sustainability

MDPI

Article

A Spatial Landslide Risk Assessment Based on Hazard, Vulnerability, Exposure, and Adaptive Capacity

Thong Xuan Tran [1,2,*], Sihong Liu [1], Hang Ha [3], Quynh Duy Bui [3], Long Quoc Nguyen [4], Dinh Quoc Nguyen [5], Cong-Ty Trinh [2] and Chinh Luu [2]

[1] College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China; sihongliu@hhu.edu.cn
[2] Faculty of Hydraulic Engineering, Hanoi University of Civil Engineering, Hanoi 11616, Vietnam; tytc@huce.edu.vn (C.-T.T.); chinhltd@huce.edu.vn (C.L.)
[3] Department of Geodesy and Geomatic Engineering, Hanoi University of Civil Engineering, Hanoi 11616, Vietnam; hanght@huce.edu.vn (H.H.); quynhbd@huce.edu.vn (Q.D.B.)
[4] Department of Mine Surveying, Hanoi University of Mining and Geology, Hanoi 11909, Vietnam; nguyenquoclong@humg.edu.vn
[5] Environmental Chemistry and Ecotoxicology Lab, Phenikaa University, Hanoi 12116, Vietnam; dinh.nguyenquoc@phenikaa-uni.edu.vn
* Correspondence: thongtx@huce.edu.vn; Tel.: +84-983036368

Abstract: Landslides threaten human life, property, and vital infrastructure in most mountainous regions. As climate change intensifies extreme weather patterns, the landslide risk is likely to increase, resulting in challenges for disaster management, sustainability development, and community resilience. This study presents a comprehensive framework for assessing landslide risk, integrating advanced machine learning models with the Iyengar–Sudarshan method. Our case study is Son La province, the Northwest region of Vietnam, with data collected from 1771 historical landslide occurrences and fifteen influencing factors for developing landslide susceptibility maps using advanced ensemble machine learning models. The Iyengar–Sudarshan method was applied to determine the weights for landslide exposure, vulnerability, and adaptive capacity indicators. The resulting landslide risk map shows that the highest-risk districts in Son La province are located in the central and northeastern regions, including Mai Son, Phu Yen, Thuan Chau, Yen Chau, Song Ma, and Bac Yen. These districts experience high landslide hazards, exposure, and vulnerability, often affecting densely populated urban and village areas with vulnerable populations, such as young children, the elderly, and working-age women. In contrast, due to minimal exposure, Quynh Nhai and Muong La districts have lower landslide risks. Despite having high exposure and vulnerability, Son La City is situated in a low-susceptibility zone with high adaptive capacity, resulting in a low landslide risk for this region. The proposed framework provides a reference tool for mitigating risk and enhancing strategic decision making in areas susceptible to landslides while advancing our understanding of landslide dynamics and fostering community resilience and long-term disaster prevention.

Keywords: landslide risk assessment; hazard; exposure; vulnerability; adaptive capacity; machine learning; Iyengar–Sudarshan method; Son La province

Citation: Tran, T.X.; Liu, S.; Ha, H.; Bui, Q.D.; Nguyen, L.Q.; Nguyen, D.Q.; Trinh, C.-T.; Luu, C. A Spatial Landslide Risk Assessment Based on Hazard, Vulnerability, Exposure, and Adaptive Capacity. *Sustainability* **2024**, *16*, 9574. https://doi.org/10.3390/su16219574

Received: 16 September 2024
Revised: 23 October 2024
Accepted: 24 October 2024
Published: 3 November 2024

1. Introduction

Landslides are a common and significant natural hazard in many nations of Asia due to these regions' diverse geography, climate, and geological characteristics [1]. A total of 55,997 people were killed in 4862 landslide events in the period of 2004–2016 [2]. The distribution of these landslide events differs, but Asia is the region that takes the majority worldwide [2,3], which is attributed to their complex landscapes and climatic conditions [4,5]. Landslide events can cause significant damage to the environment, properties, transportation networks, infrastructures, and human lives [6]. Landslides pose substantial

dangers in mountainous areas due to various triggering factors of steep terrain, complex geological conditions, extreme climate patterns, and human activities [7,8]. Landslides can vary from small scale to large scale, depending on the cubic volume of the sliding block and the extent of the damage they cause [9]. Still, all have destructive characteristics affected by dynamic factors such as land use changes, urbanization, population growth, and climate change status [10]. These factors evolve, complicating the risk assessment process and requiring continuous updates to models and data [11].

Machine learning (ML) techniques are significant in landslide-related studies because they address complex and large-scale datasets, identify samples, and make accurate forecasts [12]. Landslide risk assessment often involves various data, including topography, geology, climate, environment, society, economy, and physical structures [13,14]. ML algorithms can efficiently process and analyze these multidimensional datasets, identifying intricate patterns and correlations within datasets [15]. In recent years, advanced ML models, such as hybrid ML models and deep learning, can capture intricate interactions among multiple variables, leading to more accurate and robust landslide risk assessments [16,17].

The UltraBoost (UB) algorithm represents an innovative advancement in ML, introducing a range of optimizations that enhance the efficiency and accuracy of predictive models based on the foundation of previously established boosting algorithms [18]. This algorithm employs advanced gradient optimization techniques that allow faster convergence during training, significantly reducing the computational time required to achieve high-performance models [19]. These make UltraBoost particularly suitable for large-scale datasets where traditional boosting algorithms might struggle with processing speed [20]. This was combined with the Weights of Evidence model to assess flood susceptibility in the Putna river basin from Romania [21], or it was employed to determine a safe construction location in the Mediterranean region during pre-earthquake disasters [18]. However, the UB model has not yet been applied in landslide susceptibility mapping.

The increasing complexity and unpredictability of landslide hazards due to climate change underscore the urgent need for advanced risk assessment methods. As a result, several studies have explored landslide risk assessment by applying various ML models. Novellino et al. [12] utilized three ML models, including Artificial Neural Networks, the Generalized Boosting Model, and Maximum Entropy, to build the landslide hazard map for the Termini–Nerano area in southern Apennines, Italy. Then, this landslide hazard map was combined with the official population and building census data to estimate the landslide risk for this region. Mallick et al. [22] developed four ensemble metaheuristic ML algorithms for modelling rainfall-induced landslide susceptibility in Aqabat Al-Sulbat, Asir region, Saudi Arabia. Landslide hazard maps were created by integrating the best susceptibility model with the estimated rainfall to assess the risk to resources exposed to landslides. Wen et al. [23] merged the Recursive Feature Elimination method and Particle Swarm Optimization–AdaBoost hybrid model to create a landslide susceptibility map for the Changshou–Fuling–Wulong–Nanchuan gas pipeline in China. The fuzzy clustering (FC) and the CRITIC method (FC-CRITIC) were combined to segment pipelines and develop a pipeline vulnerability model. This study produced a pipeline risk map combining pipeline vulnerability and landslide susceptibility. Despite significant progress in estimating landslide hazards over the past few decades, there remains a critical gap in evaluating holistic landslide risks due to the lack of historical damage data and the complexity of vulnerability and exposure elements [24].

The Iyengar–Sudarshan technique is a popular weighting method in climate change risk assessment [25–28], which allows for flexibility in determining the weights of different indicators depending on the complexity and uniqueness of assessment situations [29]. This technique makes the results interpretation easier because the calculated weights of indicators provide clear insights into the relative significance of each indicator [30,31]. Therefore, this method was used to assess various disasters of flood risk [32,33], drought hazard [34], and water quantity risk [35]. Thus, the Iyengar–Sudarshan method has the potential to be applied in landslide risk assessment studies.

Vietnam ranks among the sixth nations frequently impacted by natural hazards, including landslide events [36]. The Son La province in North Vietnam is prone to landslides [37]. This study aimed to propose a comprehensive framework that combines advanced ML models with the Iyengar–Sudarshan method for assessing landslide risk in Son La province.

2. Study Area and Theoretical Framework

2.1. Description of the Study Area

Son La is the fifth largest province in the Northwest region of Vietnam. Its topography fully describes the geographical traits of the North region of Vietnam, with many mountains, hills, rivers, and mineral sources surrounded by primitive forests [38]. It borders Yen Bai, Lao Cai, and Lai Chau provinces in the North, Dien Bien Province in the West, Phu Tho and Hoa Binh in the East, and Laos in the South. The geographical boundaries of this province are defined by the latitudes 20°39 and 22°02 N and the longitudes 103°11′ and 105°02′ E. It has 12 administrative units, including Son La City and 11 districts (Figure 1). The natural area of Son La province is 14,125 km^2. The province's population was about 1,248,415 people in 2021, of which people in urban areas account for 20.8% and people in rural areas account for 79.2%. The population density of the whole province reaches 88 people/km^2. Son La province is also home to many ethnic groups, Thai, Kinh, Mong, Muong, Tay, and Nung, of which the Thai people make the majority (available at https://sonla.gov.vn/, accessed on 1 October 2024).

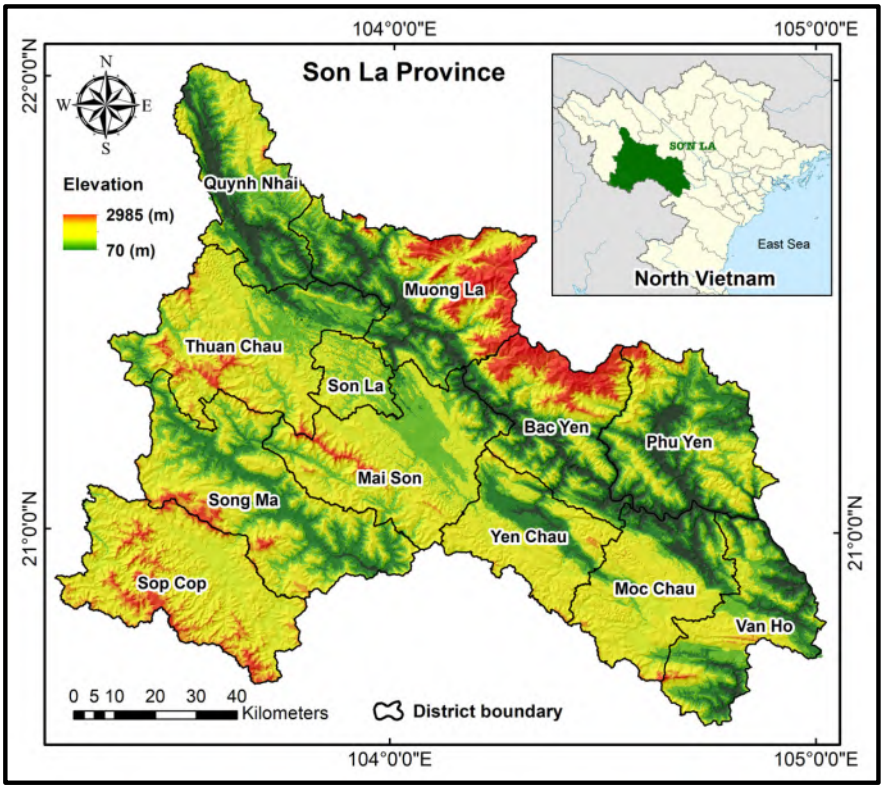

Figure 1. The research area.

Extreme weather, prolonged heavy rainfall, human activities, and complex geological conditions are the key reasons causing landslide formation in the study area [38]. The average annual rainfall ranges from about 1400 to 1700 mm/year and is heavily concentrated in June, July, and August (available at https://sonla.gov.vn/, accessed on 1 October 2024). Many factors, such as topography, geology, land cover, weathering crust thickness, soil type, and soil thickness, also contribute to the landslide occurrences in the study area [38]. In recent years, Son La has experienced numerous landslides that have damaged roads, bridges, and agricultural lands, disrupting transportation and impacting livelihoods. Heavy rain

from 24 to 25 August 2021 caused flash floods and landslides across the region. A total of 656 people were evacuated, 35 houses were damaged, 82.58 hectares of agricultural land were affected, many roads were destroyed, and over 1700 m of canals were broken [39]. In the latest event, typhoon Yagi caused widespread devastation in provinces in North Vietnam, including Son La, resulting in 318 deaths, 1976 injuries, and the evacuation of over 130,000 people, with damages exceeding VND 81,000 billion, severely impacting homes, agriculture, infrastructure, and essential services like transportation, communication, and water supply [40]. Landslides pose a recurring and formidable threat to this province, highlighting the urgent need for a comprehensive landslide risk assessment that integrates evaluations of hazards, exposure, vulnerability, and adaptive capacity into a holistic management strategy (Figure 2).

Figure 2. Some images of landslide events on Highway 6 in the Son La province in July and August 2018 through field surveys: (**a**) At Km 201 + 00; (**b**) At Km 203 + 500; (**c**) At Km 205 + 100; (**d**) At Km 206 + 800; and (**e**) At Km 230 + 300 (provided by Road Joint Stock Company 226, located in Son La province, Vietnam).

