geosciences

MDPI

Article

The Open Landslide Project (OLP), a New Inventory of Shallow Landslides for Susceptibility Models: The Autumn 2019 Extreme Rainfall Event in the Langhe-Monferrato Region (Northwestern Italy)

Michele Licata *, Victor Buleo Tebar, Francesco Seitone and Giandomenico Fubelli

Department of Earth Sciences, University of Turin, Via Valperga Caluso, 35, 10125 Turin, Italy; victor.buleotebar@unito.it (V.B.T.); francesco.seitone@unito.it (F.S.); giandomenico.fubelli@unito.it (G.F.)
* Correspondence: michele.licata@unito.it; Tel.: +39-34-6047-1366

Abstract: Landslides triggered by heavy rainfall pose significant threats to human settlements and infrastructure in temperate and equatorial climate regions. This study focuses on the development of the Open Landslide Project (OLP), an open source landslide inventory aimed at facilitating geostatistical analyses and landslide risk management. Using a multidisciplinary approach and open source, multisatellite imagery data, more than 3000 landslides triggered by the extreme rainfall of autumn 2019 in northwestern Italy were systematically mapped. The inventory creation process followed well-defined criteria and underwent rigorous validation to ensure accuracy and reliability. The dataset's suitability was confirmed through multivariate correlation and Double Pareto probably density function. The OLP inventory effectiveness in assessing landslide risks was proved by the development of a landslide susceptibility model using binary logistic regression. The analysis of rainfall and lithology revealed that regions with lower rainfall levels experienced a higher occurrence of landslides compared to areas with higher peak rainfall. This was attributed to the response of the lithological composition to rainfalls. The findings of this research contribute to the understanding and management of landslide risks in anthropized climate regions. The OLP has proven to be a valuable resource for future geostatistical analysis.

Keywords: inventory; landslide susceptibility; shallow landslide; extreme rainfall; Ordinary Kriging; binary logistic regression; Italy; Sentinel-2; Google Earth; Double Pareto

Citation: Licata, M.; Buleo Tebar, V.; Seitone, F.; Fubelli, G. The Open Landslide Project (OLP), a New Inventory of Shallow Landslides for Susceptibility Models: The Autumn 2019 Extreme Rainfall Event in the Langhe-Monferrato Region (Northwestern Italy). *Geosciences* **2023**, *13*, 289. https://doi.org/10.3390/geosciences13100289

Academic Editors: Leonhard Blesius and Jesus Martinez-Frias

Received: 20 July 2023
Revised: 15 September 2023
Accepted: 21 September 2023
Published: 23 September 2023

1. Introduction

The temperate and equatorial climate regions [1] are generally highly anthropized, and heavy rainfall is common in these areas, often occurring seasonally. The interaction between precipitation and geological/geomorphological characteristics can trigger widespread landslides [2,3], resulting in severe damage to human settlements.

In autumn 2019, high-rainfall events affected northwestern Italy: following several days of heavy rain, an intense and persistent meteorological system remained stationary for approximately 12 h on 12 October 2019 in southeast Piedmont, resulting in a peak of over 400 mm of rainfall during the afternoon [4]. Subsequently, the entire month of November experienced significant rainfall, with a second peak of 600 mm/day on November 24 within the Ligurian–Piedmont boundary [5]. Consequently, thousands of landslides were triggered, causing extensive damage to infrastructure and urban centers in the region and resulting in severe economic losses.

Numerous authors have proposed landslide susceptibility models as powerful tools for landslide risk management [6–8]. Therefore, the creation of a landslide inventory assumes great importance due to its central role in the modeling methods. However, it is crucial to establish well-defined criteria [9] and statistically validate the inventory [10].

Adhering to this rigorous procedure allows for a detailed understanding of the inventory and its appropriate utilization in the modeling process.

In this study, we introduce the Open Landslide Project (OLP), an open source landslide inventory specifically designed for landslide geostatistical analysis. As the project is in its early stages, the data are not yet available for direct download. However, interested parties can request access to the data directly from the authors. Our research proposes a multidisciplinary approach to efficiently create an inventory of recent shallow landslides, minimizing the time required for mapping while covering large spatial areas and multiple years of data. By adapting the criteria suggested by Harp et al. [9] and Taylor et al. [10], we systematically mapped over 3000 rainfall-triggered landslides that occurred during autumn 2019 in the northwest region of Italy. This extensive mapping was made possible using opensource, multisatellite imagery data.

To ensure the accuracy and reliability of the landslide inventory, it was subjected to validation [2,11–19]. Additionally, a landslide susceptibility model was developed using binary logistic regression [20] to assess the effectiveness of the dataset in a real case study.

2. Study Area

2.1. Geomorphological and Geological Setting

The study area is situated in the northwest of Italy, encompassing the regions of Piedmont and Liguria (Figure 1a). It is located approximately north of Genoa and southeast of Acqui Terme (see Figure 1b). The study focused on ten different catchments situated on the orographic right of the Bormida River, covering a total area of approximately 762 km^2. The region consists of numerous scattered small villages, and the average population density is around 40 persons per km^2. Notably, the area contains several key transportation routes, including the "A26" highway and the Asti-Genoa railway, which serve as vital connections between the Piedmont and Liguria regions (Figure 1b).

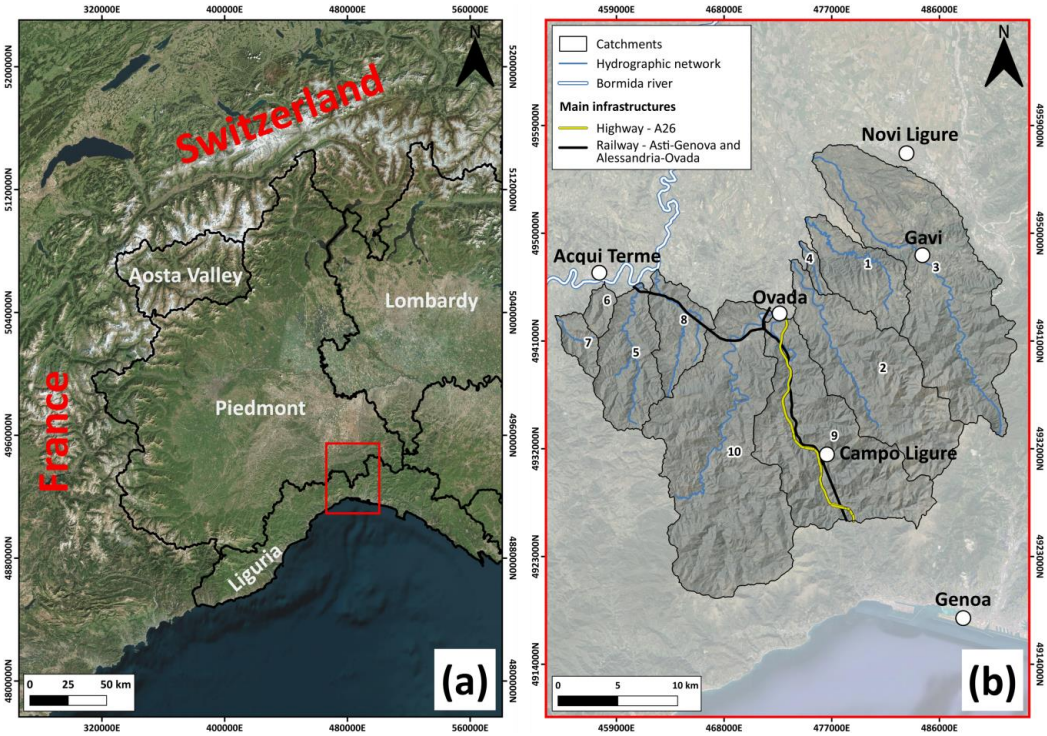

Figure 1. General overview of the study area. (**a**) Geographical setting; (**b**) main catchments and urban centers; the catchments' numbers correspond to the following: 1: Albedosa, 2: Piota 3: Lemme, 4: Cremosino, 5: Visone, 6: Ravanasco, 7: Calioga, 8: Caramagna, 9: Stura di Ovada, and 10: Orba.

The geomorphological setting in the study area can be distinguished between the southern and northern regions, each exhibiting distinct landscapes and geomorphological characteristics (Figure A2).

The southern area is situated in the upper sectors of watersheds, with elevations ranging from 400 to 1200 m above sea level (Figure A1a). The average slope steepness in this region is around 30–35 degrees (see Figure A1b), and the valleys typically exhibit V-shaped profiles. The landscape features well-preserved and easily recognizable regressive erosive morphologies (see a). The hydrographic network in this area is poorly hierarchical, partly due to the slope steepness [21]. Consequently, the terrain is mountainous and often characterized by outcrops of metamorphic rocks.

In contrast, the northern area is characterized by lower elevations ranging between 140 and 400 m above sea level (Figure A1a). The landforms in this region are relatively smoothed (see Figures A1a and A2b), with wide alluvial valleys occupying the lower areas and deep gullies present in the upper areas. The hilly morphology of southern Piedmont is primarily influenced by the nature and structure of the bedrock [22]. The slopes in this region are affected by significant colluvial deposits that rarely expose the underlying sedimentary bedrock. After the Burdigalian stage, in conjunction with a phase of rapid reversal in the drainage direction of the entire basin, this sector has experienced progressive uplift (e.g., [23]). This uplift has led to increased incision of the hydrographic network, resulting in a typical terraced morphology [21].

In the entire study area, the slope aspect does not exhibit significant isotropic behavior (Figure A1c). This suggests a complex morphology that is not dominated by structural elements at the basin scale. The slope patterns and characteristics display a diverse and intricate arrangement rather than a consistent and uniform orientation throughout the study area.

The land cover within the study area exhibits significant variation from south to north (Figure A1d). In the southern sector, lower altitudes are primarily covered by broad-leaved trees and mixed forests, while coniferous forests also appear at higher elevations. Sparse vegetation areas and transitional forested areas occupy smaller portions of the landscape. In the northern sector of the study area, agricultural crops, particularly viticultural ones, are widespread. This region includes the UNESCO Langhe-Monferrato vineyard landscapes, recognized as a World Heritage Site.

From a geological perspective, the study area exhibits a complex geodynamic evolution and has experienced both compressive and extensional tectonic phases during the Alpine and Apennine orogeneses. However, based on lithological and structural differences, two primary regions can be distinguished: the Alpine metamorphic units to the south and the Tertiary sedimentary units to the north.

The Alpine sector is situated within the southern termination of the Western Alps, as described by Piana et al. [24]. It is bounded to the east by the Sestri-Voltaggio Zone (SVZ), a transition zone between the Alps and the Apennines characterized by fragile structures. The ongoing interaction between the European and Adria plates, along with the intermediate oceanic domains, has shaped the orogenic system of the Alps and Apennines, serving as a significant tectonic element in the central-western Mediterranean region [25,26]. Starting from the Oligocene period, the Ligurian Piedmontese Ocean began subducting southeastward beneath the Adria microplate. This subduction continued until the collision between the European and Adria plates in the Middle Eocene, resulting in the formation of the western Alpine chain [27]. Subsequent subduction reversal in a northwesterly direction during the late Eocene led to the development of the Apennine chain [28]. Consequently, the southeastern termination of the Western Alps exhibits distinct characteristics compared to the rest of the range [25]. This area is characterized by metaophiolite sequences belonging to the Voltri Group and the Sestri-Voltaggio Zone, representing remnants of the oceanic crust from the Ligurian Piedmontese Ocean and the Earth's mantle [21]. The prominent lithotypes in the mapped area include schists (calc-schists, mica-schists, quartz-schists),

serpentinites, metabasites, gabbro/metagabbro, metabasites, and minor occurrences of peridotites and quartzites.

In contrast, the northern sector of the study area is situated within the Tertiary Piedmont Basin (TPB), as described by Bigi et al. [29] and Balestro et al. [30] (see Figure 2a). This is a sedimentary basin primarily located in Piedmont and is considered to be a synorogenic basin that evolved into a piggy-back basin [25]. The TPB formed as a result of tectonic and sedimentary processes between the European plate and the Adriatic microplate, with nonuniform sedimentation occurring on the developing Alpine thrust belt [31,32]. Initial deposition consisted of sandy deposits, followed by marly and evaporitic sediments typical of the ongoing transgressive phase [33] (Figure 2b). D'Atri et al. [34] identified three primary tectonic-sedimentary domains within the TPB: the Langhe Basin, the Turin Hills, and the Monferrato Hills. Within the study area, the Miocene deposits of the Alto Monferrato domains are predominantly exposed [29], forming a regional monocline with a dip of 15–20° and a dip direction toward the northwest. The main lithotypes in this region consist of carbonate-rich mudstones, arenites, siltstones, and conglomerates (Figure 2b).

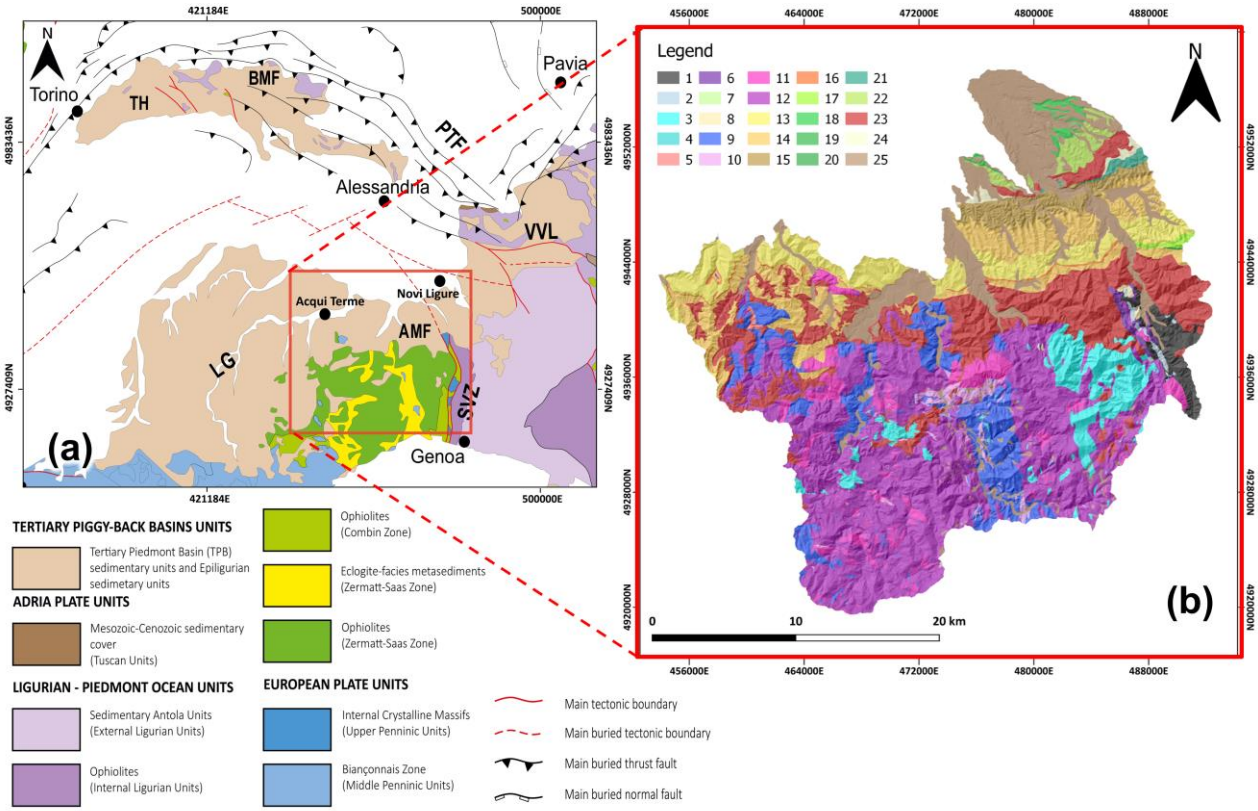

Figure 2. Geological setting of the study area. (**a**) Tectono-stratigraphic map of the Western Alps and the Northern Apennines (modified after Balestro et al. [30]). (**b**) Lithological map: (1) slate; (2) basalt; (3) peridotite; (4) quartzite; (5) orthogneiss, paragneiss; (6) limestone; (7) jasper; (8) dolomite; (9) calc schist, mica schist, quartz schist; (10) metabasite; (11) gabbro, metagabbro; (12) serpentinite; (13) arenite, mudstone; (14) carbonate rich mudstone, arenite; (15) arenite, siltstone; (16) arenite, impure limestone; (17) arenite, carbonate rich mudstone; (18) sand, gravel, clay/siltstone; (19) sand, gravel; (20) carbonate rich mudstone; (21) carbonate rich mudstone, siltstone; (22) carbonate rich mudstone, silt, sand; (23) conglomerate, arenite; (24) gypsum or anidrite, limestone, clay; and (25) gravel, sand.

