of Municipalities, Villages and Quarters, Turkish Statistical Institute. Available online: https://biruni.tuik.gov.tr/medas/?kn=95&locale=tr (accessed on 29 July 2020).
57. Solmaz Oguz, F. Changes in Coastal Use between Ardeşen-Hopa between 1970–2000. Master's Thesis, İstanbul University, Istanbul, Turkey, 2002. (In Turkish).
58. Dağ, S.; Bulut, F. An Example for Preparation of GIS-Based Landslide Susceptibility Maps: Çayeli (Rize, NE Türkiye). *J. Geol. Eng.* **2012**, *36*, 35–62. (In Turkish)
59. Keskin, İ. *1:100.000 Scale Geological Map of Turkey, No: 178 Artvin-F46 Map Sheet*; General Directorate of Mineral Research and Exploration, Geological Research Department: Ankara, Turkey, 2013. (In Turkish)
60. Lee, S.; Choi, J. Landslide susceptibility mapping using GIS and the weight-of-evidence model. *Int. J. Geogr. Inf. Sci.* **2004**, *18*, 789–814. [CrossRef]
61. Chen, W.; Li, W.; Hou, E.; Zhao, Z.; Deng, N.; Bai, H.; Wang, D. Landslide susceptibility mapping based on GIS and information value model for the Chencang District of Baoji, China. *Arab. J. Geosci.* **2014**, *7*, 4499–4511. [CrossRef]
62. Yanar, T.; Kocaman, S.; Gokceoglu, C. Use of Mamdani Fuzzy Algorithm for Multi-Hazard Susceptibility Assessment in a Developing Urban Settlement (Mamak, Ankara, Turkey). *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 114. [CrossRef]
63. Colkesen, I.; Sahin, E.K.; Kavzoglu, T. Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression. *J. Afr. Earth Sci.* **2016**, *118*, 53–64. [CrossRef]
64. Erener, A.; Mutlu, A.; Düzgün, H.S. A comparative study for landslide susceptibility mapping using GIS-based multi-criteria decision analysis (MCDA), logistic regression (LR) and association rule mining (ARM). *Eng. Geol.* **2016**, *203*, 45–55. [CrossRef]
65. Park, S.; Kim, J. Landslide Susceptibility Mapping Based on Random Forest and Boosted Regression Tree Models, and a Comparison of Their Performance. *Appl. Sci.* **2019**, *9*, 942. [CrossRef]
66. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput. Geosci.* **2013**, *51*, 350–365. [CrossRef]
67. Kumar, D.; Thakur, M.; Dubey, C.S.; Shukla, D.P. Landslide susceptibility mapping & prediction using Support Vector Machine for Mandakini River Basin, Garhwal Himalaya, India. *Geomorphology* **2017**, *295*, 115–125.
68. Chen, W.; Xie, X.; Peng, J.; Shahabi, H.; Hong, H.; Tien Bui, D.; Duan, Z.; Li, S.; Zhu, A.-X. GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method. *Catena* **2018**, *164*, 135–149. [CrossRef]
69. He, Q.; Shahabi, H.; Shirzadi, A.; Li, S.; Chen, W.; Wang, N.; Chai, H.; Bian, H.; Ma, J.; Chen, Y.; et al. Landslide spatial modelling using novel bivariate statistical based Naive Bayes, RBF Classifier, and RBF Network machine learning algorithms. *Sci. Total Environ.* **2019**, *663*, 1–15. [CrossRef]

70. Wang, Q.; Guo, Y.; Li, W.; He, J.; Wu, Z. Predictive modeling of landslide hazards in Wen County, northwestern China based on information value, weights-of-evidence, and certainty factor. *Geomat. Nat. Hazards Risk* **2019**, *10*, 820–835. [CrossRef]
71. Vakhshoori, V.; Pourghasemi, H.R.; Zare, M.; Blaschke, T. Landslide Susceptibility Mapping Using GIS-Based Data Mining Algorithms. *Water* **2019**, *11*, 2292. [CrossRef]
72. Zhang, Y.-X.; Lan, H.-X.; Li, L.-P.; Wu, Y.-M.; Chen, J.-H.; Tian, N.-M. Optimizing the frequency ratio method for landslide susceptibility assessment: A case study of the Caiyuan Basin in the southeast mountainous area of China. *J. Mt. Sci.* **2020**, *17*, 340–357. [CrossRef]
73. Bordoni, M.; Galanti, Y.; Bartelletti, C.; Persichillo, M.G.; Barsanti, M.; Giannecchini, R.; Avanzi, G.D.; Cevasco, A.; Brandolini, P.; Galvee, J.P.; et al. The influence of the inventory on the determination of the rainfall-induced shallow landslides susceptibility using generalized additive models. *Catena* **2020**, *193*, 104630. [CrossRef]
74. Bragagnolo, L.; da Silva, R.V.; Grzybowski, J.M.V. Artificial neural network ensembles applied to the mapping of landslide susceptibility. *Catena* **2020**, *184*, 104240. [CrossRef]
75. Zhang, Q.; Yu, H.; Li, Z.; Zhang, G.; Ma, D.T. Assessing potential likelihood and impacts of landslides on transportation network vulnerability. *Transp. Res. Part D Transp. Environ.* **2020**, *82*, 102304. [CrossRef]
76. Lee, S.; Min, K. Statistical analysis of landslide susceptibility at Yongin, Korea. *Environ. Geol.* **2001**, *40*, 1095–1113. [CrossRef]
77. Pourghasemi, H.R.; Pradhan, B.; Gokceoglu, C.; Moezzi, K.D. A comparative assessment of prediction capabilities of Dempster–Shafer and Weights-of-evidence models in landslide susceptibility mapping using GIS. *Geomat. Nat. Hazards Risk* **2013**, *4*, 93–118. [CrossRef]
78. Dağ, S. Landslide Susceptibility Analysis of Çayeli Region (Rize) by Statistical Methods. Ph.D. Thesis, Karadeniz Technical University, Trabzon, Turkey, 2007. (In Turkish).