2.2. Theoretical Framework

The "Risk" concept can be utilized in various ways due to the objective, approaches, or reference system [41]. According to the most comprehensive method, the "Risk" concept was defined as the expected value of losses of human life and property and the disruption of social and economic activities caused by a 'natural' disaster within a specific area over a given period [2,42]. In this sense, risk assessment emphasizes understanding disaster risk in all dimensions, including hazard characteristics, exposure, vulnerability, and adaptive capacity [41]. Integrating these elements provides a holistic view of risk by considering the likelihood and intensity of hazardous events, the exposure of assets and populations, their vulnerabilities, and their ability to adapt and recover [43]. This study focuses on a landslide risk-centred assessment framework, where landslide risk is expressed as a function of landslide hazard, exposure, vulnerability, and adaptive capacity [44,45], as illustrated by the following equation:

$$Risk = \frac{Hazard \times Exposure \times Vulnerability}{Adaptive\ Capacity} \quad (1)$$

In this study, "Hazard" refers to the potential occurrence of landslide events and is used to evaluate the spatial distribution of landslides without considering the temporal aspect. By focusing solely on the spatial dimension, the analysis of landslide hazards is directly described in landslide susceptibility modelling [46]. "Exposure" reflects the degree to which human lives, properties, infrastructure, and the environment are at risk in landslide-prone areas [47]. The human, infrastructure, and livelihood often represent areas where human lives and properties are exposed to landslide events, so they are frequently considered crucial criteria for landslide exposure [6]. "Vulnerability" reflects the susceptibility of the population and structures to damage when a landslide occurs [11]. "Adaptive Capacity" demonstrates the ability of communities and systems to prepare for, respond to, and recover from landslide events [48]. When evaluating landslide risk, selecting the right indicators is crucial for a comprehensive and insightful assessment. Seventeen indicators were collected and sorted into three categories to assess landslide risk: exposure, vulnerability, and adaptive c. The indicators representing these mentioned components may be chosen based on the availability of data and insights from previous related studies [49].

3. Data and Methods Used for Landslide Susceptibility Modelling

3.1. Data for Landslide Susceptibility Modelling

3.1.1. Landslide Inventory

Landslide susceptibility modelling significantly depends on past landslides' spatial distribution [38]. Thus, the landslide inventory map is a crucial input parameter in landslide susceptibility mapping as it provides a comprehensive database of historical landslide events [50,51]. This landslide inventory map had 1771 landslide sites gathered from various sources. Among these, 1225 landslide positions were investigated using data collected from the Institute of Geosciences and Minerals of Vietnam, while an additional 546 landslide locations were identified through a combination of field surveys in 2022, 2023 and the analysis of Google Earth imagery (Figure 3).

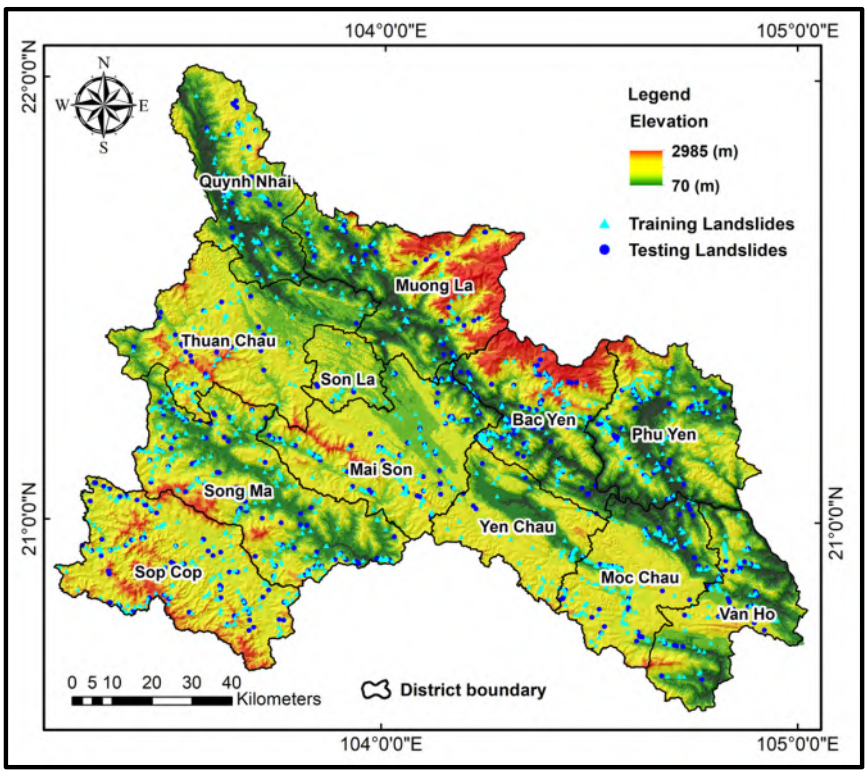

Figure 3. The landslide inventory locations.

Landslide inventory data were randomly separated into two datasets using the Sample tool in ArcGIS Pro 3.1. Specifically, 1240 landslide sites corresponding to the 70% training dataset were employed to build landslide susceptibility maps, and 531 landslide sites belonging to the 30% testing dataset were applied to verify the forecast performance of landslide susceptibility maps.

3.1.2. Landslide Conditioning Factors

Landslides are often influenced by various factors that contribute to their occurrence and severity [52]. Landslides have the potential to occur in any area characterized by unfavourable conditions related to topography, geology, hydrology, and the environment [53,54]. The gathering and identification of factors influencing landslides may rely on the availability of data and previous studies in the same study area [55]. Consequently, fifteen landslide influencing factors were chosen, consisting of altitude, slope, aspect, curvature, terrain roughness, Topographic Wetness Index (TWI), rainfall, drainage density, road density, distance from the road, distance from the rivers, hydrogeology, geology, geomorphology, and land cover.

The topographic variables, including altitude, slope angle, aspect, curvature, terrain roughness, and TWI, were derived from the ALOS DEM with a spatial resolution of 30 m. The DEM dataset was acquired in March 2021 from https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30_e.htm, accessed on 1 February 2024. The rainfall data were collected from 25 rain gauge stations in Son La province and the surrounding areas between 2012 and 2022. The hydrogeology, geology, and geomorphology maps were obtained from the Vietnamese Ministry of Natural Resources and Environment at the 1:100,000 scale in 2020. The drainage density map was collected from the provincial Department of Natural Resources and Environment in 2020. The land cover was obtained from https://livingatlas.arcgis.com/landcover/, accessed on 1 February 2024. The road network map of the Son La province was derived from the national road network map, which was supplied by the Department of Survey, Mapping, and Geographic Information in 2020.

In this study, fifteen landslide-affecting factors were collected and grouped into various categories based on their availability and contribution to landslide events. The topographic factors (altitude, slope angle, aspect, curvature, terrain roughness, and TWI) were classified into criteria based on various terrain morphology within the study area. The hydrogeology, geology, and geomorphology data were classified into criteria based on field surveys on underground water, rock formation, weathering degree, and rock stability. Other factors were determined and classified based on characteristics, structure, and spatial distribution.

Additionally, the frequency ratio analysis was utilized to verify this remark by determining the correlation between the historical landslide locations and the influencing factors. The frequency ratio is obtained by taking the number of landslide positions within a specific class, dividing it by the total number of landslide locations, and comparing it with the ratio of pixels in that class to the overall pixel count [38]. The high-frequency ratio values indicate a strong correlation between the input factors and the landslide formation in the area and vice versa. The outcome of the analysis received demonstrates the correlation between landslide causative factors and historical landslide events, resulting in the classification suitable for all input factors (Figure 4).

All landslide causative factors are transformed into the raster format at a spatial resolution of 30 m using the ArcGIS Pro environment. The comprehensive details of the fifteen landslide conditioning factors are described in Table 1 and Figure 5.

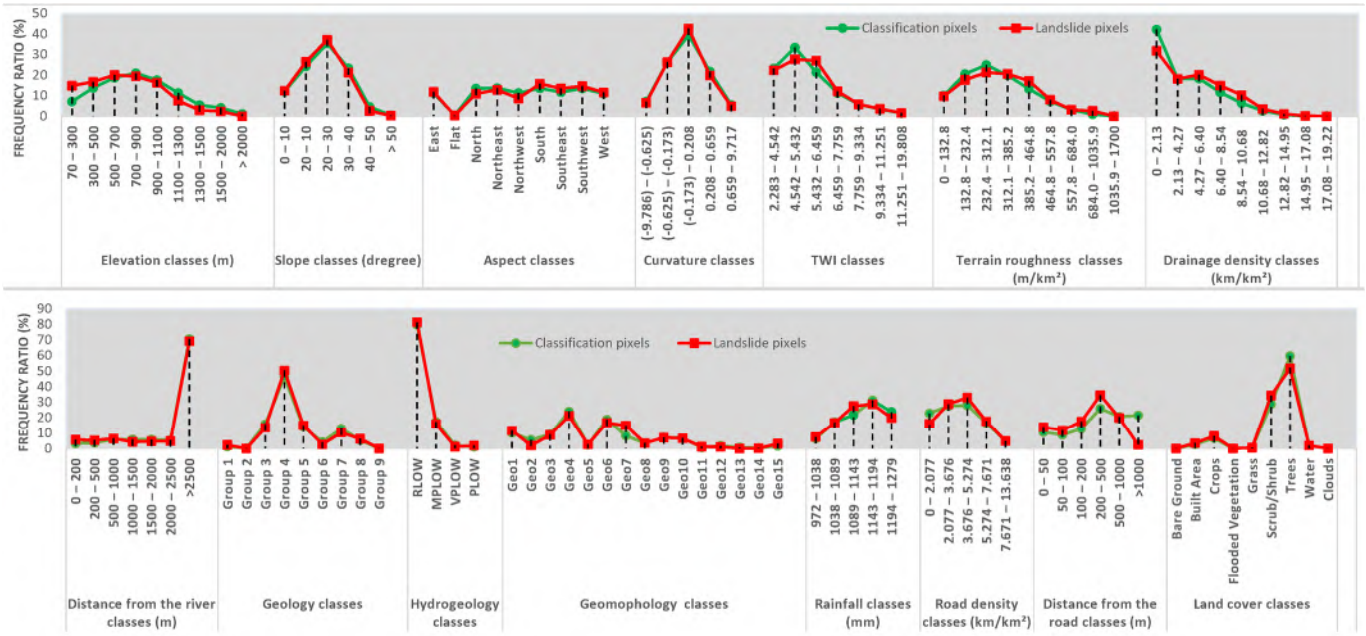

Figure 4. Frequency ratio analysis of landslide pixels on thematic maps of causative factors.

Table 1. Factors contributing to landslides and their categorization.

Factor	Classification	Classification Method
Elevation (m)	(1) 70–300, (2) 300–500, (3) 500–700, (4) 700–900, (5) 900–1100, (6) 1100–1300, (7) 1300–1500, (8) 1500–2000, (9) >2000	Natural breaks
Slope angle (degree)	(1) 0–10, (2) 10–20, (3) 20–30, (4) 30–40, (5) 40–50, (6) >50	Natural breaks
Aspect	(1) East, (2) Plat, (3) North, (4) Northeast, (5) Northwest, (6) South, (7) Southeast, (8) Southwest, (9) West	Azimuth
Curvature	(1) [(−9.786)–(−0.625)], (2) [(−0.625)–(−0.173)], (3) [(−0.173)–0.208], (4) [0.208–0.659], (5) [0.659–9.717]	Natural breaks
TWI	(1) [2.283–4.542], (2) [4.542–5.432], (3) [5.432–6.459], (4) [6.459–7.759], (5) [7.759–9.334], (6) [9.334–11.251], (7) [11.251–19.808]	Natural breaks
Terrain roughness (m/km2)	(1) [0–132.8], (2) [132.8–232.4], (3) [232.4–312.1], (4) [312.1–385.2], (5) [385.2–464.8], (6) [464.8–557.8], (7) [557.8–684.0], (8) [684.0–1035.9], (9) [1035.9–1700]	Natural breaks
Drainage density (km/km2)	(1) [0–2.13], (2) [2.13–4.27], (3) [4.27–6.40], (4) [6.40–8.54], (5) [8.54–10.68], (6) [10.68–12.82], (7) [12.82–14.95], (8) [14.95–17.08], (9) [17.08–19.22]	Natural breaks
Distance from the rivers	(1) [0–200], (2) [200–500], (3) [500–1000], (4) [1000–1500], (5) [1500–2000], (6) [2000–2500], (7) >2500	Natural breaks
Geology	(1) Group 1 (Quaternary), (2) Group 2 (Paleogene), (3) Group 3 (Jura-Creta-Cretaceous), (4) Group 4 (TriasTriassic), (5) Group 5 (Devon-Devonia), (6) Group 6 (Cambri-Ordiovic), (7) Group 7 (NeoproterozoiCambri), (8) Group 8 (Carbon-Permi), (9) Group 9 (Unknown)	Geological categories
Hydrogeology	(1) Very productive aquifer, (2) Moderately productive aquifer, (3) Poorly productive aquifer, (4) Very poorly productive aquifer	Hydrogeological categories

Table 1. *Cont.*

Factor	Classification	Classification Method
Geomorphology	(Zone 1) Valley of invasion, (Zone 2) Cavitation plateaus develop on carbonate rocks, (Zone 3) Driftwood washing plateau grows on carbonate rock, (Zone 4) Erosion plateaus develop onlimestones, (Zone 5) Erosion denudationplateaus develop on limestones, (Zone 6) Cavitation mountain range growing on carbonate rock, (Zone 7) Massive and structural mountain ranges developed on non-carbonate rocks, (Zone 8) Erosion and erosion mountain ranges develop on rocks, (Zone 9) Erosion massif develops on carbonate rock, (Zone 10) Masses and eroded mountain ranges develop on rocks, (Zone 11) The valley erodes and accumulates, (Zone 12) Cavitation mountain range growing on non-carbonate rock, (Zone 13) The mountain range erodes the structure growing on the rock, (Zone 14) Karst Funnel, (Zone 15) Invasion valley.	Geomorphological categories
Rainfall (mm)	(1) [972–1038], (2) [1038–1089], (3) [1089–1143], (4) [1143–1194], (5) [1194–1279]	Geostatistical Kriging method
Road density	(1) [0–2.077], (2) [2.077–3.676], (3) [3.676–5.274], (4) [5.274–7.671], (5) [7.671–13.638]	Natural breaks
Distance from the road	(1) [0–50], (2) [50–100], (3) [100–200], (4) [200–500], (5) [500–1000], (6) >1000	Natural breaks
Land cover	(1) Bare ground, (2) Built area, (3) Clouds, (4) Crops, (5) Flooded vegetation, (6) Grass, (7) Scrub/Shrub, (8) Tree, (9) Water.	Landcover categories