2.2. Pluviometric Setting and Rainfall 2019 Events

The rainfall patterns vary across the study area, with distinct differences between the southern and northern sectors. In the northern Monferrato hilly area, the typical annual

rainfall is around 600 mm, whereas in the southern mountainous region, the typical annual rainfall is higher, ranging from 1300 to 1600 mm [22]. A recent study on rainfall in Piedmont covering the period 2004–2016 classified the study area as having a sublittoral pluviometric regime, characterized by maximum precipitation occurring in autumn and minimum in winter [35]. In the analyzed period, the minimum annual rainfall recorded was 677 mm, while the maximum reached 1800 mm.

Autumn 2019 was particularly rainy in the study area, with two main precipitation events occurring in October and November. From October 19 to 24, intense, moist southern currents associated with an Atlantic disturbance over Spain affected Piedmont. This meteorological system brought adverse weather conditions and extremely heavy rainfall, initially impacting the eastern part of Piedmont and then spreading to the western region. In southeastern Piedmont, the peak of the event was observed on October 21, characterized by a self-regenerating, semistationary weather system that produced a V-shaped thunderstorm with intense convective precipitation [4]. The second major precipitation event took place between November 22 and 25, with intense and prolonged rainfall affecting the entire study area. It was caused by intense hot and humid sirocco currents, resulting in rainfall associated with orographic blocking phenomena between the Alps and the Apennines [5]. The two meteorological events had very different meteorological features, representing both extreme events. According to [5], in the study area, the cumulative precipitation in 2019 exceeded the annual average rainfall by 21%.

2.3. Rainfall-Induced Landslides

The cumulative precipitation during the October–November period, along with the occurrence of the two extreme rainfall events, led to a significant increase in landslides within the study area. The majority of these landslides occurred in soil and/or debris materials, exhibiting translational/rotational slide or flow kinematics (see Figure 3). Although only one casualty was reported, extensive damage was inflicted upon infrastructure, buildings, and crops across the entire territory. The northern sector, being more densely populated and cultivated, experienced particularly severe impacts.

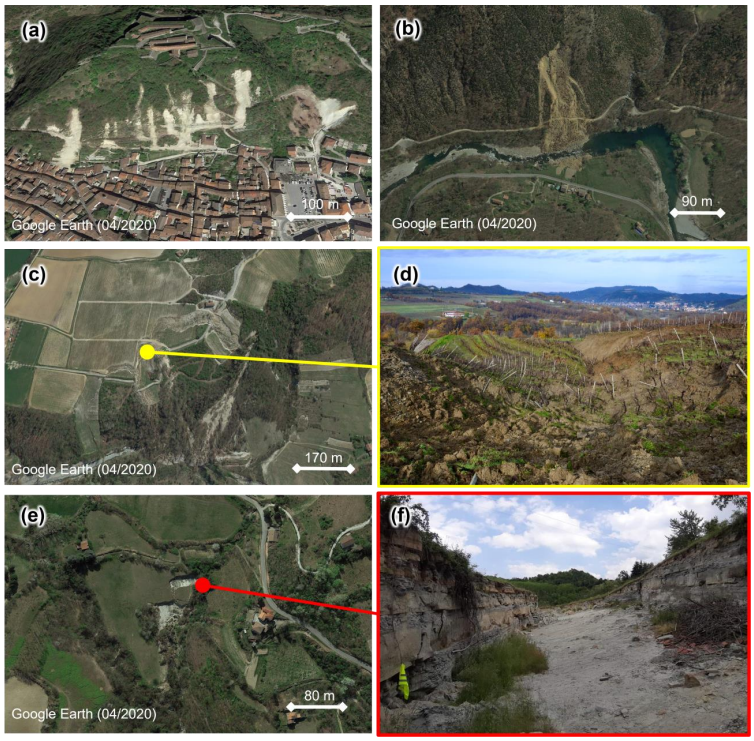

Figure 3. Satellite and field views of the typology of main landslides that occurred in autumn 2019 in the study area. (**a**) Shallow landslides near the town of Gavi; (**b**) landslides that occurred in the southern

sector of the study area, between Belforte Monferrato and Gnocchetto, near the A6 motorway; (**c**) complex landslide that occurred near Gavi; (**d**) related photograph taken from the top of the landslide (© Davide Ferrarese 2019); (**e**) a translational landslide near the village of Cavatore; and (**f**) a related photograph taken in the crown sector.

3. Materials and Methods

3.1. Rainfall Dataset

The rainfall dataset used in this research was collected from the meteorological networks of Arpa Piedmont, Arpa Liguria, and the Agrometeorological Network of Piedmont (RAM Piedmont). These datasets have undergone quality control by the respective institutional authorities and are publicly available for download from the official websites [36–38].

The Arpa Piedmont network has been recording data since 1988 and consists of 400 monitoring stations, with a density of approximately one station every 100 km^2. The rain gauges used in this network are CAE PMB2 tipping buckets, equipped with a 1000 cm^2 diameter collector [39]. The Arpa Liguria network, on the other hand, was established in 1998 and includes 70 monitoring stations using the same instrument as the Arpa Piedmonte network. The RAM Piedmont network has been active since the second half of the 1990s and comprises 120 monitoring stations. These stations were primarily installed for agricultural monitoring purposes and are located in close proximity to the main cultivation areas. The rain gauges used in the RAM Piedmont network are tipping buckets of the SIAP UM7525 model, equipped with a 1000 cm^2 diameter collector [35].

For the purpose of this research, only the stations located within the considered catchment areas and their surrounding regions (as depicted in Figure 4) were selected. The entire dataset was analyzed to determine the temporal distribution of precipitation and to ensure the temporal continuity of precipitation records. Only stations that had complete records without gaps for the entire duration of both rainfall events were considered. This allowed for standardization in data analysis between the two events.

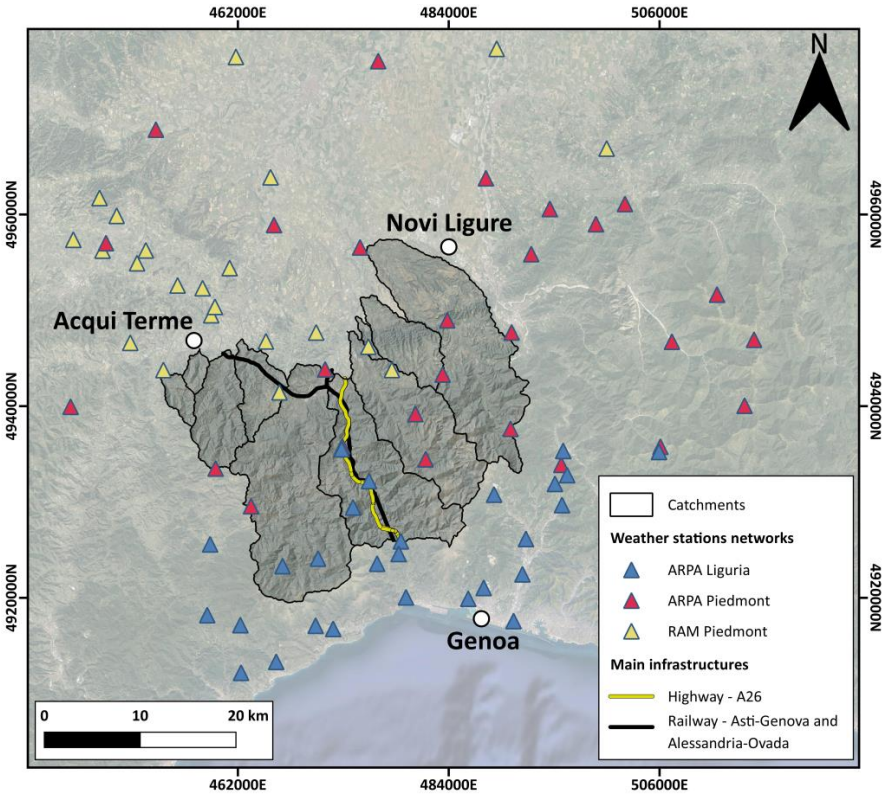

Figure 4. Map of weather station networks used for pluviometric analysis.

In order to interpolate the rainfall data within the study area, Ordinary Kriging (OK) was employed as an interpolation method. The Kriging method is commonly used for geostatistical analysis of rainfall data [40–44]. However, it should be noted that the number of stations and their spatial distribution are important parameters that influence the reliability of the final interpolated pattern [45]. To carry out the interpolation, we used the Smart-Map plugin developed by Pereira et al. [46] for the QGIS software (version 3.10 or higher). Smart-Map is an open source plugin implemented in Python 3.7 and based on the OK algorithm developed by Isaaks et al. [47]. It allows for fitting five models of isotropic theoretical semivariograms (linear, linear with sill, exponential, spherical, and Gaussian) and selecting the best model based on the cross-validation method.

3.2. Landslide Inventory

The criteria utilized for creating the OLP dataset were inspired by Harp et al. [9] and Taylor et al. [10]. Harp et al. [9] provided valuable guidelines for developing a comprehensive and statistically robust landslide inventory. The specific imagery and mapping criteria proposed by these authors, along with the strategies we adopted, can be found in Tables 1 and 2, respectively. While basing our work on these original papers, we made slight modifications to the criteria to suit the requirements of our current research. These modifications primarily involved condensing the criteria for improved conciseness and splitting the fourth imagery criterion into distinct components that consider spatial continuity and temporal resolution properties. Taylor et al. [10] introduced an additional implementation of the third mapping criterion, which assesses whether the landslide inventory aligns with the established statistical behavior of substantially complete inventories. This criterion is of fundamental importance to the OLP as it ensures the statistical robustness required for geostatistical analysis of landslides.

Table 1. Imagery criteria adopted for the creation of OLP dataset. Modified after Harp et al. [9].

N	Imagery Criterion	Strategies Adopted
1	The imagery must be continuous in space and time in the entire area to span the landslides' distribution triggered by the same event.	GEP and S2 satellite images covering all the area for the period considered in this research.
2	The imagery must have a resolution that allows identification of individual landslides as small as a few meters across.	In the area considered for this research, GEP provided very high-pixel resolution imagery.
3	The imagery must have stereo coverage or be able to be draped over a digital elevation model to obtain a stereo-like perspective view.	GEP 2.5D viewer provides a stereo-like perspective view.
4	The imagery must be as cloud-free and shadow-free as possible.	GEP images are elaborated to be cloud free. As for the S2 image, it is possible to select the maximum cloud-coverage
5	The imagery must be acquired as soon as possible after the triggering event to capture the initial features of the landslides and the terrain or infrastructure that they affect.	The GEP images available: pre-events and postevents. The S2 images available (according to the 4th criterion) were pre-events, between the events, and post-events.

Table 2. Mapping criteria adopted for the creation of OLP dataset. Modified after Harp et al. [9].

N	Mapping Criterion	Strategies Adopted
1	The landslides must be defined as polygons in a GIS system, either as a single polygon representing the entire landslide or as two or more polygons that define the landslide source and the landslide deposit.	Single polygon and derivatives.
2	The landslide polygons must be plotted on a topographic map or GIS layer that is registered to a topographic map or geo-registered image.	QGIS Geodatabase.
3	The entire population of the event's triggered landslides exceeding the minimum resolution of the imagery must be mapped.	Frequency–Area analysis.

Regarding the imagery criteria, we utilized multitemporal images from Google Earth Pro® (GEP) and Sentinel-2 (S2) for the mapping of landslides. GEP provides high spatial resolution and cloudless imagery (Figure 5a), but its temporal coverage is discontinuous with significant gaps. However, the pseudo three-dimensional (2.5D) viewer in GEP offers a stereoscopic-like perspective of the study area, and the digital elevation model (DEM) with a resolution of 30 m is sufficient for understanding the topographic relief [48]. On the other hand, S2 has offered a continuous-up-to-5-day temporal resolution of multispectral imagery since 2016, covering almost the entire globe. However, these images are affected by seasonal cloud cover conditions, which can locally impact the temporal images for wide areas. The maximum spatial resolution of S2 imagery is 10 m [49]. In terms of the area of interest, GEP images were available only for the pre-events period (April 2018) and postevents period (April 2020) (Figure 5a). S2 images available for the 4th imagery criterion included the pre-events period (14 September 2019) (Figure 5b), between the events period (13 November 2019) (Figure 5c), and postevents period (23 December 2019) (Figure 5d). Additionally, we considered S2 imagery corresponding to the postevents period of GEP (11 April 2020) (Figure 5e).

The mapping procedure was conducted using GIS software to create a geodatabase in accordance with the first and second mapping criteria. For each landslide, the following elements were mapped: (i) a polygon encompassing the source area, run-out zone, and accumulation zone; (ii) the centroid of the mapped polygon; (iii) the highest point along the crown of the landslide; (iv) the lowest point of the landslide tip; and (v) the minimum oriented bounding rectangle along the sliding direction. The combination of GEP desktop multitemporal images and the 2.5D viewer facilitated the identification and mapping of the areas affected by landslide phenomena. The geospatial analysis was carried out using QGIS and Python in order to extract the morphometric parameters for each landslide.

The field-mapping procedure for landslides encountered three technical limitations: (i) the logistical challenge of mapping a large number of landslides across a wide area, (ii) the historical period of the events occurring a few months prior to the implementation of strict COVID-19 lockdown measures, and consequently (iii) the significant time gap between the occurrence of the events and the ability to access and survey the study area. These limitations made it infeasible to conduct direct field surveys to measure parameters such as alteration thickness of geological materials and the depth of the slide surface. However, it was observed that nearly all the mapped landslides exhibited the following characteristics:

1. They were triggered by a single accumulation of seasonal precipitation during the October–November period.
2. They displayed clearly identifiable crowns, flanks, and main scarps and/or main bodies through satellite imagery.
3. They were in an initial state of activation.