79. Çevik, E.; Topal, T. GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek (Turkey). *Environ. Geol.* **2003**, *44*, 949–962. [CrossRef]
80. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **2014**, *11*, 425–439. [CrossRef]
81. Keskin, İ. *1:100.000 Scale Geological Map of Turkey, No: 179 Artvin-E47 and F47 Map Sheets*; General Directorate of Mineral Research and Exploration, Geological Research Department: Ankara, Turkey, 2013. (In Turkish)
82. Regmi, A.D.; Devkota, K.C.; Yoshida, K.; Pradhan, B.; Pourghasemi, H.R.; Kumamoto, T.; Akgun, A. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arab. J. Geosci.* **2014**, *7*, 725–742. [CrossRef]
83. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
84. Catani, F.; Lagomarsino, D.; Segoni, S.; Tofani, V. Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 2815–2831. [CrossRef]
85. Taalab, K.; Cheng, T.; Zhang, Y. Mapping landslide susceptibility and types using Random Forest. *Big Earth Data* **2018**, *2*, 159–178. [CrossRef]
86. Youssef, A.M.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Al-Katheeri, M.M. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides* **2016**, *13*, 839–856. [CrossRef]
87. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
88. Lin, Y.-P.; Chu, H.-J.; Wu, C.-F. Spatial pattern analysis of landslide using landscape metrics and logistic regression: A case study in Central Taiwan. *Hydrol. Earth Syst. Sci. Discuss* **2010**, *7*, 3423–3451. [CrossRef]
89. Wang, Q.; Li, W. A GIS-based comparative evaluation of analytical hierarchy process and frequency ratio models for landslide susceptibility mapping. *Phys. Geogr.* **2017**, *38*, 318–337. [CrossRef]
90. Nohani, E.; Moharrami, M.; Sharafi, S.; Khosravi, K.; Pradhan, B.; Pham, B.T.; Lee, S.; Melesse, A.M. Landslide Susceptibility Mapping Using Different GIS-Based Bivariate Models. *Water* **2019**, *11*, 1402. [CrossRef]
91. Çan, T.; Duman, T. Rainfall Intensity—Duration Relationship and Event Landslide Inventory of Hopa (Eastern Black Sea) Region. In Proceedings of the National Engineering Geology and Geotechnical Symposium, Çukurova University, Adana, Turkey, 12–14 October 2017; pp. 103–110. (In Turkish).
92. Althuwaynee, O.F.; Pradhan, B.; Lee, S. Application of an evidential belief function model in landslide susceptibility mapping. *Comput. Geosci.* **2012**, *44*, 120–135. [CrossRef]

93. Bulut, F.; Boynukalın, S.; Tarhan, F.; Ataoğlu, E. Causes of Landslides in the Eastern Region of Fındıklı District (Rize). In Proceedings of the 2nd National Landslide Symposium, Adapazarı, Sakarya, Turkey, 25–26 October 1995; pp. 143–152. (In Turkish).
94. Bahrami, S.; Rahimzadeh, B.; Khaleghi, S. Analyzing the effects of tectonic and lithology on the occurrence of landslide along Zagros ophiolitic suture: A case study of Sarv-Abad, Kurdistan, Iran. *Bull. Eng. Geol. Environ.* **2020**, *79*, 1619–1637. [CrossRef]
95. Yılmaz, B.S. *General Features, Causes and Prevention Methods of Landslides in the Eastern Black Sea Region*; General Directorate of Mineral Research and Exploration Report: Ankara, Turkey, 1995. (In Turkish)
96. Chung, C.-J.F.; Fabbri, A.G. Validation of spatial prediction models for landslide hazard mapping. *Nat. Hazards* **2003**, *30*, 451–472. [CrossRef]
97. Brenning, A. Spatial prediction models for landslide hazards: Review, comparison and evaluation. *Nat. Hazards Earth Syst. Sci.* **2005**, *5*, 853–862. [CrossRef]