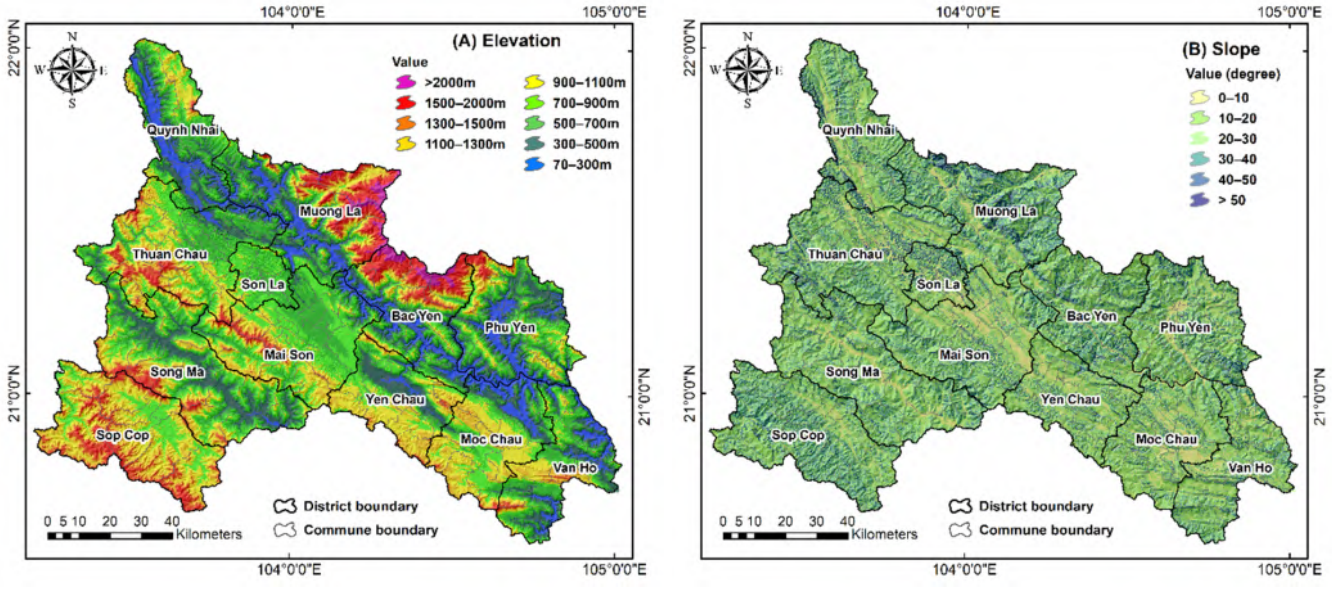

Figure 5. *Cont.*

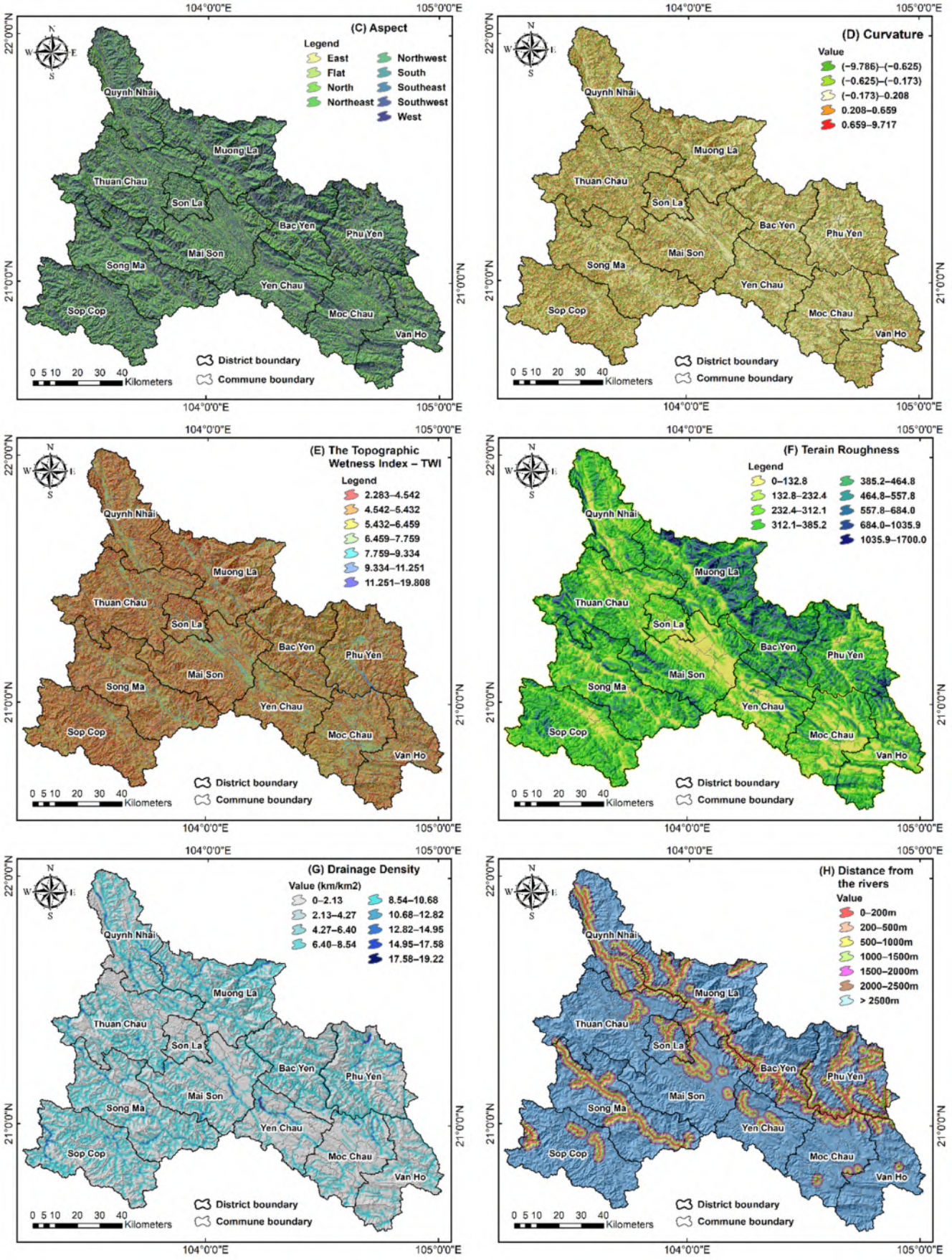

Figure 5. *Cont.*

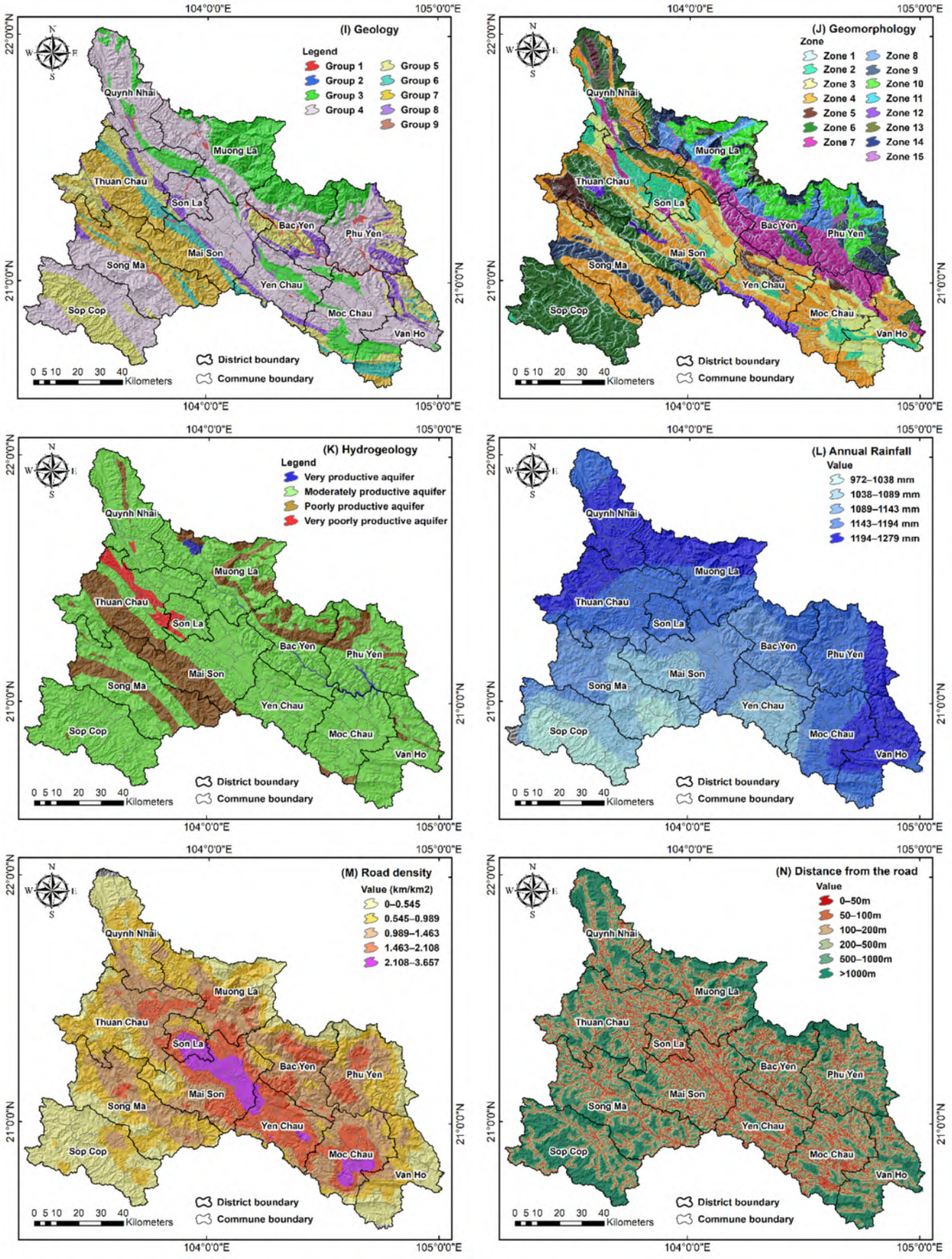

Figure 5. *Cont.*

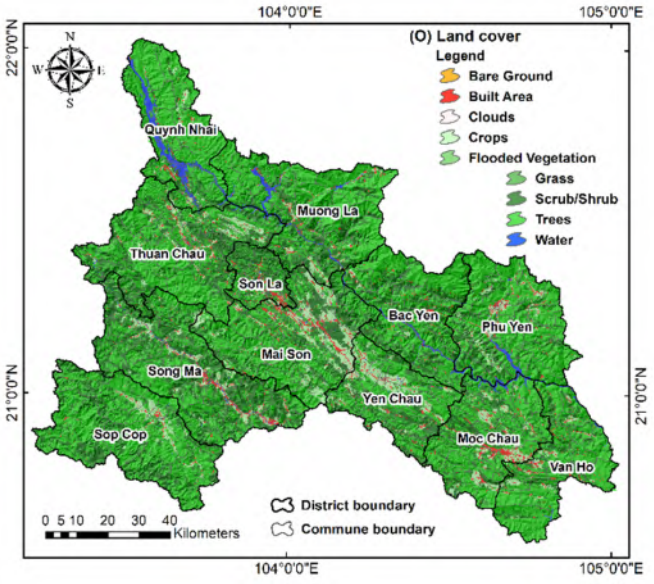

Figure 5. Thematic maps of landslide conditioning factors: (**A**) elevation, (**B**) slope, (**C**) aspect, (**D**) curvature, (**E**) TWI, (**F**) terrain roughness, (**G**) drainage density, (**H**) distance from the rivers, (**I**) geology, (**J**) geomorphology, (**K**) geohydrology, (**L**) annual rainfall, (**M**) road density, (**N**) distance from the road, and (**O**) land cover.

3.2. Multicollinearity Analysis

Determining multicollinearity between independent variables is very important to ensure the stability and reliability of the regression model [56]. Variance Inflation Factors (VIFs) and tolerance parameters can discover multicollinearity among landslide causative factors to identify the most suitable input factors for landslide susceptibility modelling [57,58]. The issue of multicollinearity arises when the VIF exceeds 10 or the tolerance falls below 0.1 [59]. These thresholds indicate the presence of strong correlations among the variables, affecting the reliability of the regression coefficients and the overall forecast model [60,61]. This study employs VIF and tolerance parameters to detect and mitigate the risks associated with multicollinearity among input variables and to build accurate predictive models.