4. They exhibited velocity rates ranging from m/min to m/s, classified as "rapid" and "very rapid" movements [50].

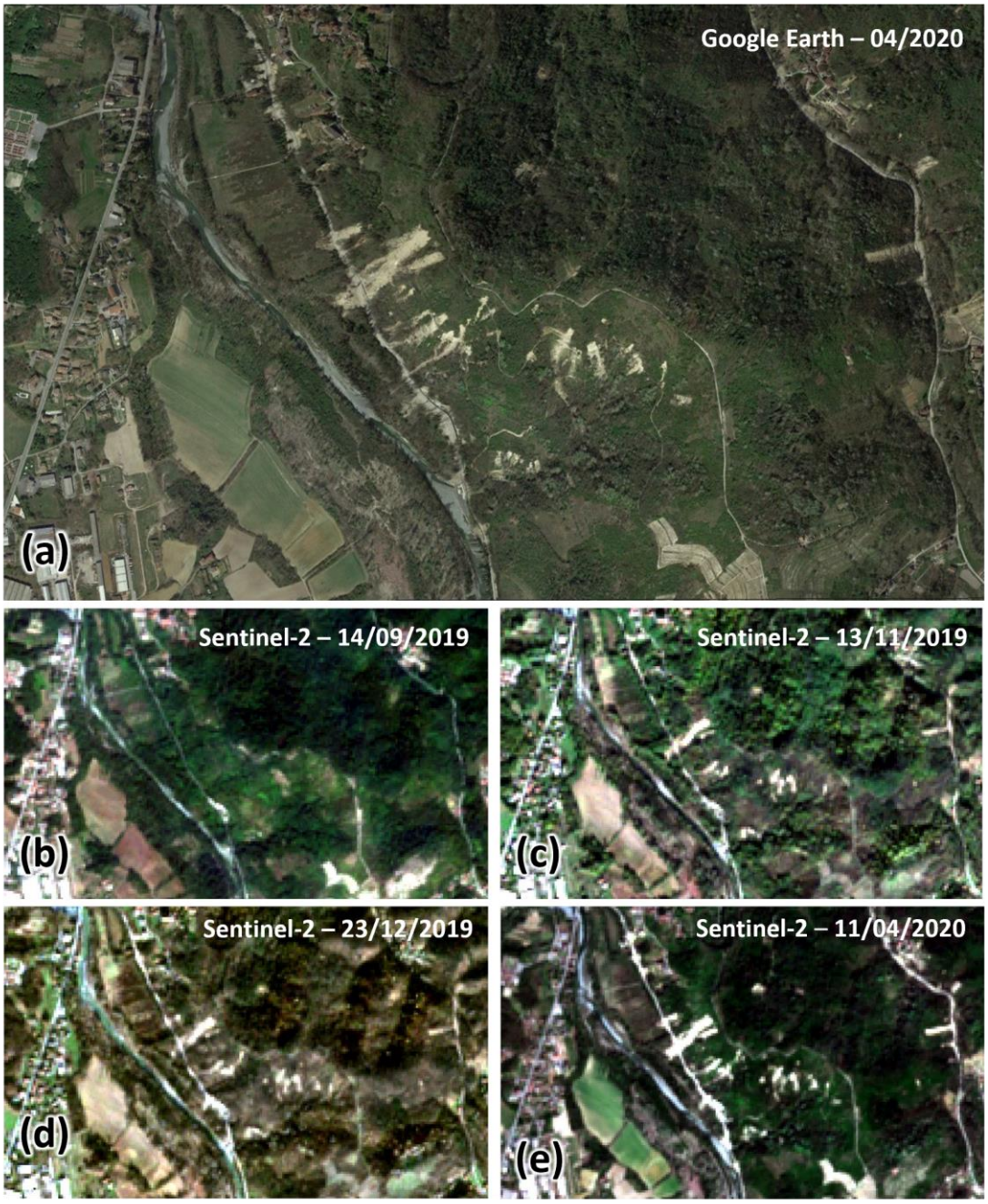

Figure 5. Multitemporal of satellite imagery used for OLP landslide mapping. (**a**) GEP high-resolution image postevents; (**b**) S2 image pre-events; (**c**) S2 image between events; (**d**) S2 image postevents; (**e**) S2 image at same time of GEP image.

Based on points 1–4, the mapped phenomena were classified as "shallow landslides" according to the general definition proposed by Cruden and Varnes [50], referring to landslides involving soil/earth materials and exhibiting rapid to very rapid movement.

In this research, we adopted a morphometrical classification approach for landslides. The criteria considered the aspect-ratio parameter [51,52] to classify shallow landslides into four categories: transversal, isometric, longitudinal, and elongated, as defined by Tian et al. [53] (Figure A3). A statistical analysis was applied to the entire dataset to identify additional morphometric parameters for classifying shallow landslides.

According to Stark and Hovius [54] and Tebbens [55], the ideal theoretical distribution of landslide areas should follow a power-law distribution. However, the mapping of small-size landslides often leads to a censoring effect, which affects the probability density estimation for this category. To quantify this underestimation, area-frequency analysis is a valuable tool (as specified in the third mapping principle in Table 2). Several landslide datasets worldwide were tested with the Double Pareto [54] probability distribution function [2,10–17,19].

In the analysis of the OLP dataset, Python statistical tools were employed, and the Double Pareto probability density function (Equation (1)) was calculated using constant logarithmic (base 10) bins with a size of 0.1 [55].

$$p(A_L) = PDF(A_L|\propto,\beta,t,c,m) = \left[\frac{\beta}{t} * \left(1 - \left(\frac{1+\left(\frac{m}{t}\right)^{-\propto}}{1+\left(\frac{c}{t}\right)^{-\propto}}\right)^{\frac{\beta}{\propto}}\right)^{-1}\right] * \left[\frac{\left(1+\left(\frac{m}{t}\right)^{-\propto}\right)^{\frac{\beta}{\propto}}}{\left(1+\left(\frac{A_L}{t}\right)^{-\propto}\right)^{(1+\frac{\beta}{\propto})}} * \left(\frac{A_L}{t}\right)^{(-\propto-1)}\right] \quad (1)$$

where A_L is the landslide area, α is the scaling exponent that controls the power function for the large-sized landslides, β is the scaling exponent that controls the power function for the small-sized landslides, t is a parameter that constrains the rollover position; c and m are, respectively, the minimum and the maximum sizes of the landslide area.

3.3. Landslide Susceptibility Model

The selection of mapping units is a crucial step in landslide susceptibility modeling and has been extensively investigated [8,56–62] due to its impact on prediction performance, resulting metrics, and mapping design. In this research, we opted to use grid-cell units, as they offer advantages such as faster procedures for data partitioning and easier and more representative assignment of covariate values. Grid-cell units have demonstrated their reliability in assessing flow and moderately sized landslides.

Various factors contribute to slope instability, including topography, climate, and geology. Drawing from the extensive scientific literature on landslide predisposing factors for susceptibility models [63–66], we considered ten different parameters as predictors based on their potential significance as proxies. These parameters include elevation (ELE), steepness (SLO), aspect (ASP), profile curvature (PRF), plan curvature (PLC), stream power index (SPI), topographic wetness index (TWI), topographic roughness index (TRI), land use obtained from the Corine Land Cover 2018 (USO), and outcropping lithology (LITO) (Figure 6).

ELE is considered to be primarily closely related to topographical conditions, but also to climatic conditions (temperature and precipitation). SLO is a key factor widely used in modeling, as it influences slope angles, water speed, shear stress, and shear strength. ASP is considered for its association with the orientation of layered materials and environmental processes related to insolation, such as soil moisture and vegetation presence. PLC and PRF are important for surface-runoff characteristics, including velocity and water concentration. The TWI serves as an indicator of soil saturation, while the TRI reflects surface texture based on elevation differences among grid-cells. Lastly, the SPI measures the potential for water-induced flow erosion, considering slope and contributing area.

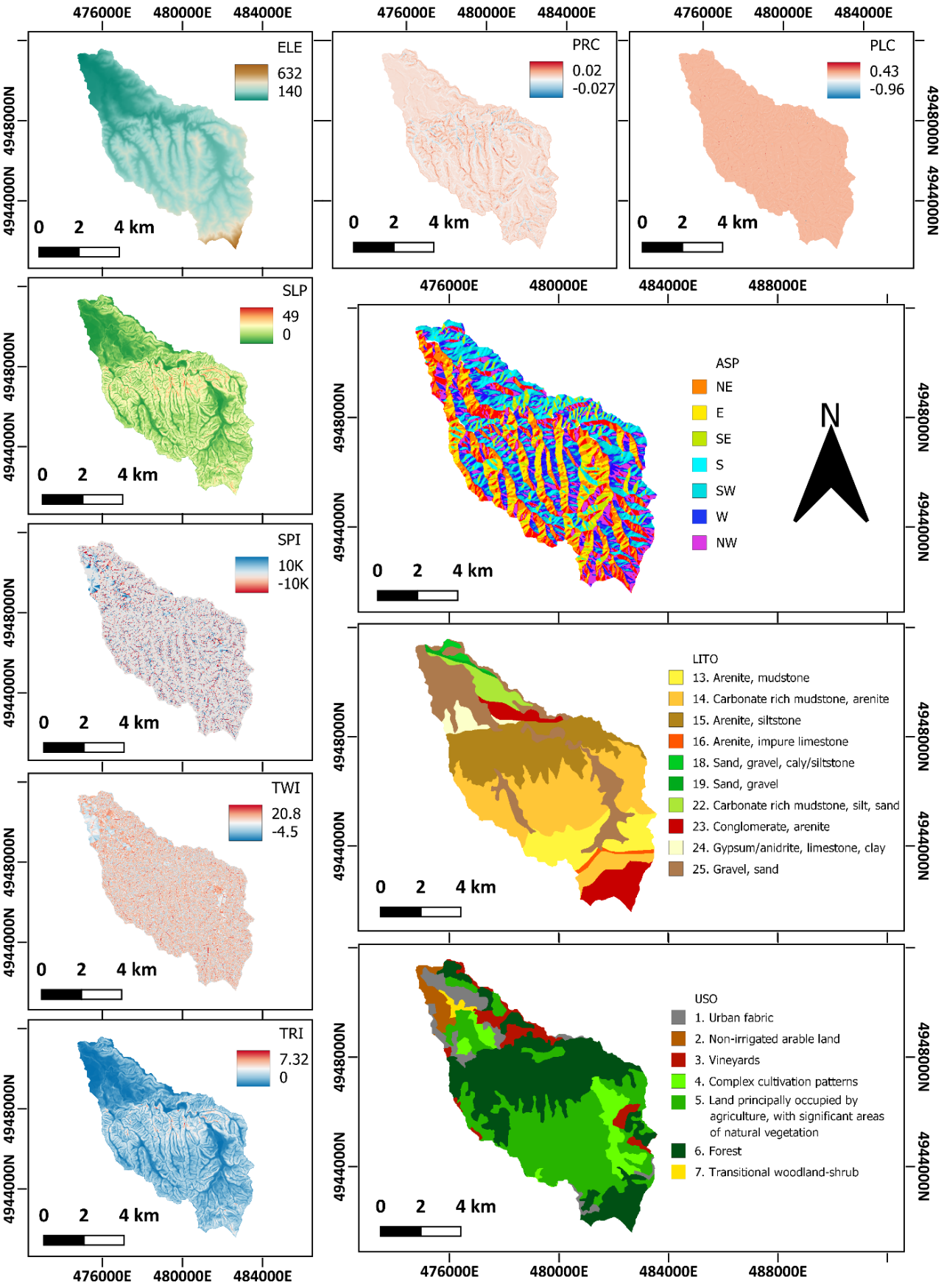

Figure 6. Predictors considered for landslide susceptibility model of the Albedosa river catchment.

Most of the selected parameters were treated as continuous variables and derived from a DEM provided by Tarquini et al. [67] with a spatial resolution of 10 m pixel size. The DEM was processed using the open source System for Automated Geoscientific Analyses (SAGA) [68]. For the categorical variables, we reclassified the ASP into eight categories: north (0–22.5; 337.5–360), northeast (22.5–67.5), east (67.5–112.5), southeast (112.5–157.5), south (157.5–202.5), southwest (202.5–247.5), west (247.5–292.5), and northwest (292.5–337.5).

The outcropping lithology (LITO) data, obtained from Arpa Piemonte at a scale of 1:250,000, was reclassified into ten categories based on the main lithotypes. These categories include (13) arenite, mudstone; (14) carbonate-rich mudstone, arenite; (15) arenite, siltstone; (16) arenite, impure limestone; (18) sand, gravel, clay/siltstone; (19) sand, gravel; (22) carbonate-rich mudstone, silt, sand; (23) conglomerates, arenite; (24) gypsum or anhydrite, limestone, clay; (25) gravel, sand (Figure 6).

For the land use (USO) variable, we reduced the number of categories to seven based on our knowledge of the area: (1) urban fabric; (2) nonirrigated arable land; (3) vineyards; (4) complex cultivation patterns; (5) land principally occupied by agriculture, with significant areas of natural vegetation; (6) forest; (7) transitional woodland-shrub.

Regarding the response variable, we extracted the highest point along the crown (LIP—Landslide Identification Point) from each polygon in our landslide inventory [65,69,70] and assigned a value of 0 or 1 to each pixel based on the absence or presence of landslides, respectively.

To assess multicollinearity among the selected parameters, a Variance Inflation Factor (VIF) analysis (Equation (2)) was conducted using the R-package "usdm" [71]. This analysis was performed prior to the statistical modeling technique to ensure the reliability of interpreting the regression coefficients as the result of a linear relationship between a set of variables that are considered independent from each other.

$$VIF_i = \frac{1}{1 - {R_i}^2} = \frac{1}{Tolerance} \quad (2)$$

where ${R_i}^2$ is defined as the unadjusted coefficient for regressing the ith independent variable on the remaining ones. High values of ${R_i}^2$ means that most of the variation in the ith covariate is explained by all other covariates [72].

The procedure was applied to both continuous and categorical variables. The results showed that, excluding the slope (SLO) and the topographic roughness index (TRI) parameters, all the variables passed the multicollinearity test. Considering the results, we decided to exclude the TRI and to keep the SLO for further analysis.

With regard to landslide susceptibility assessment, various statistical methodologies have been developed and compared in recent years to improve the ability to predict future landslides [6–8,20,73–75]. Among these techniques, Binary Logistic Regression has been widely utilized due to its reliable performance and several advantages compared to other methodologies. These advantages include its ease of implementation and interpretation of results, providing measures of predictor importance and the direction of association, and assuming a linear relationship between the response variable and the covariates [20,75,76].

Binary logistic regression is used to describe the variability of a binary response variable based on a set of covariates, and the general expression can be written as follows:

$$Y = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p \quad (3)$$

where Y represents the dependent or response variable, which in this case is the presence or absence of landslides in a given grid-cell. The β coefficients represent the regression coefficients for each independent variable, denoted as x_p.

As the response variable Y can only take two possible values (0 for absence and 1 for presence of landslides), the binary logistic regression model calculates the probability that

a particular grid-cell will have a value of 1 based on the conditions determined by the set of covariates. The expression for this probability is:

$$\eta(\pi) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p \quad (4)$$

where η is the link function that relates the right-hand side of Equation (4) to the probability that Y takes the value 1, denoted as π. In binary logistic regression, the objective is to calculate the probability that a particular grid-cell belongs to one of the two categories, either presence or absence. Therefore, the y-axis of the regression is constrained to the range of probability values between 0 and 1.

To address this constraint, the probability of occurrence is transformed using the logit function, which is the inverse of the logistic function. The logit function allows for the calculation of the log-odds of occurrence, providing a linear relationship between the predictors and the log-odds. The logit function is expressed as:

$$\eta(\pi) = loglog\{\pi/(1-\pi)\} \quad (5)$$

By applying Equation (5), the floor restriction on the probability range of 0 to 1 is removed, and the transformed probability now covers the entire real number range from negative infinity to positive infinity [77]. This transformation allows for a linear relationship between the predictors and the log-odds. The resulting outcomes of this transformation are the log-odds, also known as the logit values. The coefficients in the binary logistic regression model describe the constant unit by which the log-odds increase with the increase in every unit in the corresponding covariate. These coefficients provide insights into the strength and direction of the association between the predictors and the likelihood of landslide occurrence.

The code for the analysis was implemented using the statistical software R [78]. The data preprocessing stage involved creating 50 balanced partition samples, with an equal number of positive (presence of landslides) and negative cases (absence of landslides). A random splitting procedure was then performed, where 75% of the cases were used for training the models, and the remaining 25% were used for validation. This process was repeated 100 times to obtain multiple fitting models.

For each repetition, various outputs were generated, including regression coefficients, confusion matrices, Receiver Operating Characteristics (ROC) curves, Area Under the Curve (AUC) values, and maps showing the probabilities of landslide occurrence. The AUC values were calculated to assess the discriminatory power of the models, with higher AUC values indicating better performance.

To determine the best probability cutoff for classifying landslides, the Youden index [79] was used. This index is obtained by dividing the sum of true positives and true negatives (accuracy) by the total population, which includes true positives, true negatives, false positives, and false negatives. The different probability cutoffs were evaluated to find the one that maximizes the average accuracy across the 100 replicates.

4. Results

4.1. Rainfall Analysis Results

The pluviometric stations selected for the analysis included 27 stations from Arpa Piedmont, 27 stations from Arpa Liguria, and 22 stations from RAM Piedmont.

When examining the daily rainfall data collected between 1 October and 31 December 2019 (Figure 7a), three distinct clusters can be observed. The first cluster occurred in October, followed by a second cluster in November, and finally a third cluster in December. However, it is important to note that the December cluster was excluded from rainfall data interpolation. This exclusion was determined based on the observation that between 4 December and 31 December, a total of 22 out of the 77 stations experienced at least one day of nonoperation, which could introduce inconsistencies in the data.