3.3. Machine Learning Models

3.3.1. UltraBoost (UB)

UB is an advanced ML algorithm that belongs to the family of boosting algorithms that aims to improve prediction accuracy by combining multiple weak learners (usually decision trees) into a single strong model [19]. In this way, each new weak learner is trained to correct the errors made by the previously added learners [20]. The details of this algorithm can be described as follows [19]:

1. Input a training dataset $D = (x_i, y_i)$, $i = 1 - N$, where N represents the total number of training samples; x_i denotes the input values; and y_i denotes the target values. An initial simple prediction $F_0(x)$ for all training samples is set by minimizing the loss function L to find an initial constant prediction.

$$F_0(x) = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i, \gamma) \quad (2)$$

where $F_0(x)$ is the initial simple prediction, L describes the loss function, and γ is the initial constant prediction.

2. This technique iteratively trains weak learners, typically decision trees, for each iteration $t = 1, 2, \ldots, T$, where T denotes the number of iterations or the total number of weak models. At each iteration t, a new weak learner is trained to correct the errors of the combined model from previous iterations. Then, the model is updated with the new weak learner's predictions, scaled by a learning rate ε to control its contribution.

$$F_t(x) = F_{t-1}(x) + \varepsilon . h_t(x) \tag{3}$$

where $F_t(x)$ and $F_{t-1}(x)$ signify the predictions of the model at iteration t and $t-1$, respectively; ε is the learning rate; and $h_t(x)$ denotes the t-th weak learner.

3. Continue adjusting the weights of the training samples based on errors from previous iterations. The goal is to find the weak learner $h_t(x)$ that minimizes the weighted loss function.

$$h_t = \underset{h}{\operatorname{argmin}} \sum_{i=1}^{N} w_i L(y_i, F_{t-1}(x_i) + h(x_i)) \tag{4}$$

where h_t is the t-th weak learner; w_i represents assigned weights for the training samples, adjusted based on the errors of previous learners; and $h(x_i)$ is the prediction of the weak learner for the i-th sample.

4. Finally, the predictions of all weak learners are combined to form the final strong model. The combination is often a weighted sum of the individual learners' predictions.

$$F(x) = \sum_{t=1}^{T} \varepsilon . h_t(x) \tag{5}$$

where $F(x)$ is the final prediction and T describes the total number of weak learners.

3.3.2. Decorate (DC)

The DC algorithm was introduced and developed by Melville and Mooney [62]. It is an ensemble learning technique that enhances the performance of a base classifier by combining it with additional classifiers trained on modified versions of the original dataset [62,63]. This algorithm operates based on the following principle [62]:

1. Input a training dataset $D = (x_i, y_i),\ i = 1 - N$ with N samples. Decorate begins by creating a set of base classifiers $B = B_1,\ B_2, \ldots,\ B_j,\ j = 1 - M$, where M is the total number of base classifiers. Each base classifier B_j is trained on a weighted version of the original dataset, with sample weights adjusted iteratively based on previous classification errors.
2. After training the base classifiers, a meta-learner is trained using the outputs of the base classifiers as input features. Let $h(x_i)$ represent the meta-learner's prediction. The meta-learner h is trained on the base classifier predictions B and the target label y_i. The objective of the meta-learner training is to minimize the loss function L for the meta-learner h.

$$L(h) = \sum_{i=1}^{N} L\left(y_i, h\left\{\sum_{j=1}^{M} B_j(x_i)\right\}\right) \tag{6}$$

where L is the loss function and $h\left\{\sum_{j=1}^{M} B_j(x_i)\right\}$ represents the meta-learner's output given the predictions of the base classifiers B for sample x_i.

3. The final prediction F for a new sample x is obtained by first using the base classifiers B to generate predictions and then applying the meta-learner h.

$$F(x) = \sum_{j=1}^{M} h\left(B_j(x)\right) \tag{7}$$

3.3.3. Dagging (DG)

The DG algorithm was first introduced by Ting & Witten in (1997) [64]. It is a machine learning ensemble method that divides the training dataset into disjoint subsets, training a separate classifier on each subset and then aggregating their predictions [65]. The Dagging algorithm can be summarized as follows

1. Input a training dataset $D = (x_i, y_i)$, $i = 1 - N$, with N samples. Dagging starts by splitting the training dataset D into k disjoint subsets, such that $D = D_1 \cup D_2 \cup \ldots \cup D_k$, with no overlap between the subsets, ensuring $D_i \cap D_j = \varnothing$ for any $i \neq j$.
2. Each disjoint subset D_i is used to train a base classifier B_i with $i = 1, 2, \ldots, k$ to create a classification model.
3. The final prediction F for a new sample x is received based on the aggregation of predictions using majority voting and can be expressed as follows:

$$F(x) = \underset{y}{\operatorname{argmax}} \sum_{i=1}^{k} I(B_i(x) = y) \tag{8}$$

where I is the indicator function that equals 1 if $B_i(x) = y$, and 0 if $B_i(x) \neq y$.

3.3.4. Bagging (BG)

The BG algorithm is a technique used to improve ML algorithms' predictive capacity, accuracy, and stability (Breiman, 1996 [66]). It works by creating multiple bootstrap samples from the original training data, and each is used to train a separate base classifier [67]. The final prediction is made by aggregating the predictions of these individual classifiers. This technique can avoid overfitting and acts well on strong classifiers [68]. Mathematically, the content of this algorithm can be presented as shown below:

1. Input a training dataset $D = (x_i, y_i), i = 1 - N$, with N samples, and generate bootstrap samples D_b for $b = 1, 2, \ldots, B$. Each bootstrap sample D_b is created by sampling N observations from the initial training dataset D with replacement, $D_b = (x_{bi}, y_{bi})$.
2. A separate model is trained on each bootstrap sample D_b, leading to the creation of a classifier $B_k(x)$ from each corresponding bootstrap sample D_b.
3. The final prediction F for a new sample x is determined by taking the majority vote for classification tasks.

$$F(x) = \underset{y}{argmax} \sum_{b=1}^{B} (F_b(x)) \tag{9}$$

3.3.5. MultiScheme (MS)

The MS algorithm is designed for ensemble learning, combining multiple models to improve prediction accuracy [69]. It involves weighting and combining the predictions of different base learners to produce a final prediction [70]. The method involves the following key steps [69]:

1. Input a training dataset $D = (x_i, y_i)$, $i = 1 - N$, with N samples. Let $B_1, B_2, \ldots, B_M$ represent the base learners, where M is the total number of base models B and $B_i(x)$ is the prediction of the i-th base learner for an input feature x.
2. Each base learner B_i is assigned a weight w_i based on its error rate, with models that perform better receiving higher weights.

$$w_i = \frac{\log\left(\frac{1-\varepsilon_i}{\varepsilon_i}\right)}{2} \tag{10}$$

where ε_i represents the error rate of the i-th base learner on a validation set.

3. The final ensemble prediction $F(x)$ is made by taking a weighted sum of the base models' predictions.

$$F(x) = \text{sign}\left(\sum_{i=1}^{M} w_i.B_i(x)\right) \quad (11)$$

where $F(x)$ is the final ensemble prediction for a new input x, M is the total number of base learners, and sign(z) denotes the sign function and reaches +1 if $z > 0$ and -1 if $z < 0$.

3.3.6. MultiBoostAB (MB)

The MB algorithm was first proposed by Webb [71]. This algorithm is an ensemble learning method that extends AdaBoost by incorporating the concepts of wagging (bootstrap aggregation) and BrownBoost, making it more robust to noisy data and outliers [71]. The content of this algorithm can be described as follows:

1. Input a training dataset $D = (x_i, y_i), i = 1 - N$, with N samples. This algorithm begins by assigning equal weights w_i to all training samples.
2. For each iteration $t = 1, 2, \ldots, T$.
   Create a bootstrapped dataset $D\prime_t$ by sampling with replacement from the original training dataset D.
   Train a weak classifier $h_t(x)$ on the new bootstrapped dataset $D\prime_t$.
   Calculate the classifier's error ε_t.

$$\varepsilon_t = \frac{\sum_{i=1}^{N} w_i.I(h_t(x_i) \neq y_i)}{\sum_{i=1}^{N} w_i} \quad (12)$$

   where I is the indicator function that equals 1 if $h_t(x_i) \neq y_i$ and 0 if $h_t(x_i) = y_i$.
   Based on the classifier's error ε_t, a weight, w_t, is assigned to the classifier.

$$w_t = \frac{1}{2}\, ln\left(\frac{1 - \varepsilon_t}{\varepsilon_t}\right) \quad (13)$$

   The sample weights w_i are updated to emphasize misclassified samples.
   This process is repeated for several iterations, each adding a new weak classifier to the ensemble.
3. The final strong classifier $F(x)$ is determined based on a weighted combination of all weak classifiers.

$$F(x) = \text{sign}\left(\sum_{t=1}^{T} w_t.h_t(x)\right) \quad (14)$$

   where $F(x)$ is the final strong classifier for a new input x, T is the total number of boosting iterations or the total number of weak classifiers, and sign(z) denotes the sign function and reaches +1 if $z > 0$ and -1 if $z < 0$.

3.3.7. Cascade Generalization (CG)

The CG algorithm was proposed by Gama and Brazdil [72] and is an ensemble learning technique that enhances prediction accuracy by combining multiple models in a cascading manner [72]. The core concept involves training models sequentially, where each model's prediction becomes an input for subsequent models [68]. The general framework of this algorithm can be expressed as follows [72]:

1. Input a training dataset $D = (x_i, y_i)$, $i = 1 - N$, with N samples. This technique starts with a base model that uses the original input features to predict the target variable.

$$F_i^{(1)} = f_1(x_i) \quad (15)$$

2. The prediction from this base model $F_i^{(1)}$ is then fed as an input feature into the next model, which combines it with the original features to make a more refined prediction.

$$F_i^{(2)} = f_2\left(x_i, F_i^{(1)}\right) \tag{16}$$

3. This cascading process continues through multiple levels, with each subsequent model using the predictions of all previous models as additional features.

$$F_i^{(k)} = f_k\left(x_i, F_i^{(1)}, F_i^{(2)}, \ldots, F_i^{(k-1)}\right) \tag{17}$$

4. The final prediction is made by the last model in the cascade.

$$F_i = F_i^{(K)} \tag{18}$$

where F_i is the final predicted output after K levels of cascading and $F_i^{(K)}$ denotes the prediction from the final with K-th level mode.

3.4. Model Validation and Performance Comparison

Estimating the prediction performance of ML models in landslide susceptibility mapping plays a crucial role in ensuring the reliability and usefulness of the forecast models in real-world applications [73]. A diverse set of statistical indices is often used to evaluate the forecasting capacity of ML models [56]. The evaluation is critical for validating the current ML model and guiding future improvements and comparative analyses with other ML models [70]. In the present study, a diverse set of statistical indices, including sensitivity, specificity, accuracy (ACC), F-measure, Jaccard, Receiver Operating Characteristic (ROC) Curve, and Area Under the ROC Curve (AUC), were employed to gauge the predictive performance of six proposed ensemble ML models. These mentioned indices are calculated in the following equations:

$$Sensitivity = \frac{TP}{TP + FN} \tag{19}$$

$$Specificity = \frac{TN}{FP + TN} \tag{20}$$

$$F - measure = \frac{2TP}{2TP + FP + FN} \tag{21}$$

$$Jaccard = \frac{TP}{TP + FP + FN} \tag{22}$$

$$ACC = \frac{TN + TP}{TN + FN + TP + FP} \tag{23}$$

where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively.

Sensitivity and specificity reveal the model's ability to correctly identify landslide-prone and non-landslide areas, respectively. Sensitivity measures how well the model identifies true positives, while specificity assesses the model's accuracy in recognizing true negatives [74]. The Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) Curve provides a comprehensive measure of a model's ability to discriminate between landslide-prone and non-landslide areas [73]. It reflects how well the model distinguishes between the two classes across all thresholds, with a higher AUC value indicating better performance. Specifically, an AUC between 0.5 and 0.6 reflects very poor performance, 0.6 to 0.7 indicates poor performance, 0.7 to 0.8 represents moderate performance, 0.8 to 0.9 denotes good performance, and 0.9 to 1.0 signifies excellent performance [56].

3.5. Experimental Procedure

This study applied the Waikato Environment for Knowledge Analysis (WEKA) software version 3.9.6 to create landslide susceptibility models. Weka is a popular open-source software tool designed to provide a comprehensive suite of tools for data analysis and predictive modelling. The input data consisted of a landslide inventory dataset, which included landslide locations marked with a value of 1 (LS = 1, indicating a confirmed landslide) and corresponding non-landslide points marked with a value of 0 (LS = 0, indicating non-landslide sites). The dataset was divided into a training set (70%, approximately 1240 sites for each type) and a testing set (30%, approximately 531 sites for each type) to serve as input data for ML training. Fifteen predictor variables were extracted for these points: altitude, slope angle, aspect, curvature, terrain roughness, TWI, rainfall, hydrogeology, geology, geomorphology, drainage density, land cover, road density, distance from rivers, distance from roads, and road network. All input data were converted into the CSV format to ensure consistency in attributes across the training and testing datasets. The training and testing datasets were then uploaded into Weka's Explorer interface for model building. The target variable in this model represents the predicted landslide susceptibility, expressed as a value ranging from 0 to 1. This value indicates the probability of a landslide occurring at each specific point within the study area. A value closer to 0 suggests a low likelihood of a landslide, while a value approaching 1 indicates a higher probability of landslide occurrence.

In this study, seven ML algorithms were employed to develop six ML ensemble models, namely Decorate-UltraBoost (DCUB), Dagging-UltraBoost (DGUB), Bagging-UltraBoost (BGUB), MultiScheme-UltraBoost (MSUB), Cascade Generalization-UltraBoost (CGUB), and MultiBoostAB-UltraBoost (MBUB) for landslide susceptibility mapping. The UltraBoost model served as the base classifier, and the Decorate, Dagging, Bagging, MultiScheme, Cascade Generalization, and MultiBoostAB models acted as meta-classifiers. The parameter settings for each ML algorithm in the WEKA software are described in Table 2.

Table 2. Parameter settings for each ML model in WEKA software.

ML Algorithms	Parameter Settings
Decorate-UltraBoost	classifiers.UltraBoost -S 1 -B "classifiers.Decorate -E 15 -R 1.0 -S 1 -I 50 -W classifiers.SysFor -- -L 10 -N 60 -G 0.3 -S 0.3 -C 0.25" -B "classifiers.RotationForest -G 3 -H 3 -P 50 -S 1 -num-slots 1 -I 10 -W classifiers.RandomTree -- -K 0 -M 1.0 -V 0.001 -S 1"
MultiSearch-UltraBoost	classifiers.UltraBoost -S 1 -B "classifiers.MultiSearch -E CC -class-label 1 -S 1 -W classifiers.SysFor -- -L 10 -N 60 -G 0.3 -S 0.3 -C 0.25" -B "classifiers.RotationForest -G 3 -H 3 -P 50 -S 1 -num-slots 1 -I 10 -W classifiers.RandomTree -- -K 0 -M 1.0 -V 0.001 -S 1"
Dagging-UltraBoost	classifiers.UltraBoost -S 1 -B "classifiers.Dagging -F 10 -S 1 -W classifiers.SysFor -- -L 10 -N 60 -G 0.3 -S 0.3 -C 0.25" -B "classifiers.RotationForest -G 3 -H 3 -P 50 -S 1 -num-slots 1 -I 10 -W classifiers.RandomTree -- -K 0 -M 1.0 -V 0.001 -S 1"
Bagging-UltraBoost	classifiers.UltraBoost -S 1 -B "classifiers.Bagging -P 100 -S 1 -num-slots 1 -I 10 -W classifiers.SysFor -- -L 10 -N 60 -G 0.3 -S 0.3 -C 0.25" -B "classifiers.RotationForest -G 3 -H 3 -P 50 -S 1 -num-slots 1 -I 10 -W classifiers.RandomTree -- -K 0 -M 1.0 -V 0.001 -S 1"
CascadeGeneralization-UltraBoost	classifiers.UltraBoost -S 1 -B "classifiers.CascadeGeneralization -U -L -C -X 5 -M \"classifiers.SysFor -L 10 -N 60 -G 0.3 -S 0.3 -C 0.25\" -S 1 -num-slots 1 -B \"classifiers.SysFor -L 10 -N 60 -G 0.3 -S 0.3 -C 0.25\"" -B "classifiers.RotationForest -G 3 -H 3 -P 50 -S 1 -num-slots 1 -I 10 -W classifiers.RandomTree -- -K 0 -M 1.0 -V 0.001 -S 1"
MultiBoostAB-UltraBoost	classifiers.UltraBoost -S 1 -B "classifiers.MultiBoostAB -C 3 -P 100 -S 1 -I 10 -W classifiers.SysFor -- -L 10 -N 60 -G 0.3 -S 0.3 -C 0.25" -B "classifiers.RotationForest -G 3 -H 3 -P 50 -S 1 -num-slots 1 -I 10 -W classifiers.RandomTree -- -K 0 -M 1.0 -V 0.001 -S 1"

Remarks: S: set the random seed; B: base classifier; E: number of iterations; R: ratio of artificial examples; I: iterations for boosting; W: base learner; L: iterations; N: instances per iteration; G: gradient; S: split threshold; C: confidence; H: hidden layers; P: percentage; num-slots: CPU slots; K: attributes; M: minimum instances; V: variance; class label: label; F: folds; U: multi-layer; X: maximum layers.

4. Data and Methods Used for Landslide Exposure, Vulnerability, and Adaptive Capacity Maps

4.1. Data for Risk Assessment

4.1.1. Exposure Indicators

Humans, infrastructure, and agriculture are the main criteria for evaluating exposure (Table 3). Population density is a crucial indicator in landslide exposure analysis for human criteria because it directly correlates with the concentration of human lives and valuable assets within a specific area [75].

Table 3. Exposure, vulnerability, and adaptive capacity indicators for risk assessment.

Components	Main Criteria	Sub-Criteria	Unit	Sources
Exposure (E)	Human (E1)	Population density (E1-1)	Person/km^2	District statistical yearbooks in 2022
	Infrastructure (E2)	Residential land area (E2-1)	ha	
		Length of roads (E2-2)	km	
		Number of large primary and secondary schools (E2-3)	Per 1000 pupil	The website of Son La province (available at https://data.sonla.gov.vn/list-data, accessed on 1 October 2024)
	Agriculture (E3)	Agricultural land area (E3-1)	ha	
		Aquaculture land area (E3-2)	ha	
Vulnerability (V)	Population (V1)	Percentage of children under 6 years old (V1-1)	%	General Department of Natural Disaster Prevention in 2020
		Percentage of elderly people over 60 years old (V1-2)	%	
		Percentage of females between 17 and 59 years old (V1-3)	%	
	Infrastructure (V2)	Percentage of unsolid houses (V2-1)	%	
	Demographics (V3)	Birth rate (V3-1)	Per 1000 people	District statistical yearbooks in 2022
		Death rate (V3-2)	Per 1000 people	
Adaptive Capacity (AC)	Healthcare (AC1)	Number of medical facilities in each commune (AC1-1)	piece	District statistical yearbooks in 2022
	Infrastructure (AC2)	Asphalt road density (AC2-1)	km/km^2	
		Number of high schools, universities, and colleges (AC2-2)	Per 1000 students	The website of Son La province (available at https://data.sonla.gov.vn/list-data, accessed on 1 October 2024)
		Number of enterprises (AC2-3)	Per 1000 people	
	Population (AC3)	Percentage of population aged 17–59 (AC3-1)	%	General Department of Natural Disaster Prevention in 2022

For infrastructure criteria, residential land area, the length of roads, and a number of large primary and secondary schools are sub-criteria. The percentage of residential land area highlights the extent of human settlements that landslides could directly affect [76]. This indicator quantifies land use and reflects the spaces where local communities live, work, and build their lives [77]. The presence of primary and secondary schools in areas susceptible to landslides significantly contributes to the overall exposure of a community to such natural hazards [78]. These educational institutions are critical components of community infrastructure, often accommodating large numbers of children who are particularly vulnerable during a landslide [79]. Road networks are vital infrastructure for daily life and are essential routes for evacuation and emergency response [80]. The disruption of road networks due to landslides can have catastrophic effects on accessibility and emergency management, amplifying the overall exposure of the area [81].

For agriculture criteria, agricultural land area and aquaculture land area are sub-criteria. In the Son La province, the livelihoods of the local community are closely related to farming and aquaculture activities [38]. Agriculture is not just a source of food for the whole province but also a livelihood for many generations relying on the land for their daily needs [82]. Similarly, aquaculture plays a vital role in the livelihoods of those living near rivers and reservoirs, where fish farming is a significant source of food and income [83].

4.1.2. Vulnerability Indicators

Population, infrastructure, and demographics are the main criteria for evaluating vulnerability (Table 3). For population criteria, the percentage of children under 6 years old, the percentage of older people over 60 years old, and the percentage of females between 17 and 59 years old are sub-criteria. Children, elders, and females are particularly vulnerable to the adverse impacts of landslide events due to potential mobility limitations, health issues, and increased dependence on social networks [84]. Children and elderly individuals are particularly vulnerable during landslides due to their unique physical and health-related needs [85]. Females aged 17–59 are often responsible for household management and childcare, making them especially vulnerable during landslides [86].

For infrastructure criteria, the percentage of unsolid houses is a sub-criteria. The study area contains two housing types: solid houses, built with robust concrete or brick, and unsolid houses, made from less durable wood or bamboo [87]. Unsolid houses highlight the quality and stability of housing structures, as poorly constructed or unstable homes are more susceptible to damage from landslides, exacerbating the risks faced by residents [38].

For demographic criteria, birth rate and death rate are sub-criteria. Birth and death rates represent demographic aspects that are crucial in assessing human vulnerability to landslides, as they provide insights into the population's capacity to stand and recover from such events [85].

4.1.3. Adaptive Capacity Indicators

Healthcare, infrastructure, and population are the main criteria for evaluating adaptive capacity (Table 3). Adaptive capacity signifies the adaptability of humans and organizations to apply skills and resources to overcome the adverse impacts of landslide events [88]. The number of medical facilities in each commune is the sub-criteria of healthcare criteria. The number of medical facilities in each commune reflects the availability and accessibility of healthcare services, which is vital for addressing injuries, illnesses, and emergencies resulting from landslides [89].

Asphalt road density, the percentage of solid houses, and number of high schools, universities, colleges, and enterprises per 1000 people are sub-criteria for infrastructure criteria. Asphalt road density is important in transportation infrastructure resilience, facilitating evacuation efforts and emergency response operations in landslide-affected areas [90]. Solid houses can protect human lives from the risk of injury and ensure displacement for local communities [87]. The presence of high schools, universities, and colleges plays a crucial role in enhancing the community's resilience and preparedness for landslides [78]. Furthermore, these educational institutions can contribute to community awareness and education about landslide risks through workshops and training courses, helping to enhance the overall knowledge and preparedness of the local population [91]. Enterprises contribute to building community resilience by generating employment opportunities, fostering economic stability, and supporting infrastructure development [92].

The percentage of the population aged 17–59 is a sub-criteria of population criteria since labour resources serve the main roles in disaster preparedness, response, and recovery efforts, facilitating the restoration of livelihoods after landslide events [93].

4.2. Weighting Method

4.2.1. Iyengar–Sudarshan Method

Factors often denote various degrees of influence on landslide risk, so determining each factor's weight is important [94]. This study applied the Iyengar–Sudarshan technique to identify the weight of landslide exposure, vulnerability, and adaptive capacity. This technique was introduced by Iyengar and Sudarshan [30] to create a combined index from multivariate data. This method has found widespread application in determining multi-criteria weights in risk assessments. According to this, the weights for each main criterion and each sub-criterion in landslide risk assessment are determined using the following equations:

$$w_{pq} = \frac{C}{\sqrt{Var(y_{pq})}} \tag{24}$$

$$C = \left[\sum_{q=1}^{Q} \frac{1}{\sqrt{Var(y_{pq})}}\right]^{-1} \tag{25}$$

$$Var = \frac{1}{P}\sum_{p=1}^{P}(y_{pq} - \overline{y_{pq}})^2 \tag{26}$$

$$\overline{y_{pq}} = \frac{1}{P}\sum_{p=1}^{P} y_{pq} \tag{27}$$

where w_{pq} represents the weight assigned to the p^{th} sub-criterion in each component (landslide exposure, landslide vulnerability, and landslide adaptive capacity); y_{pq} indicates the standardized value; C denotes a normalizing constant; p = 1, 2, . . ., P and q = 1, 2, . . ., Q are the standardized values; Q denotes the indicators that contribute to landslide risk; Var represents the variance in the corresponding indicators to the landslide risk; and $\overline{y_{pq}}$ is considered as the result of a linear combination of y_{pq}.

Finally, the composite weighted values of each factor (exposure, vulnerability, and adaptive capacity) are calculated according to the following equation:

$$W_{(E,V,AC)} = \frac{\sum_{p=1}^{P} w_{pq} y_{pq}}{P} \tag{28}$$

where W denotes the composite weighted value of exposure (E), vulnerability (V), and adaptive capacity (AC); P represents the total number of criteria in each component; and w_{pq} indicates the criterion weight.

4.2.2. Factor Normalization

These input data include relevant criteria or parameters indicating landslide exposure, vulnerability, and adaptive capacity. They were collected from field surveys, statistical records, and remote sensing resources. Because the measurement units of each parameter are different, normalization needs to be conducted to standardize all these values on a scale ranging from 0 to 1. After gathering and processing the input data within the GIS application, they were standardized from 0 to 1 using the approach proposed by Connor and Hiroki [95]. The content of this method is outlined as follows:

$$y_{pq} = \frac{X_{pq} - MinX_{pq}}{MaxX_{pq} - MinX_{pq}} \tag{29}$$

$$y_{pq} = \frac{MaxX_{pq} - X_{pq}}{MaxX_{pq} - MinX_{pq}} \tag{30}$$

where y_{pq} represents the standardized value of the q parameter in the p^{th} region; X_{pq} denotes the starting value q in each criterion p^{th}; and $MaxX_{pq}$, $MinX_{pq}$ correspond to the maximum and minimum values of the indicator for sub-criteria q within criteria p^{th}.

Suppose a positive relationship exists between the parameters and the risk elements. In this case, Equation (29) is employed for normalization, while if a negative relationship exists between the indicators and the vulnerability components, Equation (30) is utilized to standardize the indicator. Subsequently, these indicators were integrated into Iyengar and Sudarshan's method to calculate weights and create a combined index to assess landslide risk.