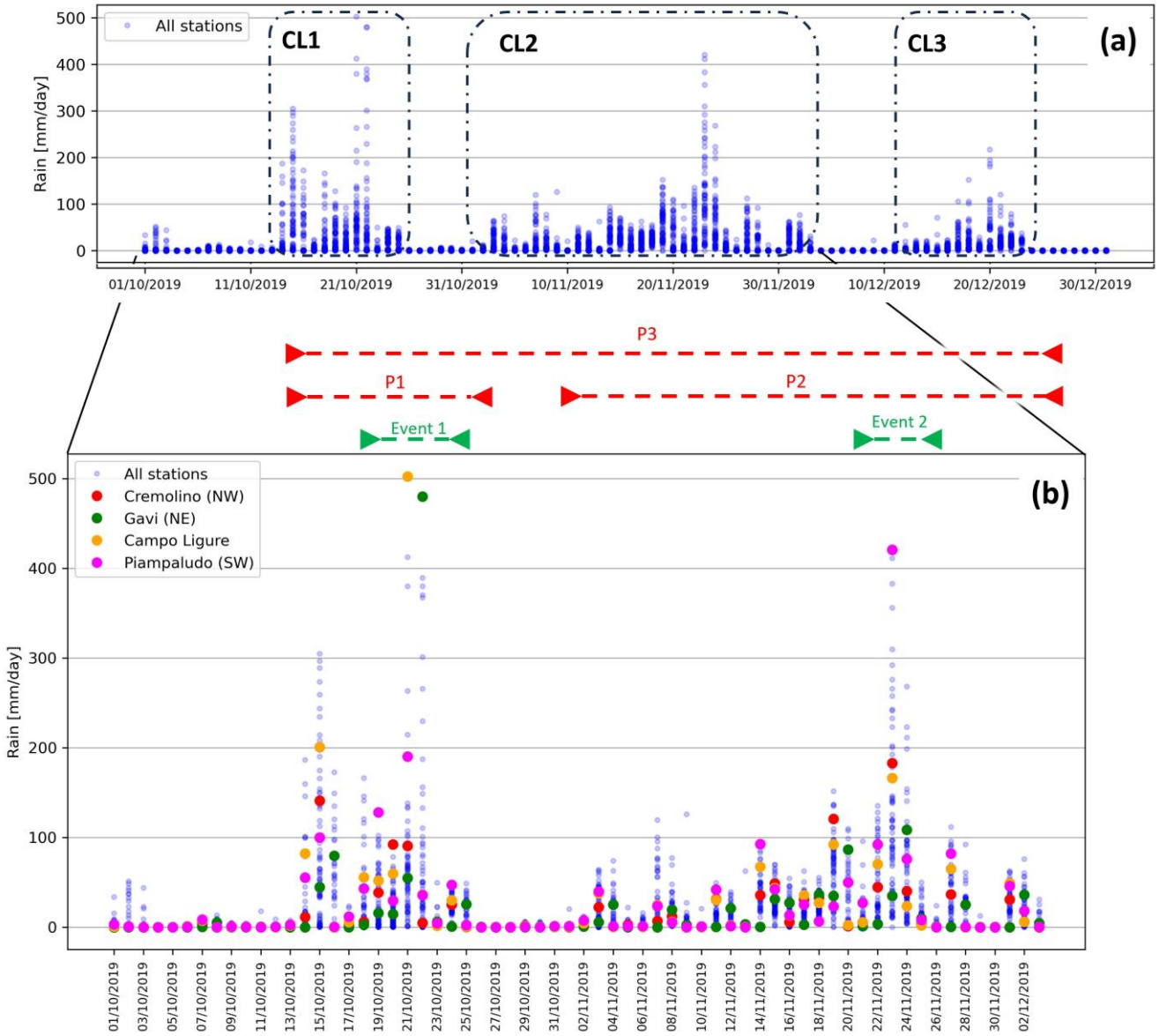

Figure 7. Daily rainfall data recorded for the 77 stations considered for the October–December time range. (**a**) Full time range and main pluviometric clusters highlighted in black dash-dotted rectangles; (**b**) focus on the time range of the main clusters (CL1 and CL2): the Arpa events' timings (Event 1 and Event 2) are highlighted in green dashed lines, while the periods (P1, P2, and P3) chosen for Kriging interpolation are represented by red dashed lines.

In Figure 7b, two distinct rainfall periods can be observed. The first period (P1) occurred from 14 October to 25 October, with the maximum daily rainfall recorded on 21 October. The second period (P2) occurred from 2 November to 3 December, with the peak daily rainfall observed on 23 November. In addition to P1 and P2, we also identified a third period (P3) from 14 October to 3 December (Figure 7b) that embraced the whole time-range from the start of P1 until the end of P2.

The P1 was characterized by a twelve-day period of rainfall, including two days of extremely heavy rain. The highest peak of rainfall (505.2 mm/day) occurred on 21 October in the Campo Ligure village, which is located near the middle of the study area in the Stura di Ovada river catchments (see Figure 1b). In contrast, P2, which lasted for about a month, experienced almost continuous rainfall. Only one extreme rainfall peak (420.6 mm/day)

was recorded on 23 November at the Piampaludo station, situated in the southwest of the Orba river catchments (see Figure 1b).

The cross-validation analysis of OK revealed R^2 values ranging between 0.832 and 0.873, with an RMSE between 93.37 and 148.79 (see Figure A4a). The Smart-Map plugin suggested using the spherical model only for P2 based on the R^2 comparison. However, considering the minimal variation in R^2 values between the spherical and linear models, we used the linear model (see Figure A4b) for all period interpolations to facilitate model comparisons. Ordinary Kriging maps are showed in Figure 8.

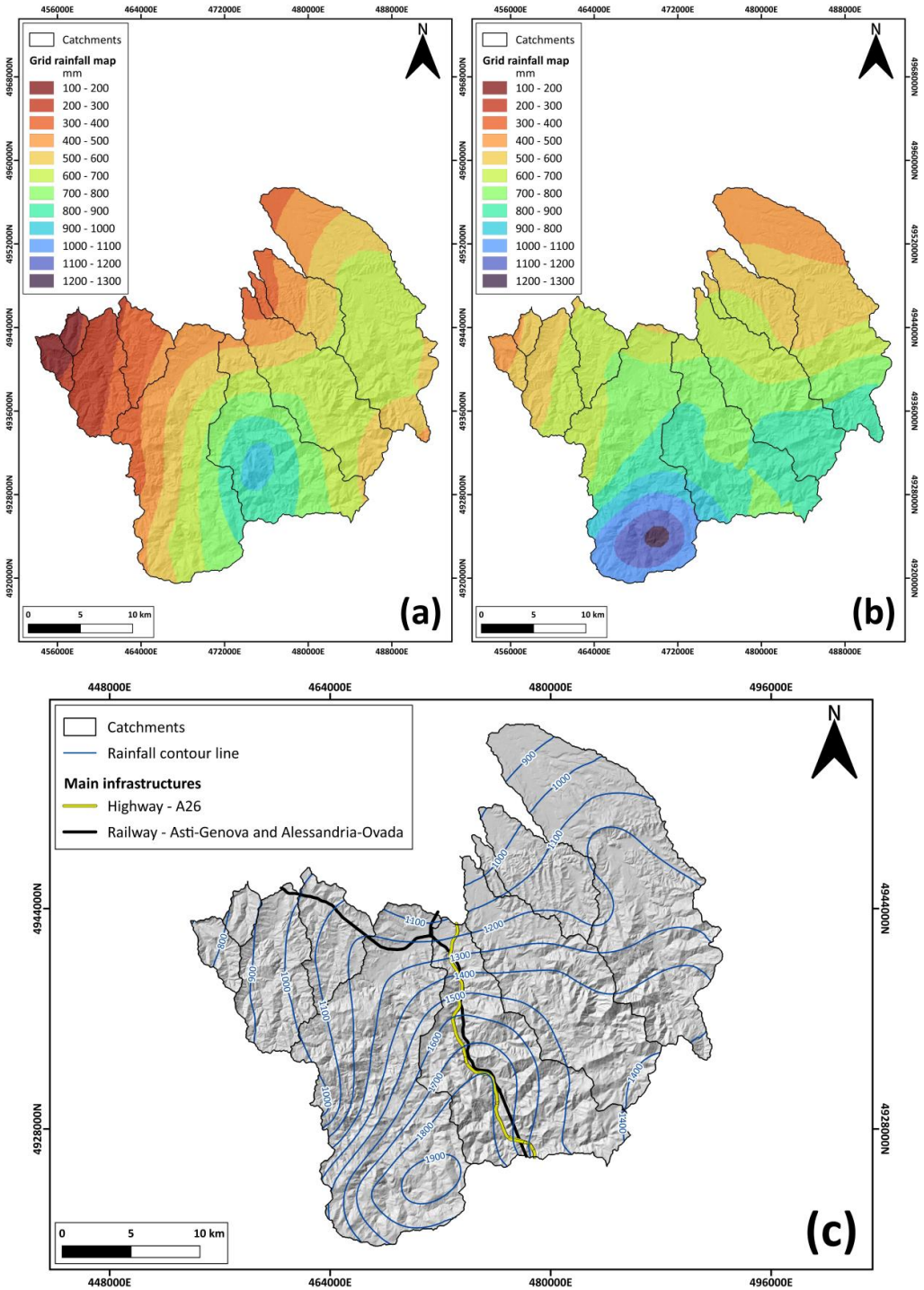

Figure 8. Precipitation maps of the study area. (**a**) Heatmap of P1; (**b**) heatmap of P2; (**c**) contour map of P3 overlayed on hillshade and main anthropic infrastructure in the study area.

The minimum and maximum cumulative rainfall in each catchment highlight significant differences. The Stura di Ovada river catchments and the Orba river catchments (Figure 1b) exhibit a considerable gap between the minimum and maximum values. This disparity is particularly noticeable in P1 (Figure 9a) compared to P2 (Figure 9b). However, in P3, which considers both the P1 and P2 periods, there is a smoothing effect on the difference between the minimum and maximum values (Figure 9c).

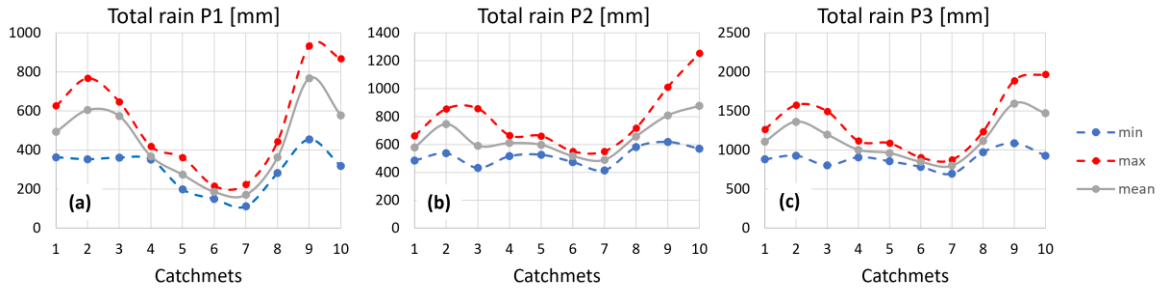

Figure 9. Minimum and maximum rainfall cumulative for catchments in the study area. (**a**) Kriging interpolation of P1, (**b**) Kriging interpolation of P2, (**c**) Kriging interpolation of P3.

4.2. Landslide Inventory

The OLP inventory for landslides triggered by the extreme precipitation events of autumn 2019 record a total of 3300 landslides over a studied area of approximately 762 km^2 (see Figure 10a). Using the QGIS software, we extracted the area value and average raster values of DEM, slope steepness, and slope aspect for each landslide polygon. The overall statistics for the entire dataset are summarized in Table 3.

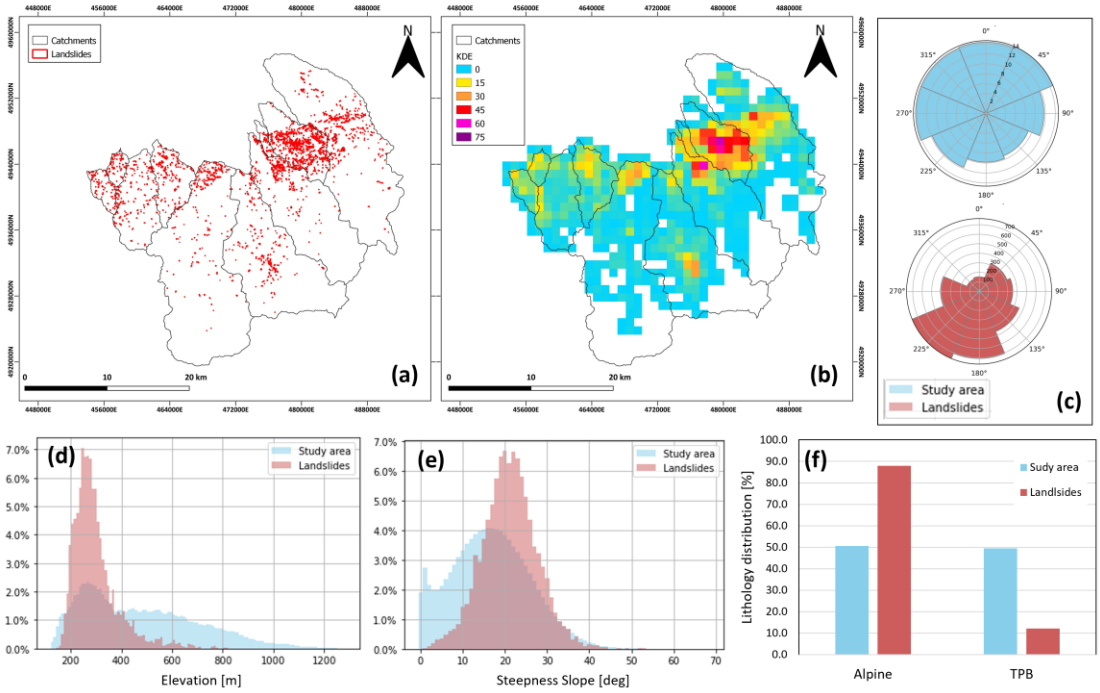

Figure 10. Landslides on the study area. (**a**) Landslides; (**b**) Kernel Density Estimation heatmap of landslides (landslides/km^2); (**c**) slope aspect of study area versus landslides; (**d**) elevation frequency distribution of study area versus landslides; (**e**) slope steepness frequency distribution of study area versus landslides; (**f**) tectonostratigraphic units of study area versus landslides.

Table 3. Summary of the main morphometric parameters of the slope where the landslides occurred.

Parameter	Elevation [m]	Slope Steepness [°]	Slope Aspect [°]	Area [m^2]
min	148	1.95	3	11
max	825	53.48	360	7875
std	90	6.48	80	690
median	281	21.56	207	280

The spatial distribution of landslides was analyzed using the QGIS Kernel Density Estimation (KDE) interpolation toolbox (see Figure 10b); the morphometric characteristics (elevation, slope steepness, slope aspect) of the landslide source areas were compared with those of the entire study area (see Figure 10c–e).

The KDE interpolation reveals a widespread density of values up to 15 landslides/km^2, with a higher concentration in the northern sector of the area, except for the area of Prato Ligure, which exhibits an anomalous density between 16 and 30 landslides/km^2. The main cluster corresponds to basins 1 and 4, with densities reaching up to 66 landslides/km^2, and to a lesser extent in the northern sector of river basins 2 and 3. A second smaller cluster can be identified in the northern sectors of basins 5, 6, 7, 8, and 10, with density values up to 29 landslides/km^2.

Diagrams in Figure 10c illustrate the slope aspect orientation of the area and the landslides. In the catchment area, it exhibits slight isotropy with a preferential orientation toward the west, north, and east. Regarding the orientation of the landslides, a strong isotropic trend toward the south and southwest is evident. The elevation in the mapped area (Figure 10d) displays a bimodal behavior, with the first absolute frequency peak around 300 m above sea level (a.s.l.) and a second flattened peak around 500 m a.s.l. Conversely, the landslides exhibit only one frequency peak around 300 m a.s.l. Both the catchment area and the landslides demonstrate a unimodal distribution in terms of slope steepness (Figure 10e), with peaks slightly out of phase, measuring 18 and 21 degrees, respectively. Lastly, Figure 10f provides information about the area coverage for both Alpine and TPB lithologies in the entire study area. It indicates that the area coverage is similar for both lithologies. However, there is a higher frequency of landslide occurrence in TPB lithologies (89.9%) compared to Alpine lithologies (12.1%) (Figure 11a).