5. Methodology Flowchart

The research methodology is structured around three stages: (1) modelling landslide susceptibility, (2) collecting landslide exposure, vulnerability, and adaptive capacity indicators, and (3) producing a landslide risk map. The methodological approach of this research is represented in Figure 6. In the initial step, we utilized six ML ensemble models of Decorate-UltraBoost (DCUB), Dagging-UltraBoost (DGUB), Bagging-UltraBoost (BGUB), MultiScheme-UltraBoost (MSUB), Cascade Generalization-UltraBoost (CGUB), and MultiBoostAB-UltraBoost (MBUB) to model landslide susceptibility. The landslide susceptibility map that demonstrated the highest accuracy through validation was selected as the landslide hazard map. The second step focused on developing landslide exposure, vulnerability, and adaptive capacity maps by gathering a range of socio-economic and physical indicators. The Iyengar–Sudarshan method was used to identify the weights of these indicators, resulting in maps that illustrate landslide exposure, vulnerability, and adaptive capacity. Finally, the landslide hazard, exposure, vulnerability, and adaptive capacity maps with their respective weights were integrated for spatial analysis to build a landslide risk map for the study area.

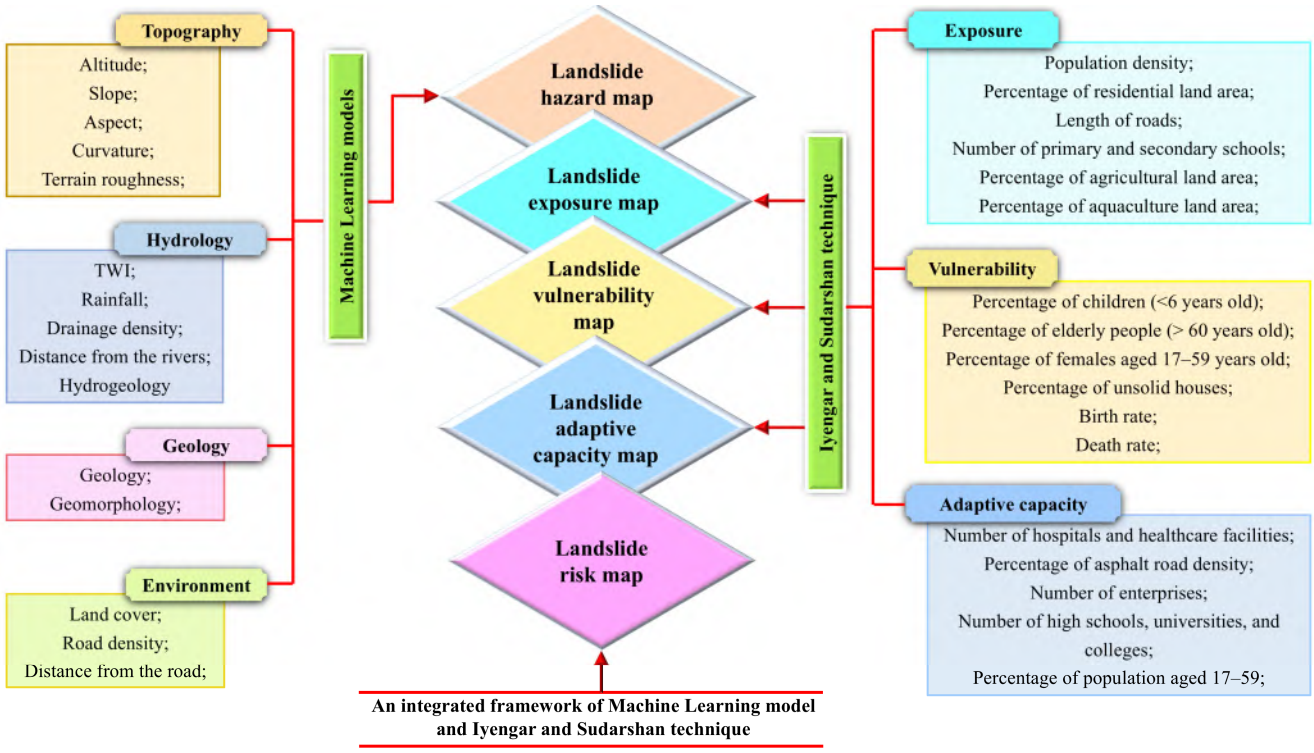

Figure 6. The proposed framework for landslide risk assessment.

6. Results

6.1. Landslide Susceptibility Modelling

6.1.1. Multicollinearity Analysis and Factor Selection

The results of VIF and tolerance calculations for fifteen factors influencing landslides are presented in Table 4. All these factors exhibit VIF values smaller than 10 and tolerance values greater than 0.1. Consequently, they are suitable for constructing landslide susceptibility models in this study.

Table 4. Indices of multicollinearity diagnosis for independent variables.

Independent Variables	Collinearity Statistics	
	VIF	Tolerance
Terrain roughness	1.477	0.677
Drainage density	1.279	0.782
Geomorphology	1.171	0.854
TWI	2.002	0.499
Slope	1.818	0.550
Road density	1.337	0.748
Distance from the road	1.439	0.695
Distance from the rivers	1.464	0.683
Rainfall	1.162	0.861
Land cover	1.172	0.853
Elevation	1.689	0.592
Geohydrology	1.100	0.909
Curvature	1.474	0.679
Aspect	1.009	0.992
Geology	1.123	0.891

6.1.2. Model Validation and Comparison

The validation of susceptibility models is necessary to identify the degree of effectiveness and the applied capability of these models. This study developed six ensemble ML models (DCUB, DGUB, BGUB, MSUB, CGUB, and MBUB) to build landslide susceptibility maps for Son La province in Vietnam. The predictive performance of these models was evaluated using cross-validation, with several standard quantitative indices on both the training and validation datasets in Table 5 and Figure 7.

Table 5. The calculated metrics for assessing the model's performance.

Sample	Indices	DCUB	MSUB	DGUB	BGUB	CGUB	MBUB
Training	TP	1188	1168	1129	1180	1153	1143
	TN	1162	1050	1048	1095	896	1067
	FP	77	189	191	144	343	172
	FN	52	72	111	60	87	97
	PPV(%)	93.91	86.07	85.53	89.12	77.07	86.92
	NPV(%)	95.72	93.58	90.42	94.81	91.15	91.67
	SST(%)	95.81	94.19	91.05	95.16	92.98	92.18
	SPF(%)	93.79	84.75	84.58	88.38	72.32	86.12
	ACC (%)	94.80	89.47	87.82	91.77	82.65	89.15
	F-Measure (%)	94.85	89.95	88.20	92.04	84.28	89.47
	Jaccard (%)	90.21	81.74	78.90	85.26	72.84	80.95
	MCC (%)	90.13	81.08	78.56	84.86	70.72	80.62
	Kappa	0.89590	0.7894	0.7563	0.8354	0.6531	0.783

Table 5. *Cont.*

Sample	Indices	DCUB	MSUB	DGUB	BGUB	CGUB	MBUB
Validating	TP	445	463	418	445	476	438
	TN	388	343	367	351	314	375
	FP	144	189	165	181	218	157
	FN	86	68	113	86	55	93
	PPV(%)	75.55	71.01	71.70	71.09	68.59	73.61
	NPV(%)	81.86	83.45	76.46	80.32	85.09	80.13
	SST(%)	83.80	87.19	78.72	83.80	89.64	82.49
	SPF(%)	72.93	64.47	68.98	65.98	59.02	70.49
	ACC (%)	78.36	75.82	73.85	74.88	74.32	76.48
	F-Measure (%)	79.46	78.28	75.04	76.92	77.71	77.80
	Jaccard (%)	65.93	64.31	60.06	62.50	63.55	63.66
	MCC (%)	65.89	62.39	61.19	61.79	60.03	63.77
	Kappa	0.56730	0.5166	0.4770	0.4977	0.4865	0.5297

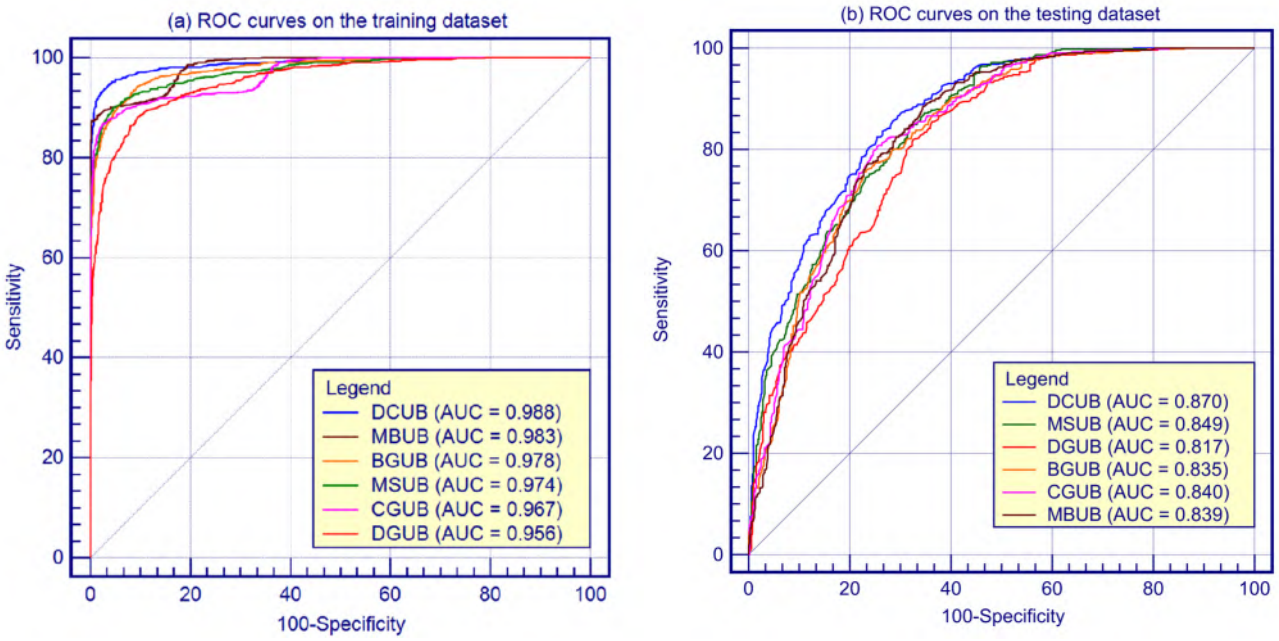

Figure 7. ROC and AUC of six ensemble ML models: (**a**) the training dataset and (**b**) the validating dataset.

Focusing on the validating sample, the DCUB model has the highest performance metrics compared to the left model of MSUB, DGUB, BGUB, CGUB, and MBUB for specificity (72.93%), accuracy (78.36%), F-Measure (79.46%), Jaccard (65.93%), and MCC (65.89%). The CGUB model exhibits the highest sensitivity (89.64%). Analysis of the ROC Curve further shows that the DCUB model achieves superior performance with an AUC of 0.870, followed by MSUB (0.849), CGUB (0.840), MBUB (0.839), BGUB (0.835), and DGUB (0.817). The results from the validating dataset suggest that the DCUB model has the highest predictive capabilities.

Regarding the training sample, the DCUB model achieves the highest performance metrics compared to the remaining models (MSUB, DGUB, BGUB, CGUB, and MBUB) with sensitivity (95.81%), specificity (93.79%), accuracy (94.80%), F-Measure (94.85%), Jaccard (90.21%), and MCC (90.13%). Furthermore, the ROC analysis indicates that the DCUB model has the highest AUC (0.988), followed by MBUB (0.983), BGUB (0.978), MSUB (0.974), CGUB (0.967), and DGUB (0.956). The findings from the training dataset indicate that the DCUB model offers the highest predictive accuracy.

6.1.3. Landslide Hazard Mapping

Based on the results in Section 6.1.2, the DCUB model was identified as the best predictive model for generating the landslide susceptibility map for the study area. This map highlights how different regions are spatially prone to landslide occurrences. The landslide hazard map was classified into five categories, including very low, low, moderate, high, and very high levels, using the quantile technique in ArcGIS Pro 3.1.0 software (Figure 8). The received results highlight a distinct zonal distribution pattern of landslide susceptibility across Son La province, with areas classified as having high and very high susceptibility predominantly situated in the Northeast, particularly in Muong La, Bac Yen, and Phu Yen districts. Similarly, the Southwest also showed significant landslide susceptibility, especially in Thuan Chau, Song Ma, parts of Mai Son, and Sop Cop districts. The regions with low and very low landslide susceptibility are primarily found stretching across the central region of the study area from west to east, covering Son La, parts of Mai Son, and Moc Chau districts.