Regarding the TPB lithologies, landslides are more frequent where alternations of mudstones and arenite in various proportions are present. There are also minor occurrences in alternations of arenite and siltstones or arenite and conglomerates (Figure 11b). For the Alpine lithologies, the highest frequency of landslides is observed in schists, with minor occurrences in Serpentinites lithologies (See Figure 11c). Figure 11d shows the existent link between number of landslides, rainfall, and main lithological units. The two different clusters correspond to TPB and Alpine lithologies with precipitation peaks around 1200 mm/km^2 and 1800 mm/km^2, respectively.

The OLP dataset was subjected to statistical analysis to quantify the correlation between landslide morphometric and topographical parameters across the entire landslide inventory. A correlation matrix was calculated on the entire dataset and the results are shown in the correlogram depicted in Figure 12.

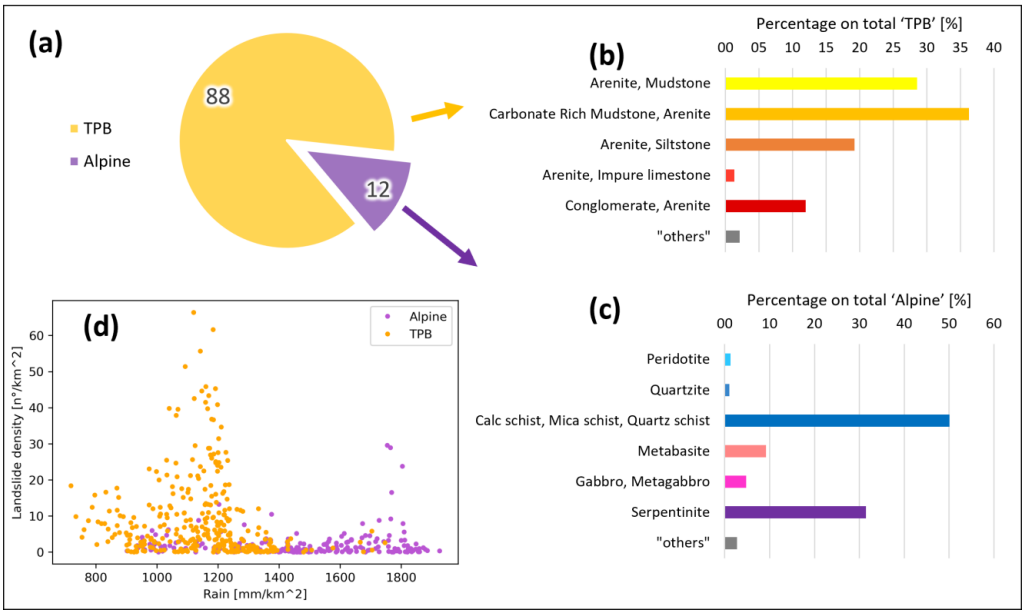

Figure 11. Landslides lithologies. (**a**) Alpine versus TPB units; (**b**) TPB lithologies; (**c**) alpine lithologies; (**d**) landslide density versus total 2019 autumn rainfall cumulation per square kilometer.

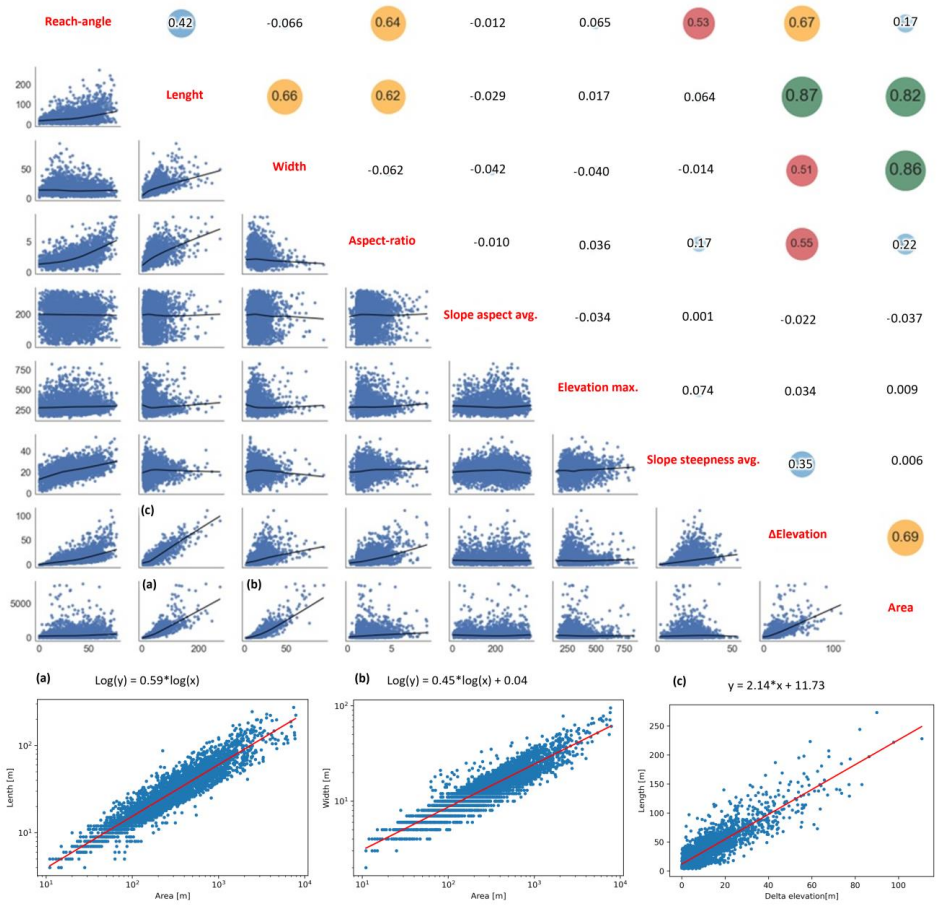

Figure 12. Correlogram of OLP inventory. (**a**) Linear regression of length versus area in log–log space; (**b**) linear regression of width versus area in log–log space; (**c**) linear regression of length versus Δ elevation.

The correlogram shows that the slope aspect has a very weak statistical influence. Similarly, the average slope steepness (slope steepness avg) and the maximum elevation (elevation max.) exhibit light significance, showing only a relationship with the reach angle and the maximum difference in elevation between the crown and tip (Δ elevation). On the other hand, statistically significant correlations are observed for the landslide width, length, and derived aspect-ratio. The most notable correlations are observed between area and width (Corr. = 0.86), area and length (Corr. = 0.82), as well as Δ elevation and length (Corr. = 0.87) (see Figure 12a–c).

Following the approach suggested by Tian et al. [53], we divided the landslides into four classes based on aspect-ratio thresholds (Figure 13a). The majority of the mapped landslides (66%) fall under the "elongated" type, while the remaining percentages include "longitudinal" (20.8%) and "isometric" (11.9%) types. The classification identified only 1.3% of landslides as "transversal" (see Figure 13b).

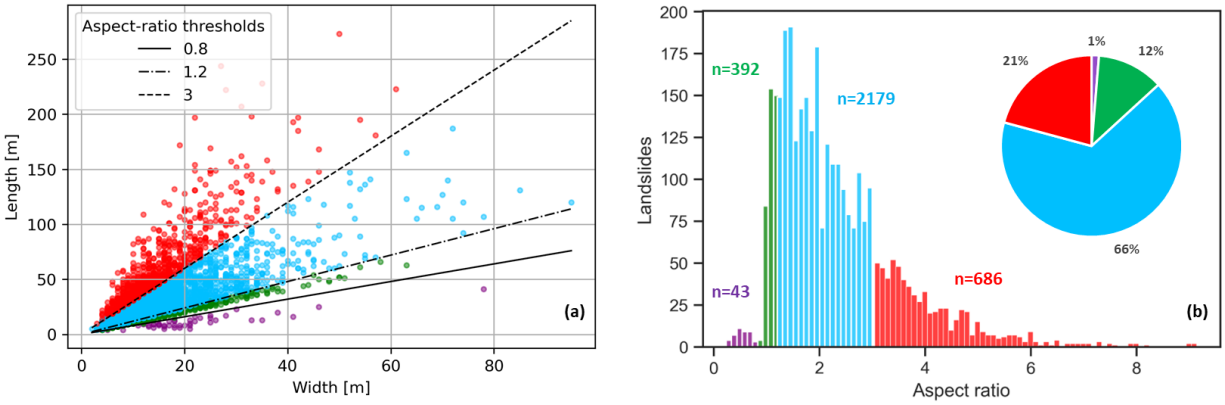

Figure 13. Aspect-ratio classification (purple: transverse, green: isometric, light-blue: longitudinal, red: elongated) and correlation with morphometrical parameters and lithologies. (**a**) Landslides' length versus width distribution; (**b**) aspect-ratio frequency distribution.

The frequency-area distribution of the OLP inventory is shown in Figure 14a. To fit the observed data, we employed a maximum likelihood estimation approach to fit the Double Pareto model to the observed data. The obtained values for α and β are 1.4 and 2.4, respectively, in radians. Additionally, the rollover for area values occurs at 140 m^2.

In Figure 14b, the frequency-area distributions of landslides are shown for different aspect-ratio classes. The area distribution of landslides exhibits a similar power-law behavior, and the rollover is consistent across the classified subset. However, when considering small landslides below the rollover, the frequency distribution of the longitudinal type is an order of magnitude lower than that of other classes. The transverse class, on the other hand, deviates from this pattern, displaying a lack of areas above 2×10^3. The rollover is not clearly identifiable, and the distribution does not fit the Double Pareto distribution.

Further analysis of the frequency-area distribution of landslides was conducted by classifying them based on lithotypes. The distributions for both "Alpine" (Figure 14c) and "TPB" bedrock (Figure 14d) lithologies demonstrate a good fit with the Double Pareto distribution. However, the peridotite and quartzite lithotypes exhibit a gap in landslide sizes greater than 2×10^2 m^2.

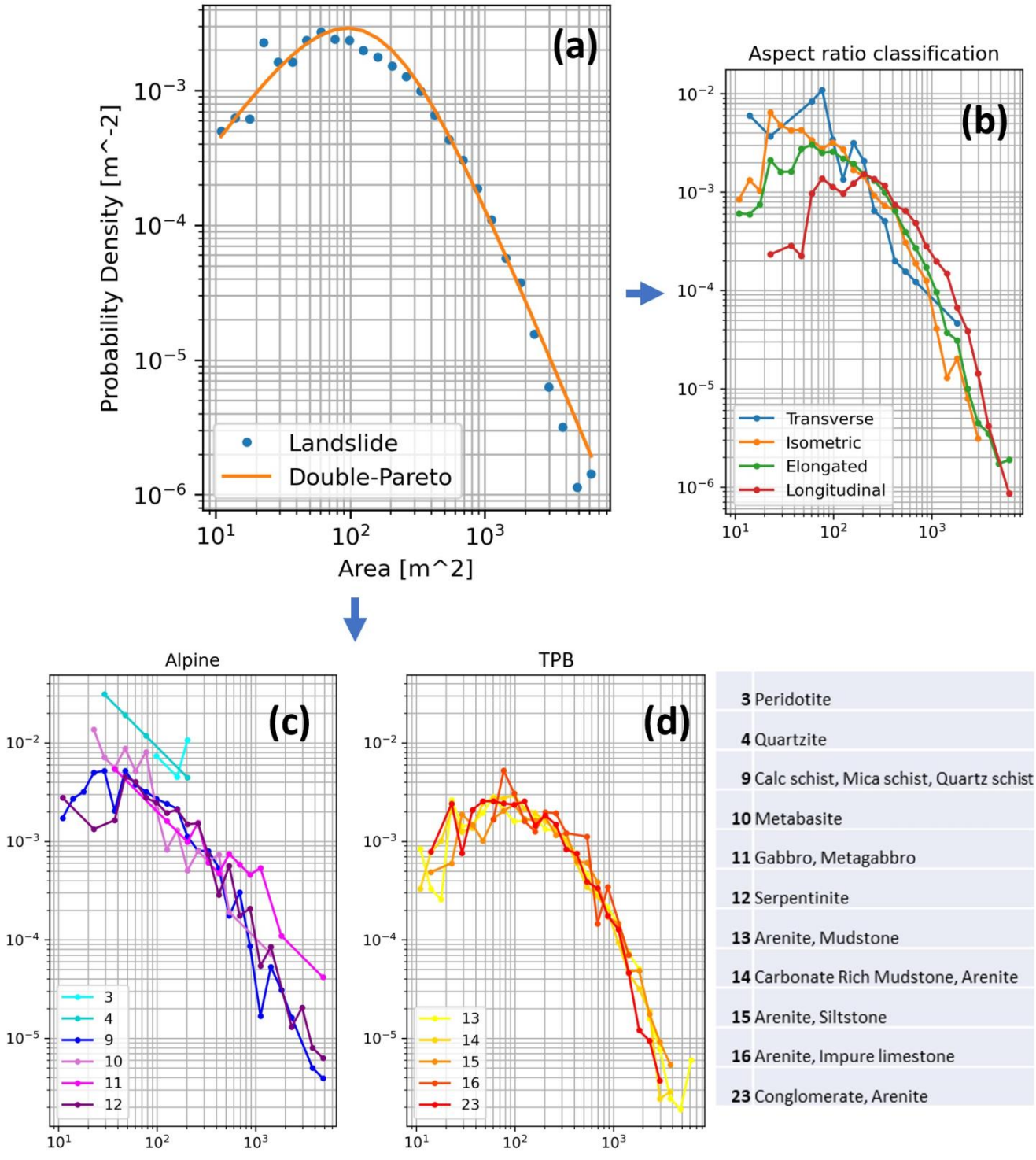

Figure 14. Frequency-area analysis of OLP inventory. (**a**) Full dataset analysis and Double Pareto probability density function; (**b**) frequency-area distribution for aspect-ratio classes; (**c**) frequency-area distribution for main Alpine lithologies; (**d**) frequency-area distribution for main TPB lithologies.

4.3. Landslide Susceptibility Model

Multicollinearity analysis was conducted on both continuous and categorical variables (Figure 15). The analysis revealed that all variables, except for SLO and TRI, passed the multicollinearity test. Based on these results, the decision was made to exclude TRI from further analysis, while retaining SLO as a parameter for subsequent analysis.

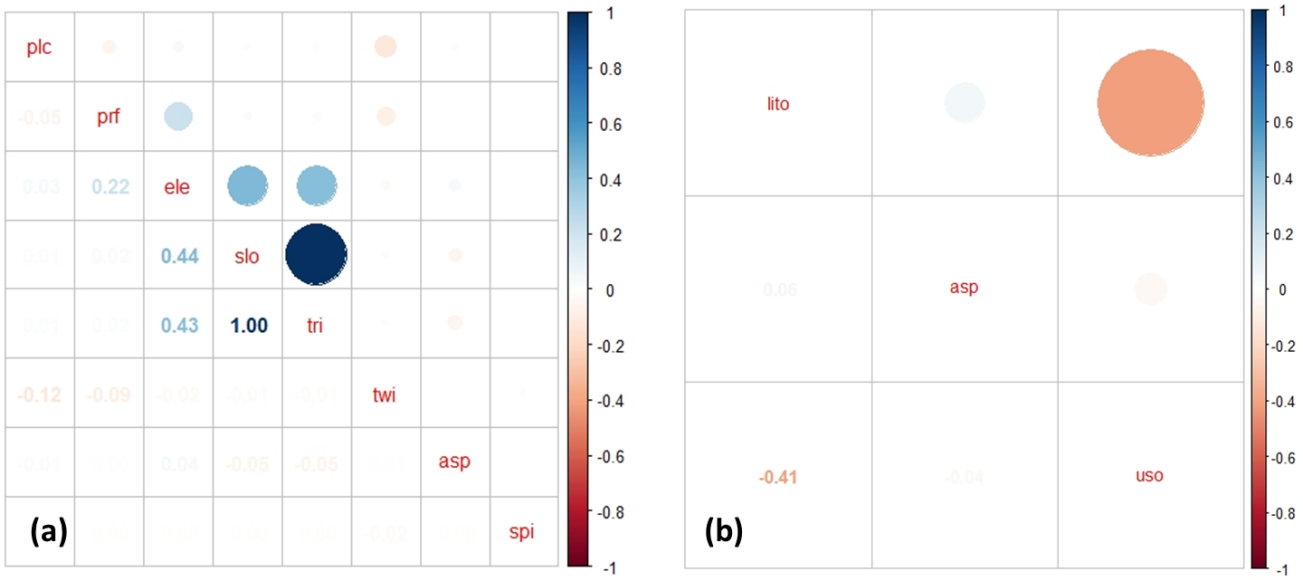

Figure 15. Multicollinearity analysis. (**a**) Continuous variables; (**b**) categorical variables.

Figure 16a presents the results of the confusion matrix calculation, indicating the performance metrics of the validation model. These metrics demonstrate a relatively high performance of the model in terms of prediction accuracy and precision. The precision score of 0.81 indicates the rate between true positives (TP) and all positive cases predicted by the model, including both true and false positives. Sensitivity (recall), which measures the model's ability to correctly predict the presence of landslides, is slightly higher than specificity, which measures the model's ability to predict absence cases correctly.

Figure 16a shows the ROC curve mean derived from 100 replicates, resulting in an area under the curve (AUC) of 0.85 with a low standard deviation of 0.09. The best probability cutoff point, determined using the Youden index, is 0.52.

Finally, the bar chart in Figure 16a shows the most important predictors using t-statistics and considering, for the 100 replicates, those variables with the most significant *p*-values. SSLO, PRF, and TWI are the most important DEM-derived morphometric parameters. Regarding the categorical variables, ASP classes corresponding to west (ASP.7), southwest (ASP.6), and south (ASP.5), and lithology classes representing 'Arenite, Siltstone' (15), 'Conglomerates, Arenite' (23), and 'Arenite, Mudstone' (13), are those ones which play the most important role in the classification of stable/unstable grid-cells.

The susceptibility map was generated by calculating the mean fitted probability values per pixel using data from the 100 replicates (Figure 16b). The map reveals that the central sector of the catchment and the steepness slopes of the valleys exhibit higher susceptibility values. However, it is worth noting that the bottom valleys in the same area show low susceptibility levels. The northern sector of the catchment displays low susceptibility values, while the southern sector exhibits varying levels of susceptibility, ranging from low to moderate.

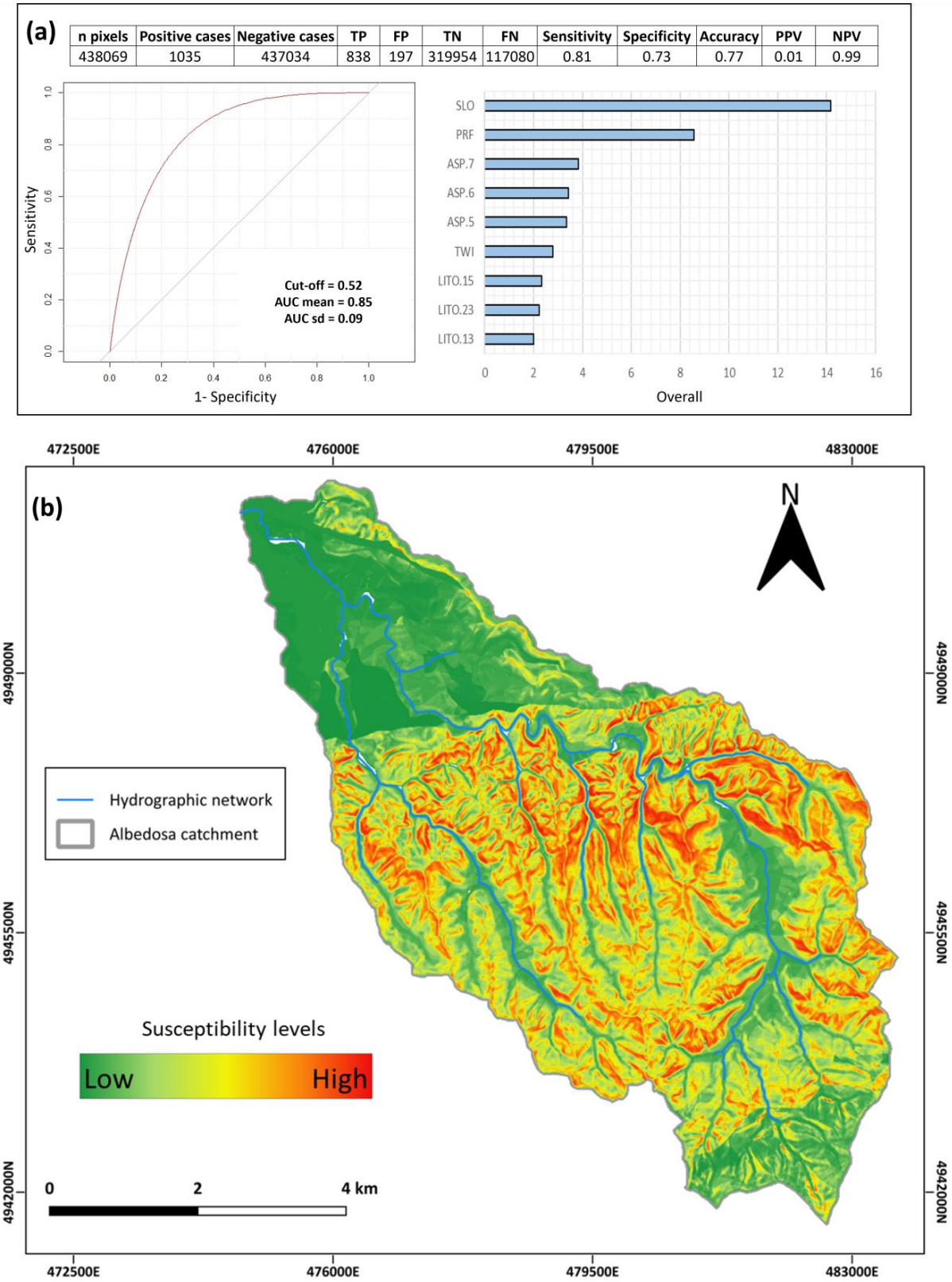

Figure 16. Landslide susceptibility model for the Albedosa river catchment performed with binary logistic regression algorithm and the OLP landslide dataset. (**a**) Landslide susceptibility model metrics: validation model, ROC curve, and principal predictors using t-statistics. (**b**) Landslide susceptibility map.

5. Discussion

The analysis of precipitation data through rainfall maps provided a comprehensive assessment of the spatial and temporal variations in rainfall across the study area.

By interpolating the data for P1 and P2, two grid maps of cumulative rainfall were generated. The map for P1 (Figure 8a) clearly indicates the Campo Ligure zone as the most affected by precipitation. In particular, the thunderstorm event on 21 October had a significant impact, with an affected area extending from the southwest to the northeast. The grid map clearly illustrates the development direction of the V-shaped thunderstorm. In contrast, the hydrographic basins located in the northwest recorded considerably lower rainfall accumulations, ranging from 100 to 200 mm.

For the P2 grid map (Figure 8b), the highest rainfall was observed in the proximity of the village of Urbe, situated southwest of the study area. Conversely, the northern regions reported lower rainfall amounts, approximately half of those in the southern regions. The precipitation data for P3 were presented as a contour map (Figure 8c), which confirmed significant differences in precipitation patterns between the northern and southern areas. Specifically, the Orba and Stura di Ovada river catchments were the most affected, with cumulative daily rainfall values ranging from 1100 to over 1900 mm over a 50-day period.

Overall, the interpolation of pluviometric data confirmed the results presented in Arpa Piemonte's reports, improving precision in local rainfall observations. The notable difference in rainfall from north to south within the study area can be mainly attributed to the terrain's orography, which features significant elevation variations of approximately 1000 m from north to south.

These results provide a solid basis for a better understanding of precipitation distribution and its impact on the territory, providing valuable information for further detailed studies on the trigger thresholds of rainfall-induced landslides.

The distribution of landslides in the area is found to be nonuniform, as shown in Figure 10. Notably, there is a marked difference in the orientation (slope aspect) of the landslides compared to the surrounding study area. The slope steepness data reveal a unimodal distribution, with the highest frequency of landslides occurring around a slope angle of approximately 21°. Additionally, an analysis of the elevation range indicates that landslides are more common in the elevation band between 200 and 400 m. It is worth noting that this distribution is not related to relief energy but rather to variations in lithology in the study area.

Interestingly, the frequency of landslides does not appear to be solely dependent on precipitation patterns. The analysis shown in Figure 11d reveals that the correlation between landslides and rainfall is not absolute. Other factors related to lithology and geological characteristics likely play a substantial role in landslide occurrence.

Further analysis of the lithology and landslide occurrence reveals interesting patterns. The frequency of landslides is significantly influenced by the underlying lithological composition. In the case of TPB (Figure 11b), the highest number of landslides occurs in lithologies characterized by alternating layers of permeable sandstones and impermeable mudstones or siltstones. Examples of lithological combinations associated with increased landslide occurrence include carbonate-rich mudstone and arenite, arenite and mudstone, and arenite and siltstone. In contrast, the landslides in the Alpine region show a different pattern (Figure 11c). Here, landslides are primarily concentrated in schist formations, which suggests a dominant influence of foliation and alteration processes. The lithology of serpentinites is also significantly influenced by landslides; probable causes are to be found in the foliate structure (such as serpentinite-shists and talc-rich formations) and in their vulnerability to hydrometeorological alteration.

According to Stark and Hovious [54], the landslide inventory compiled for the study area is considered comprehensive, and the Double Pareto distribution fits well with the observed data of the entire inventory (as shown in Figure 14a). The aspect-ratio-classified landslides also demonstrate a good distribution (Figure 14b). However, it should be noted

that transverse landslides are relatively less represented in the inventory, indicating their rarity within the area.

Regarding landslide parameters, a correlation analysis reveals strong linear relationships between length–area, width–area, and length–Δ elevation (as depicted in Figure 12a,b). These findings allow for the extrapolation of coefficients, suggesting the possibility of estimating landslide length and width based on the area and Δ elevation derived from the landslide polygons and digital elevation model (DEM). This approach differs from the current procedure, which involves minimum bounding box calculations and major and minor axis analyses.

The statistical model demonstrates a strong performance for the entire validation area, as evidenced by the calculated performance metrics, including the confusion matrix and the ROC curve (Figure 16a). Despite the lower Positive Predictive Values (PPV) and the generation of false positive cases, it is important to consider that these instances may represent potential future landslides rather than type I errors. This interpretation is supported by the high Negative Predictive Values (NPV), indicating successful predictions of absence cases.

The model performance is further reinforced by the metrics derived from the one-hundred replicates, including the AUC, the low standard deviation (SD), and the high accuracy (0.77) and precision (0.81). Additionally, both sensitivity and specificity values exhibit significantly good performance. These results confirm the model reliability and effectiveness in accurately predicting landslides.

The statistical model's ability to correctly classify landslide and nonlandslide areas is robust, and the high precision and accuracy values further emphasize its performance. These findings, in conjunction with the strong performance indicated by sensitivity, specificity, and the AUC, provide solid evidence of the model's high performance and reliability in the validation process.

The susceptibility map was successfully generated using the mean fitted probability values per pixel derived from the 100 replicates models. The map clearly depicts three distinct sectors within the catchment area, each characterized by different susceptibility values (Figure 16b). The northern sector of the catchment exhibits low susceptibility, which can be attributed to the presence of lithological characteristics such as anhydrites, conglomerates, and gravel–sand formations. The slope steepness in this area ranges from low to medium. In contrast, the middle sector shows high susceptibility, with varying slope steepness ranging from low to high. The lithological composition in this region consists of layers of arenites (sandstones) and mudstones/siltstones, contributing to the increased landslide susceptibility. The southern sector displays low to medium susceptibility levels, with medium to high slope steepness. The predominant lithology in this area is conglomerates. Overall, the susceptibility values observed in the northern and southern parts of the catchment are strongly influenced by the lithological characteristics, in addition to the slope steepness.

6. Conclusions

Landslides are a common occurrence in regions with temperate and equatorial climates, especially during periods of abundant and intense rainfall. These rainfall events have the potential to trigger numerous landslides, resulting in devastating consequences for the surrounding environment and human settlements. Therefore, it is crucial to assess landslide risks to ensure sustainable and resilient development in these areas.

To conduct accurate geostatistical analyses, large datasets of landslides are necessary. In order to contribute to this field, we have established the Open Landslide Project (OLP). Our approach focused on efficiently creating an inventory of recent shallow landslides, while minimizing the time required for mapping and covering large spatial areas and multiple years of data. To achieve this, we mapped over 760 km^2 in northwest Italy using freely accessible satellite imagery from the extreme rainfall events in autumn 2019. A total of 3300 landslides were mapped and classified according to Tian et al. [53],

based on aspect-ratio. The OLP inventory has successfully undergone statistical testing and validation.

Our analysis has revealed correlations between morphometric parameters of landslides. This finding opens up new possibilities for future research to expedite the classification of landslides based on aspect-ratio, a significant characteristic of landslides. Further investigations in this area could greatly enhance the efficiency of the landslide classification process.

Interestingly, when examining the relationship between rainfall patterns and lithologies where landslides occur, we found that the frequency of landslides is strongly influenced by lithology rather than solely being tied to precipitation levels. Surprisingly, areas with lower levels of rainfall, such as TPB, have recorded a higher number of landslides compared to regions with higher peak precipitation, like the Alpine region. This discrepancy cannot be attributed to morphometry factors such as slope steepness or elevation; instead, it is the lithological composition of the region that plays a significant role, as highlighted in the results of our analysis.

In addition to our landslide mapping efforts, we have developed a susceptibility model for the Albedosa river catchment, corresponding to the maximum peak of landslide occurrence. It was developed using binary logistic regression, and the resulting model achieved excellent metrics, particularly in terms of the statistical significance of the input data, indicating that the dataset used for the model is suitable for assessing landslide susceptibility.