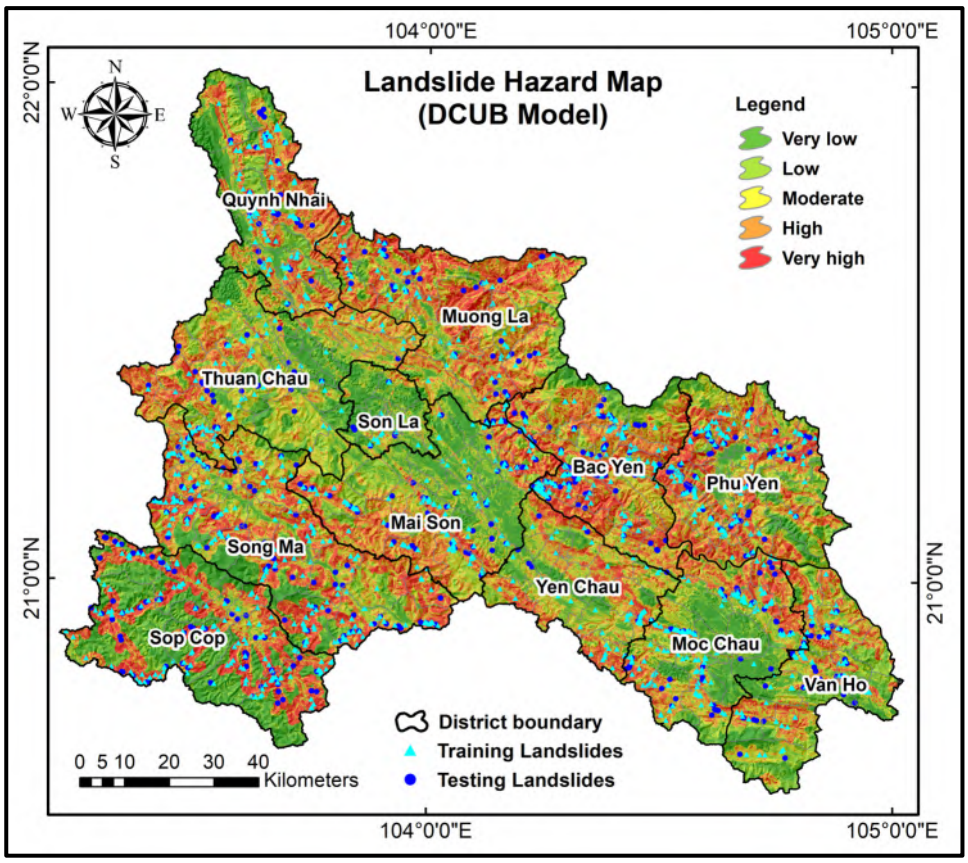

Figure 8. Landslide hazard map for Son La province using DCUB model.

6.2. Landslide Exposure Analysis

The indicators representing landslide exposure, vulnerability, and adaptive capacity for each commune were gathered from various sources, including the 2022 statistical yearbooks of the 11 districts within the province and the National Data Portal website (https://data.gov.vn/SitePages/Index.aspx#/index, accessed on 1 October 2024). A comprehensive range of datasets was employed and prepared in the ArcGIS Pro environment to ensure a unique format to serve the landslide risk assessment, including topographic, geological, hydrological, environmental, physical, and socio-economic factors. These factors were converted into the raster format with a spatial resolution of 30 m to ensure detailed analysis for landslide risk assessment in the Son La province. A total of 18 indicators were collected and categorized into three groups: exposure (E), vulnerability (V), and adaptive capacity (AC) to estimate landslide risk.

The landslide exposure indicators and their respective weights were calculated using the Iyengar-Sudarshan method, as outlined in Table 6. These indicators were then aggregated using the weighted sum tool in the GIS application, and the resulting maps are shown in Figure 9.

Table 6. Landslide exposure criteria and sub-criteria weighting using Iyengar–Sudarshan method.

Components	Main Criteria			Sub-Criteria		
	Label	Code	Weight	Label	Code	Weight
Exposure (E)	Human	E1	0.364	Population density	E1-1	1.000
	Infrastructure	E2	0.369	Residential land area	E2-1	0.392
				Length of roads	E2-2	0.303
				Number of large primary and secondary schools	E2-3	0.305
	Agriculture	E3	0.266	Agricultural land area	E3-1	0.503
				Aquaculture land area	E3-2	0.497

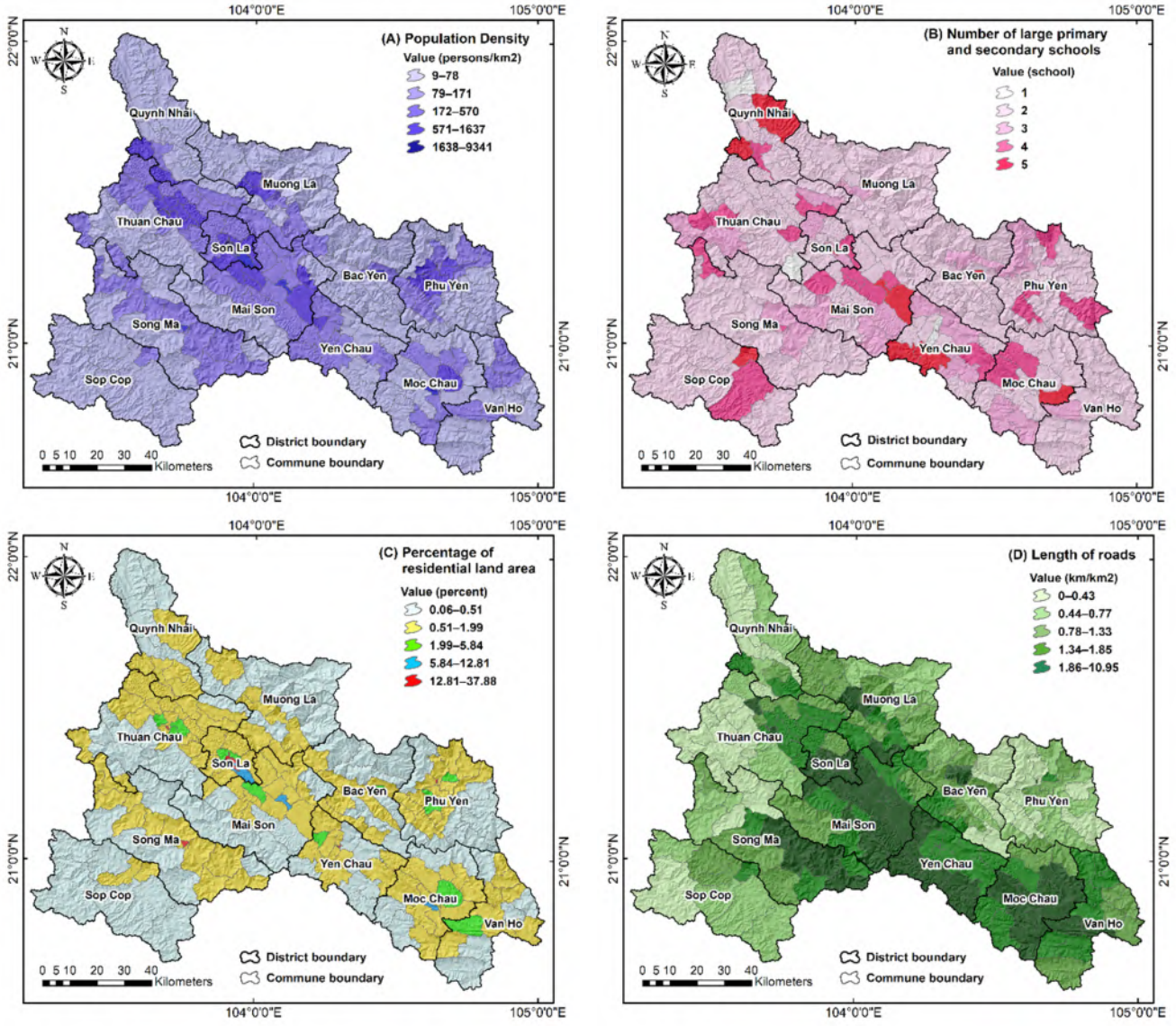

Figure 9. *Cont.*

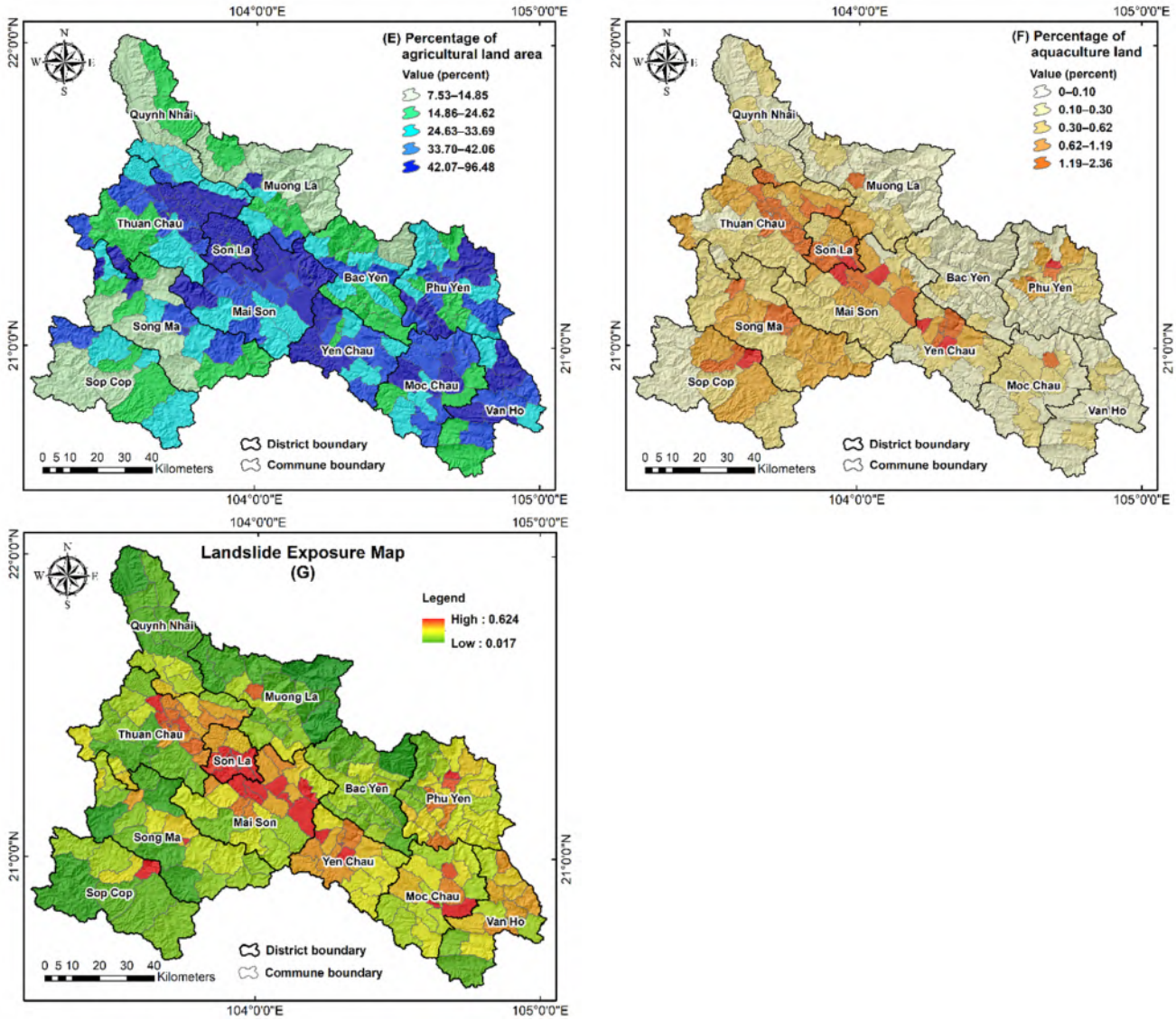

Figure 9. Landslide exposure maps: (**A**) population density, (**B**) the number of large primary and secondary schools, (**C**) percentage of residential land area, (**D**) length of roads, (**E**) percentage of agricultural land area, (**F**) percentage of aquaculture land, and (**G**) final landslide exposure map.

The main criteria of humans (E1) and infrastructure (E2) have relatively high weights, suggesting that residential areas and built environments are more vulnerable to landslides. In contrast, agriculture (E3) has a lower weight, indicating a reduced level of exposure. From a human perspective, population density is the sole indicator and plays a critical role. The residential land area emerges as the most significant indicator of infrastructure. However, the length of roads and the number of large primary and secondary schools also substantially contribute to assessing infrastructure exposure to landslides. In agriculture, the weights for agricultural land and aquaculture land areas are nearly identical, implying that both have a comparable impact on determining the exposure of agricultural zones to landslides.

The final landslide exposure map reveals that districts most exposed to landslides are located around major villages known for their tourist attractions and city centres where provincial administrative offices are concentrated. These areas generally have high population density and infrastructure, including Son La City, Mai Son, Yen Chau, Moc Chau, Van Ho, Phu Yen, and parts of Thuan Chau.

6.3. Landslide Vulnerability Analysis

The landslide vulnerability indicators and their respective weights are also calculated using the Iyengar–Sudarshan method, as presented in Table 7. These indicators were then combined using the weighted sum tool in the GIS environment, and the final maps generated from this process are illustrated in Figure 10.

Table 7. Landslide vulnerability criteria and sub-criteria weighting using the Iyengar–Sudarshan method.

Components	Main Criteria			Sub-Criteria		
	Label	Code	Weight	Label	Code	Weight
Vulnerability (V)	Population	V1	0.553	Percentage of children under 6 years old	V1-1	0.353
				Percentage of elderly people over 60 years old	V1-2	0.291
				Percentage of females between 17 and 59 years old	V1-3	0.356
	Infrastructure	V2	0.146	Percentage of unsolid houses	V2-1	1.000
	Demographics	V3	0.301	Birth rate	V3-1	0.465
				Death rate	V3-2	0.535

The population criterion (V1) has a significantly higher weight compared to the infrastructure (V2) and demographic (V3) criteria, reflecting the importance of population structure in determining landslide vulnerability. For the population group, the structure of the community (children under 6 years old, elderly people over 60 years old, females aged 17–59 years old) has a significant impact on vulnerability to landslide hazards. In the demographic aspect, the death rate has a slightly greater influence than the birth rate in determining vulnerability. A high death rate may reflect difficulties in living conditions and healthcare, leading to increased vulnerability. For the infrastructure aspect, the percentage of non-solid houses is the only indicator and plays a crucial role in representing the significant impact of housing conditions on vulnerability.