Author Contributions: Conceptualization, M.L., V.B.T., F.S. and G.F.; methodology, M.L., V.B.T. and F.S.; software, M.L., V.B.T. and F.S.; validation, M.L., V.B.T., F.S. and G.F.; formal analysis, M.L., V.B.T. and F.S.; investigation, M.L., V.B.T. and F.S.; resources, M.L., V.B.T., F.S. and G.F.; data curation, M.L., V.B.T. and F.S.; writing—original draft preparation, M.L., V.B.T., F.S. and G.F.; writing—review and editing, M.L., V.B.T., F.S. and G.F.; visualization, M.L., V.B.T. and F.S.; supervision, M.L. and G.F.; project administration, G.F.; funding acquisition, G.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

Appendix A

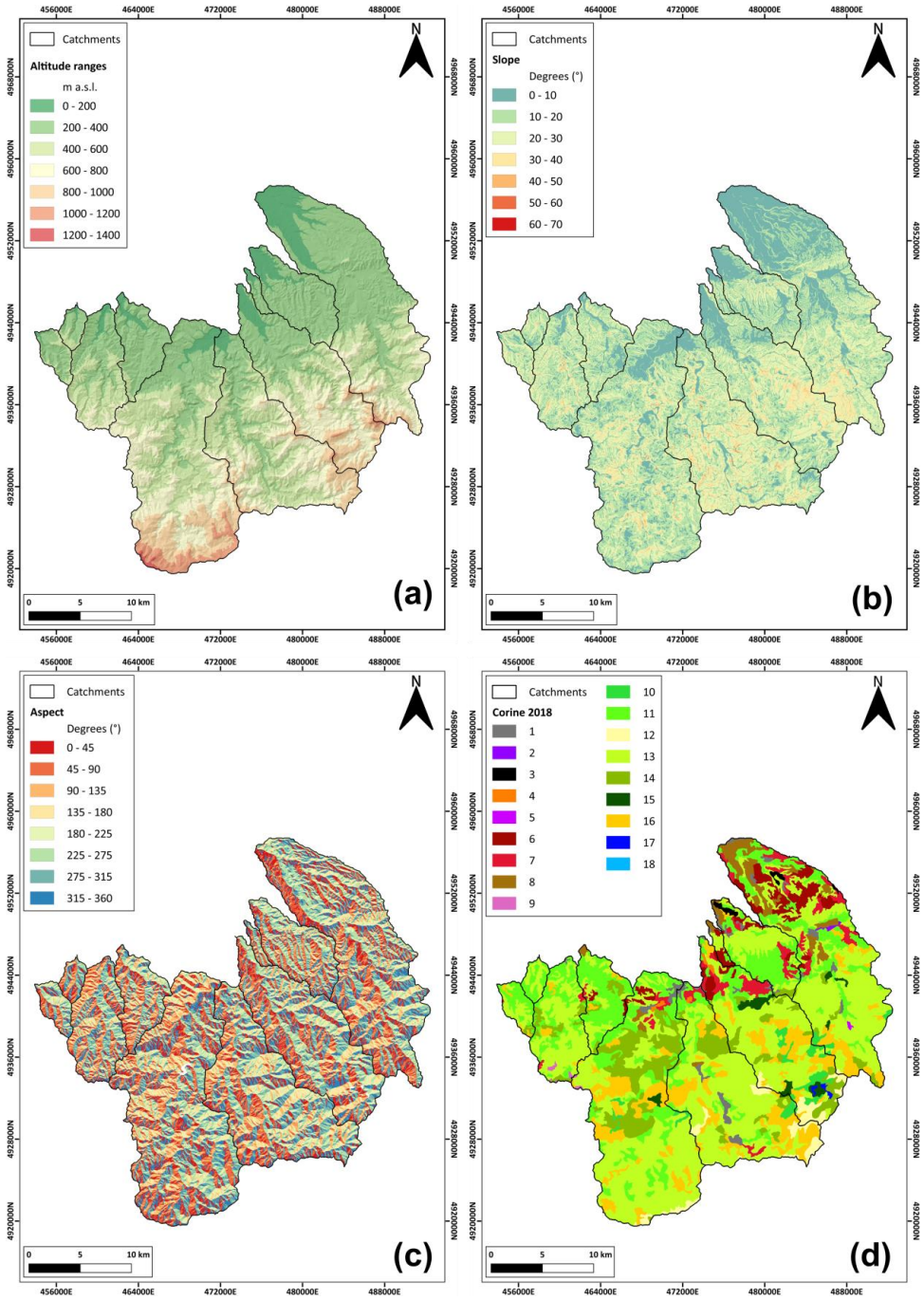

Figure A1. Morphometry and land cover of the studied area. (**a**) Elevation; (**b**) slope steepness; (**c**) slope aspect; (**d**) CORINE 2018: (1) discontinuous urban fabric; (2) road and rail networks and associated land; (3) sport and facilities; (4) industrial or commercial units; (5) mineral extraction sites; (6) vineyards; (7) complex cultivation patterns; (8) nonirrigated arable land; (9) pastures; (10) sparsely vegetated areas; (11) land principally occupied by agriculture, with significant areas of natural vegetation; (12) natural grassland; (13) broad-leaved forest; (14) mixed forest; (15) coniferous forest; (16) transitional woodland-shrub; (17) water bodies; (18) water courses.

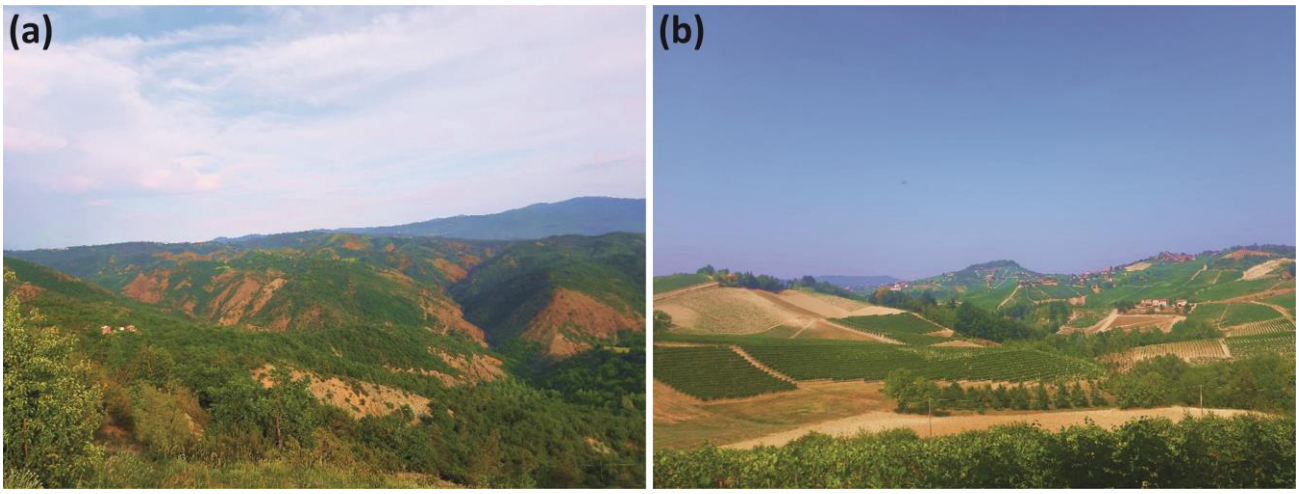

Figure A2. Photographs of the typical landscape of the study area. (**a**) Southern area and (**b**) northeast area.

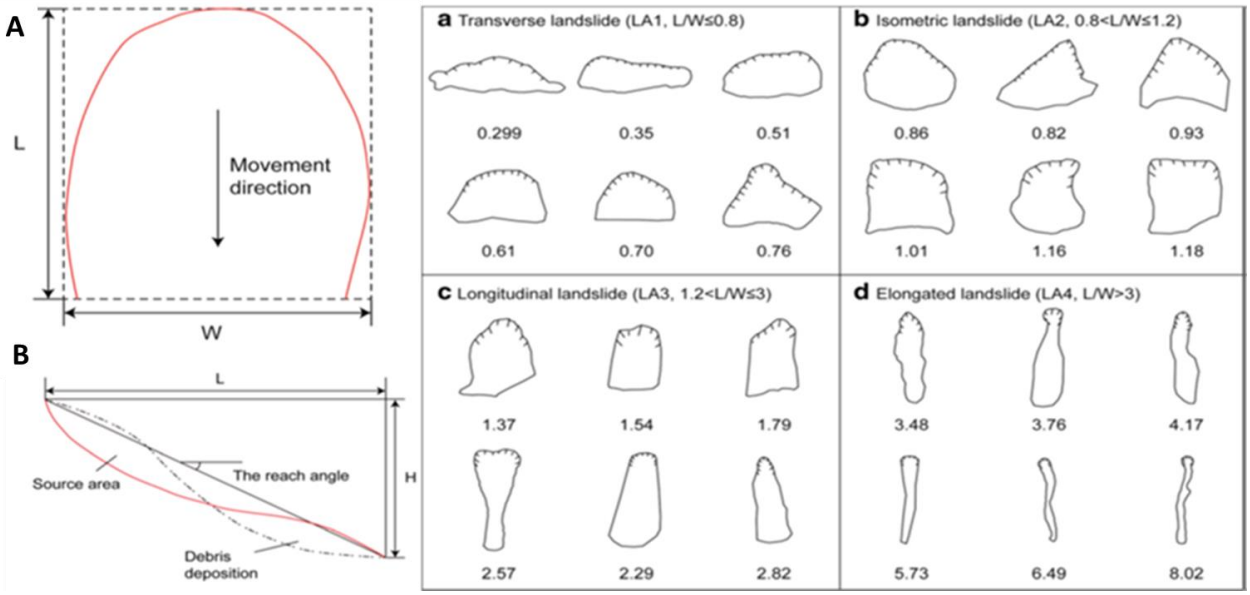

Figure A3. Landslide aspect-ratio classification. Modified after Tian et al. [53]. (**A**) Orthogonal view; (**B**) longitudinal view. The pictures (**a**–**c**) and (**d**) showed some orthogonal examples of the four aspect-ratio classified landslides.

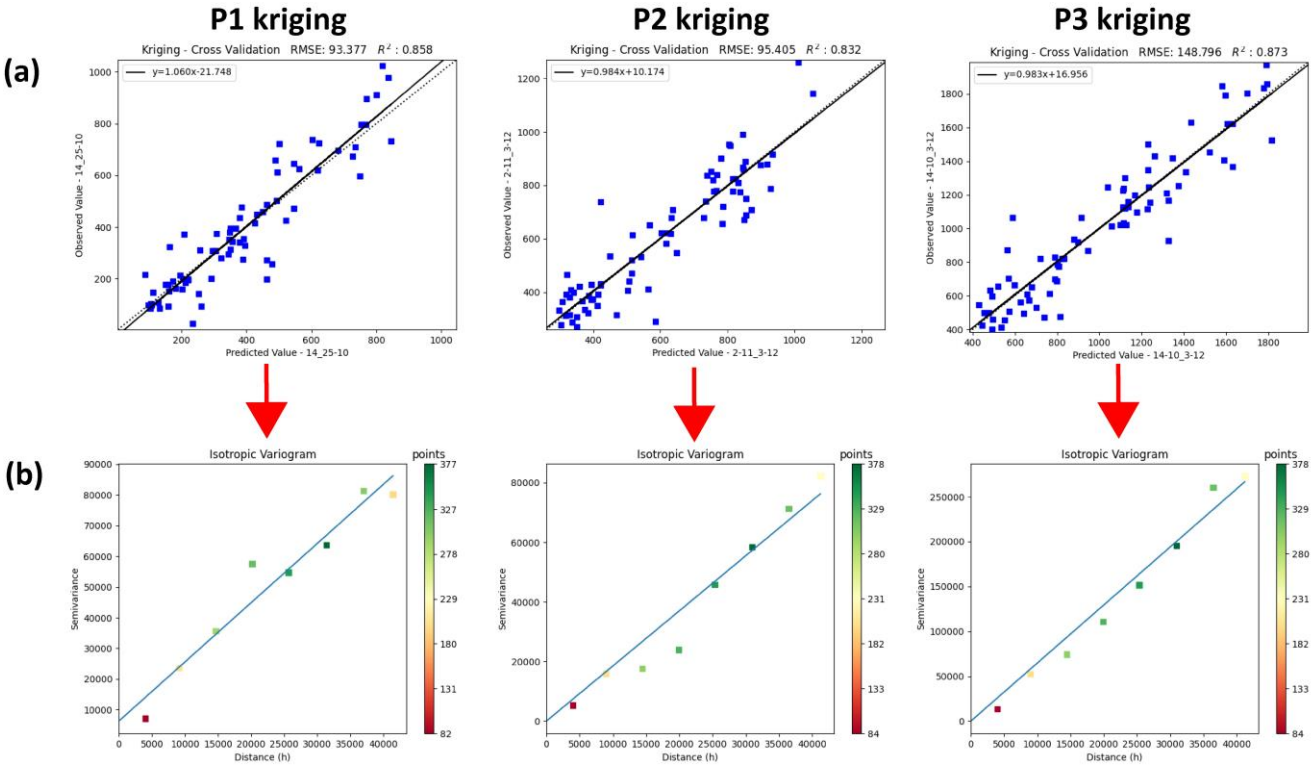

Figure A4. Kriging model metrics calculated by Smart-Map plugin. (**a**) Cross validation; (**b**) semivariograms.

References

1. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated World Map of the Köppen-Geiger Climate Classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [CrossRef]
2. Galli, M.; Ardizzone, F.; Cardinali, M.; Guzzetti, F.; Reichenbach, P. Comparing Landslide Inventory Maps. *Geomorphology* **2008**, *94*, 268–289. [CrossRef]
3. Gullà, G.; Conforti, M.; Borrelli, L. A Refinement Analysis of the Shallow Landslides Susceptibility at Regional Scale Supported by GIS-Aided Geo-Database. *Geomat. Nat. Hazards Risk* **2021**, *12*, 2500–2543. [CrossRef]
4. Arpa Piemonte—Evento 22, 25 Novembre 2019. Available online: http://www.arpa.piemonte.it/pubblicazioni-2/relazioni-tecniche/analisi-eventi/ (accessed on 18 July 2023).
5. Arpa Piemonte—Evento 19, 24 Ottobre 2019. Available online: http://www.arpa.piemonte.it/pubblicazioni-2/relazioni-tecniche/analisi-eventi/eventi2019/ (accessed on 18 July 2023).
6. Lombardo, L.; Cama, M.; Conoscenti, C.; Märker, M.; Rotigliano, E. Binary Logistic Regression versus Stochastic Gradient Boosted Decision Trees in Assessing Landslide Susceptibility for Multiple-Occurring Landslide Events: Application to the 2009 Storm Event in Messina (Sicily, Southern Italy). *Nat. Hazards* **2015**, *79*, 1621–1648. [CrossRef]
7. Felicísimo, Á.M.; Cuartero, A.; Remondo, J.; Quirós, E. Mapping Landslide Susceptibility with Logistic Regression, Multiple Adaptive Regression Splines, Classification and Regression Trees, and Maximum Entropy Methods: A Comparative Study. *Landslides* **2013**, *10*, 175–189. [CrossRef]
8. Martinello, C.; Cappadonia, C.; Rotigliano, E. Investigating the Effects of Cell Size in Statistical Landslide Susceptibility Modelling for Different Landslide Typologies: A Test in Central–Northern Sicily. *Appl. Sci.* **2023**, *13*, 1145. [CrossRef]
9. Harp, E.L.; Keefer, D.K.; Sato, H.P.; Yagi, H. Landslide Inventories: The Essential Part of Seismic Landslide Hazard Analyses. *Eng. Geol.* **2011**, *122*, 9–21. [CrossRef]
10. Taylor, F.E.; Malamud, B.D.; Witt, A.; Guzzetti, F. Landslide Shape, Ellipticity and Length-to-Width Ratios: What Shape Is a Landslide? *Earth Surf. Process. Landforms* **2018**, *43*, 3164–3189. [CrossRef]
11. Sahrane, R.; Bounab, A.; El Kharim, Y. Investigating the Effects of Landslides Inventory Completeness on Susceptibility Mapping and Frequency-Area Distributions: Case of Taounate Province, Northern Morocco. *CATENA* **2023**, *220*, 106737. [CrossRef]
12. Samodra, G.; Ngadisih, N.; Malawani, M.N.; Mardiatno, D.; Cahyadi, A.; Nugroho, F.S. Frequency–Magnitude of Landslides Affected by the 27–29 November 2017 Tropical Cyclone Cempaka in Pacitan, East Java. *J. Mt. Sci.* **2020**, *17*, 773–786. [CrossRef]
13. Schlögel, R.; Malet, J.-P.; Reichenbach, P.; Remaître, A.; Doubre, C. Analysis of a Landslide Multi-Date Inventory in a Complex Mountain Landscape: The Ubaye Valley Case Study. *Nat. Hazards Earth Syst. Sci.* **2015**, *15*, 2369–2389. [CrossRef]

14. Hurst, M.D.; Ellis, M.A.; Royse, K.R.; Lee, K.A.; Freeborough, K. Controls on the Magnitude-Frequency Scaling of an Inventory of Secular Landslides. *Earth Surf. Dynam.* **2013**, *1*, 67–78. [CrossRef]
15. Regmi, N.R.; Giardino, J.R.; Vitek, J.D. Characteristics of Landslides in Western Colorado, USA. *Landslides* **2014**, *11*, 589–603. [CrossRef]
16. Qiu, H.; Hu, S.; Yang, D.; He, Y.; Pei, Y.; Kamp, U. Comparing Landslide Size Probability Distribution at the Landscape Scale (Loess Plateau and the Qinba Mountains, Central China) Using Double Pareto and Inverse Gamma. *Bull. Eng. Geol Environ.* **2021**, *80*, 1035–1046. [CrossRef]
17. Bhuyan, K.; Tanyaş, H.; Nava, L.; Puliero, S.; Meena, S.R.; Floris, M.; Van Westen, C.; Catani, F. Generating Multi-Temporal Landslide Inventories through a General Deep Transfer Learning Strategy Using HR EO Data. *Sci. Rep.* **2023**, *13*, 162. [CrossRef] [PubMed]
18. Maki Mateso, J.-C.; Bielders, C.L.; Monsieurs, E.; Depicker, A.; Smets, B.; Tambala, T.; Bagalwa Mateso, L.; Dewitte, O. Characteristics and Causes of Natural and Human-Induced Landslides in a Tropical Mountainous Region: The Rift Flank West of Lake Kivu (Democratic Republic of the Congo). *Nat. Hazards Earth Syst. Sci.* **2023**, *23*, 643–666. [CrossRef]
19. Tanyaş, H.; Görüm, T.; Kirschbaum, D.; Lombardo, L. Could Road Constructions Be More Hazardous than an Earthquake in Terms of Mass Movement? *Nat. Hazards* **2022**, *112*, 639–663. [CrossRef]
20. Lombardo, L.; Mai, P.M. Presenting Logistic Regression-Based Landslide Susceptibility Results. *Eng. Geol.* **2018**, *244*, 14–24. [CrossRef]
21. Capponi, G.; Crispini, L. *Note Illustrative del Foglio 213 e 230 «Genova» Della Carta Geologica D'Italia Alla Scala 1:50.000.*; APAT—Servizio Geologico d'Italia: Firenze, Italy, 2008.
22. Luino, F. Sequence of Instability Processes Triggered by Heavy Rainfall in the Northern Italy. *Geomorphology* **2005**, *66*, 13–39. [CrossRef]
23. Ghibaudo, G.; Clari, P.; Perello, M. Litostratigrafia, Sedimentologia Ed Evoluzione Tettonico-Sedimentaria Dei Depositi Miocenici Del Margine Sud-Orientale Del Bacino Terziario Ligure-Piemontese (Valli Borbera, Scrivia e Lemme). *Boll. Della Soc. Geol. Ital.* **1985**, *104*, 349–397.
24. Piana, F.; Fioraso, G.; Irace, A.; Mosca, P.; d'Atri, A.; Barale, L.; Falletti, P.; Monegato, G.; Morelli, M.; Tallone, S.; et al. Geology of Piemonte Region (NW Italy, Alps–Apennines Interference Zone). *J. Maps* **2017**, *13*, 395–405. [CrossRef]
25. Federico, L.; Crispini, L.; Vigo, A.; Capponi, G. Unravelling Polyphase Brittle Tectonics through Multi-Software Fault-Slip Analysis: The Case of the Voltri Unit, Western Alps (Italy). *J. Struct. Geol.* **2014**, *68*, 175–193. [CrossRef]
26. Molli, G.; Crispini, L.; Malusà, M.G.; Mosca, P.; Piana, F.; Federico, L. Geology of the Western Alps-Northern Apennine Junction Area: A Regional Review. *J. Virtual Explor.* **2010**, *36*, 9. [CrossRef]
27. Ford, M.; Duchêne, S.; Gasquet, D.; Vanderhaeghe, O. Two-Phase Orogenic Convergence in the External and Internal SW Alps. *J. Geol. Soc.* **2006**, *163*, 815–826. [CrossRef]
28. Castellarin, A. Alps-Apennines and Po Plain-Frontal Apennines Relations. In *Anatomy of an Orogen: The Apennines and Adjacent Mediterranean Basins*; Vai, G.B., Martini, I.P., Eds.; Springer: Dordrecht, The Netherlands, 2001; pp. 177–195. ISBN 978-90-481-4020-6.
29. Bigi, G.; Castellarin, A.; Coli, M.; Dal Piaz, G.V.; Sortori, R.; Scandone, P.; Vai, G.B. *Structural Model of Italy, Sheets 1-2: CNR, Progetto Fin. Geodinamica*; SELCA: Firenze, Italy, 1990.
30. Balestro, G.; Festa, A.; Dilek, Y.; Tartarotti, P. Pre-Alpine Extensional Tectonics of a Peridotitelocalized Oceanic Core Complex in the Late Jurassic, High-Pressure Monviso Ophiolite (Western Alps). *Episodes* **2015**, *38*, 266–282. [CrossRef]
31. Mosca, P.; Polino, R.; Rogledi, S.; Rossi, M. New Data for the Kinematic Interpretation of the Alps–Apennines Junction (Northwestern Italy). *Int. J. Earth Sci.* **2010**, *99*, 833–849. [CrossRef]
32. Rossi, M.; Mosca, P.; Polino, R.; Rogledi, S.; Biffi, U. New Outcrop and Subsurface Data in the Tertiary Piedmont Basin (NW-Italy): Unconformity-Bounded Stratigraphic Units and Their Relationships with Basin-Modification Phases. *Riv. Ital. Paleontol. Stratigr.* **2009**, *115*, 305–335.
33. Gelati, R.; Gnaccolini, M. Synsedimentary Tectonics and Sedimentation in the Tertiary Piedmont Basin, Northwestern Italy. *Riv. Ital. Paleontol. Stratigr.* **1998**, *104*, 193–214.
34. d'Atri, A.; Irace, A.; Piana, F.; Tallone, S.; Varrone, D.; Bellino, L.; Fioraso, G.; Cadoppi, P.; Fusetti, E.; Morelli, M.; et al. *Note Illustrative Della Carta Geologica d'Italia Alla Scala 1:50,000, Foglio 194, Acqui Terme*; ISPRA, Istituto Superiore per La Protezione e La Ricerca Ambientale: Roma, Italy, 2014.
35. Baronetti, A.; Acquaotta, F.; Fratianni, S. Rainfall Variability from a Dense Rain Gauge Network in North-Western Italy. *Clim. Res.* **2018**, *75*, 201–213. [CrossRef]
36. Arpa Dati Meteoidrografici. Available online: https://www.arpa.piemonte.it/dati-ambientali/dati-meteoidrografici-giornalieri-richiesta-automatica (accessed on 18 July 2023).
37. Arpa Liguria Dati Meteoidrografici. Available online: https://omirl.regione.liguria.it/#/map (accessed on 18 July 2023).
38. Sistema Piemonte RAM. Available online: http://www.sistemapiemonte.it/fedwcodagt/centralineDB.jsp (accessed on 18 July 2023).
39. Acquaotta, F.; Fratianni, S.; Venema, V. Assessment of Parallel Precipitation Measurements Networks in Piedmont, Italy. *Int. J. Climatol.* **2016**, *36*, 3963–3974. [CrossRef]
40. Hattermann, F.F.; Wattenbach, M.; Krysanova, V.; Wechsung, F. Runoff Simulations on the Macroscale with the Ecohydrological Model SWIM in the Elbe Catchment-Validation and Uncertainty Analysis. *Hydrol. Process.* **2005**, *19*, 693–714. [CrossRef]

41. Buytaert, W.; Celleri, R.; Willems, P.; Bièvre, B.D.; Wyseure, G. Spatial and Temporal Rainfall Variability in Mountainous Areas: A Case Study from the South Ecuadorian Andes. *J. Hydrol.* **2006**, *329*, 413–421. [CrossRef]
42. Berezowski, T.; Szcześniak, M.; Kardel, I.; Michałowski, R.; Okruszko, T.; Mezghani, A.; Piniewski, M. CPLFD-GDPT5: High-Resolution Gridded Daily Precipitation and Temperature Data Set for Two Largest Polish River Basins. *Earth Syst. Sci. Data* **2016**, *8*, 127–139. [CrossRef]
43. Brinckmann, S.; Krähenmann, S.; Bissolli, P. High-Resolution Daily Gridded Data Sets of Air Temperature and Wind Speed for Europe. *Earth Syst. Sci. Data* **2016**, *8*, 491–516. [CrossRef]
44. Frazier, A.G.; Giambelluca, T.W. Spatial Trend Analysis of Hawaiian Rainfall from 1920 to 2012. *Int. J. Climatol.* **2017**, *37*, 2522–2531. [CrossRef]
45. Lucas, M.P.; Longman, R.J.; Giambelluca, T.W.; Frazier, A.G.; Mclean, J.; Cleveland, S.B.; Huang, Y.-F.; Lee, J. Optimizing Automated Kriging to Improve Spatial Interpolation of Monthly Rainfall over Complex Terrain. *J. Hydrometeorol.* **2022**, *23*, 561–572. [CrossRef]
46. Pereira, G.W.; Valente, D.S.M.; Queiroz, D.M.D.; Coelho, A.L.D.F.; Costa, M.M.; Grift, T. Smart-Map: An Open-Source QGIS Plugin for Digital Mapping Using Machine Learning Techniques and Ordinary Kriging. *Agronomy* **2022**, *12*, 1350. [CrossRef]
47. Isaaks, E.H.; Srivastava, R.M. *An Introduction to Applied Geostatistics*; Oxford University Press: New York, NY, USA, 1989; ISBN 0-19-505012-6.
48. Shuttle Radar Topography Mission. Available online: https://www2.jpl.nasa.gov/srtm/ (accessed on 18 July 2023).
49. Sentinel-2 Mission. Available online: https://sentinel.esa.int/web/sentinel/missions/sentinel-2 (accessed on 18 July 2023).
50. Cruden, D.M.; Varnes, D.J. *Landslide Types and Processes*; Transportation Research Board; U.S. National Academy of Sciences: Washington, DC, USA, 1996; pp. 36–75.
51. Li, L.; Lan, H.; Strom, A.; Macciotta, R. Landslide Length, Width, and Aspect Ratio: Path-Dependent Measurement and a Revisit of Nomenclature. *Landslides* **2022**, *19*, 3009–3029. [CrossRef]
52. IAEG Commission on Landslides. Suggested Nomenclature for Landslides. *Bull. Int. Assoc. Eng. Geol.* **1990**, *41*, 13–16. [CrossRef]
53. Tian, Y.; Xu, C.; Chen, J.; Zhou, Q.; Shen, L. Geometrical Characteristics of Earthquake-Induced Landslides and Correlations with Control Factors: A Case Study of the 2013 Minxian, Gansu, China, Mw 5.9 Event. *Landslides* **2017**, *14*, 1915–1927. [CrossRef]
54. Stark, C.P.; Hovius, N. The Characterization of Landslide Size Distribution. *Geophys. Res. Lett.* **2001**, *28*, 1091–1094. [CrossRef]
55. Tebbens, S.F. Landslide Scaling: A Review. *Earth Space Sci.* **2020**, *7*, e2019EA000662. [CrossRef]
56. Van Den Eeckhaut, M.; Reichenbach, P.; Guzzetti, F.; Rossi, M.; Poesen, J. Combined Landslide Inventory and Susceptibility Assessment Based on Different Mapping Units: An Example from the Flemish Ardennes, Belgium. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 507–521. [CrossRef]
57. Erener, A.; Düzgün, H.S.B. Landslide Susceptibility Assessment: What Are the Effects of Mapping Unit and Mapping Method? *Environ. Earth Sci.* **2012**, *66*, 859–877. [CrossRef]
58. Rotigliano, E.; Cappadonia, C.; Conoscenti, C.; Costanzo, D.; Agnesi, V. Slope Units-Based Flow Susceptibility Model: Using Validation Tests to Select Controlling Factors. *Nat. Hazards* **2012**, *61*, 143–153. [CrossRef]
59. Ba, Q.; Chen, Y.; Deng, S.; Yang, J.; Li, H. A Comparison of Slope Units and Grid Cells as Mapping Units for Landslide Susceptibility Assessment. *Earth Sci. Inform.* **2018**, *11*, 373–388. [CrossRef]
60. Bornaetxea, T.; Rossi, M.; Marchesini, I.; Alvioli, M. Effective Surveyed Area and Its Role in Statistical Landslide Susceptibility Assessments. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 2455–2469. [CrossRef]
61. Domènech, G.; Alvioli, M.; Corominas, J. Preparing First-Time Slope Failures Hazard Maps: From Pixel-Based to Slope Unit-Based. *Landslides* **2020**, *17*, 249–265. [CrossRef]
62. Martinello, C.; Cappadonia, C.; Conoscenti, C.; Agnesi, V.; Rotigliano, E. Optimal Slope Units Partitioning in Landslide Susceptibility Mapping. *J. Maps* **2021**, *17*, 152–162. [CrossRef]
63. Carrara, A.; Cardinali, M.; Guzzetti, F.; Reichenbach, P. Gis Technology in Mapping Landslide Hazard. In *Geographical Information Systems in Assessing Natural Hazards*; Carrara, A., Guzzetti, F., Eds.; Advances in Natural and Technological Hazards Research; Springer: Dordrecht, The Netherlands, 1995; Volume 5, pp. 135–175. ISBN 978-90-481-4561-4.
64. Guzzetti, F.; Reichenbach, P.; Cardinali, M.; Galli, M.; Ardizzone, F. Probabilistic Landslide Hazard Assessment at the Basin Scale. *Geomorphology* **2005**, *72*, 272–299. [CrossRef]
65. Costanzo, D.; Rotigliano, E.; Irigaray, C.; Jiménez-Perálvarez, J.D.; Chacón, J. Factors Selection in Landslide Susceptibility Modelling on Large Scale Following the Gis Matrix Method: Application to the River Beiro Basin (Spain). *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 327–340. [CrossRef]
66. Amato, G.; Eisank, C.; Castro-Camilo, D.; Lombardo, L. Accounting for Covariate Distributions in Slope-Unit-Based Landslide Susceptibility Models. A Case Study in the Alpine Environment. *Eng. Geol.* **2019**, *260*, 105237. [CrossRef]
67. Tarquini, S.; Vinci, S.; Favalli, M.; Doumaz, F.; Fornaciai, A.; Nannipieri, L. Release of a 10-m-Resolution DEM for the Italian Territory: Comparison with Global-Coverage DEMs and Anaglyph-Mode Exploration via the Web. *Comput. Geosci.* **2012**, *38*, 168–170. [CrossRef]
68. Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L.; Wehberg, J.; Wichmann, V.; Böhner, J. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* **2015**, *8*, 1991–2007. [CrossRef]
69. Lombardo, L.; Cama, M.; Maerker, M.; Rotigliano, E. A Test of Transferability for Landslides Susceptibility Models under Extreme Climatic Events: Application to the Messina 2009 Disaster. *Nat. Hazards* **2014**, *74*, 1951–1989. [CrossRef]

70. Esposito, C.; Mastrantoni, G.; Marmoni, G.M.; Antonielli, B.; Caprari, P.; Pica, A.; Schilirò, L.; Mazzanti, P.; Bozzano, F. From Theory to Practice: Optimisation of Available Information for Landslide Hazard Assessment in Rome Relying on Official, Fragmented Data Sources. *Landslides* **2023**, *20*, 2055–2073. [CrossRef]
71. Naimi, B. R-GIS. Available online: https://r-gis.net/ (accessed on 18 July 2023).
72. Zuur, A.F.; Ieno, E.N.; Walker, N.J.; Saveliev, A.A.; Smith, G.M. GLM and GAM for Absence–Presence and Proportional Data. In *Mixed Effects Models and Extensions in Ecology with R*; Statistics for Biology and Health; Springer New York: New York, NY, USA, 2009; pp. 245–259. ISBN 978-0-387-87457-9.
73. Conoscenti, C.; Ciaccio, M.; Caraballo-Arias, N.A.; Gómez-Gutiérrez, Á.; Rotigliano, E.; Agnesi, V. Assessment of Susceptibility to Earth-Flow Landslide Using Logistic Regression and Multivariate Adaptive Regression Splines: A Case of the Belice River Basin (Western Sicily, Italy). *Geomorphology* **2015**, *242*, 49–64. [CrossRef]
74. 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]
75. 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]
76. Hosmer, D.; Lemeshow, S. Stepwise Logistic Regression. Applied Logistic Regression. In *Chapter 4, Model-building Strategies and Methods for Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013; pp. 116–121.
77. Hartmann, K.; Krois, J.; Rudolph, A. Statistics and Geodata Analysis Using R (SOGA-R); Department of Earth Sciences, Freie Universitaet Berlin. Available online: https://www.geo.fu-berlin.de/en/v/soga-r/index.html (accessed on 18 July 2023).
78. R Core Team. Available online: http://www.r-project.org/index.html (accessed on 18 July 2023).
79. Youden, W.J. Index for Rating Diagnostic Tests. *Cancer* **1950**, *3*, 32–35. [CrossRef]