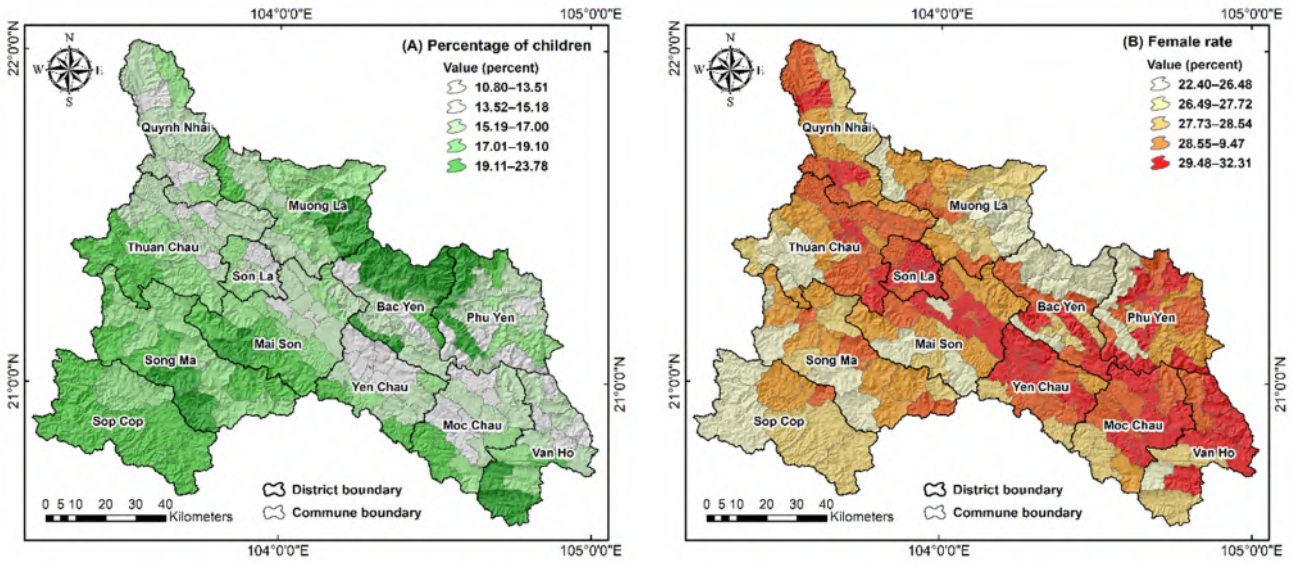

Figure 10. *Cont.*

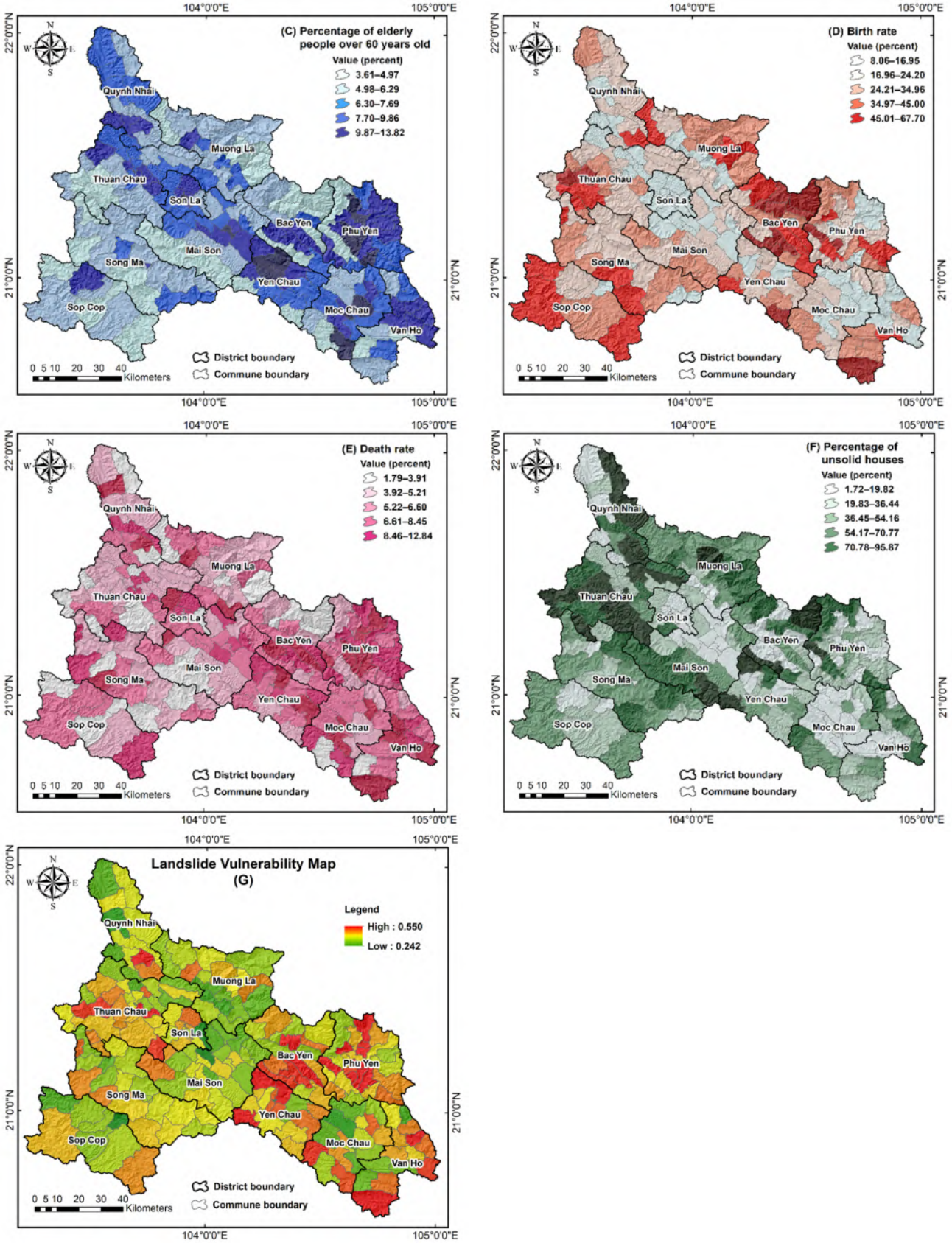

Figure 10. Landslide vulnerability maps: (**A**) percentage of children under 6 years old, (**B**) female ratio, (**C**) percentage of elderly people over 60 years old, (**D**) birth rate, (**E**) death rate, (**F**) percentage of unsolid houses, and (**G**) final landslide vulnerability map.

The final landslide vulnerability map indicates that districts with high vulnerability to landslide hazards are primarily concentrated in the southeastern districts of Son La Province, such as Bac Yen, Phu Yen, Yen Chau, parts of Moc Chau, and Van Ho. Similarly, the western districts also have high vulnerability to landslides, including Thuan Chau, parts of Song Ma, and Sop Cop. The north contains parts of the Quynh Nhai and Muong La districts. These are mostly areas with a higher population structure (children under 6 years old, elderly people over 60 years old, females aged 17–59 years old) and demographic indicators compared to other regions in the province.

6.4. Landslide Adaptive Capacity Analysis

The weights assigned to the landslide adaptive capacity indicators, as presented in Table 8, were combined using the weighted sum tool in the GIS workspace. The outcomes of this integration are visualized in Figure 11.

Table 8. Landslide adaptive capacity criteria and sub-criteria weighting using Iyengar–Sudarshan method.

Components	Main Criteria			Sub-Criteria		
	Label	**Code**	**Weight**	**Label**	**Code**	**Weight**
Adaptive Capacity (AC)	Healthcare	AC1	0.280	Number of medical facilities in each commune	AC1-1	1.000
	Infrastructure	AC2	0.436	Asphalt road density	AC2-1	0.387
				Number of high schools, universities, and colleges	AC2-2	0.353
				Number of enterprises	AC2-3	0.270
	Population	AC3	0.283	Percentage of population aged 17–59 years old	AC3-1	1.000

The infrastructure criterion (AC2) has a higher weight than the medical (AC1) and population (AC3) criteria, emphasizing the importance of physical infrastructure in enhancing adaptive capacity to landslide hazards. Asphalt road density is significantly more important in the infrastructure aspect than other sub-criteria. Well-developed asphalt roads facilitate quick response and relief activities during emergencies. The number of enterprises reflects economic capacity, supporting resilience through resource and infrastructure development. The number of high schools, universities, and colleges reflects community awareness and adaptive capacity. In terms of population, the percentage of the population aged 17–59 years old represents the working population that is actively involved in disaster preparedness and response, making it a key factor in adaptive capacity. For the medical aspect, the number of medical facilities in each commune is the sole indicator. It is crucial to enhance the community's adaptation and recovery capabilities in landslide events.

The final landslide adaptive capacity map demonstrates that the districts with higher coping capacity are primarily located in the central and eastern parts of Son La province, including Son La City, Yen Chau, Moc Chau, parts of Mai Son, Phu Yen, and Bac Yen. The north contains parts of Quynh Nhai and Muong La districts. In the west and Southwest are parts of the Thuan Chau, Song Ma, and Sop Cop districts. These areas generally have the highest physical infrastructure in the province, as indicated by the high density of asphalt roads, the number of enterprises, the number of high schools, universities, and colleges, and the percentage of solid houses. Additionally, these regions also have a significantly higher number of medical facilities in each commune and percentage of the population aged 17–59 years old compared to other districts in Son La Province.

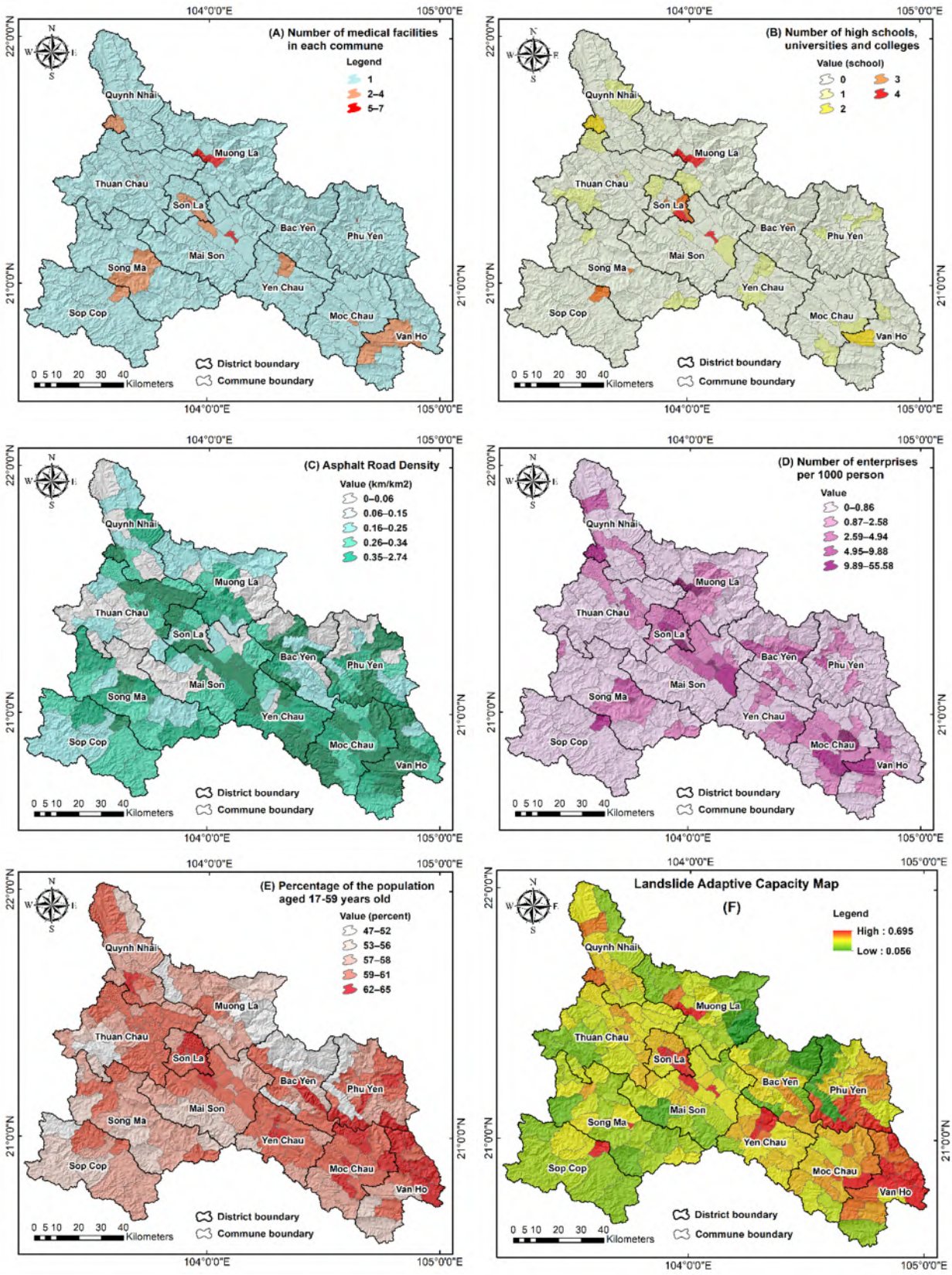

Figure 11. Landslide adaptive capacity maps: (**A**) number of medical facilities in each commune, (**B**) number of high schools, universities, and colleges, (**C**) asphalt road density, (**D**) number of enterprises per 1000 people, (**E**) percentage of the population aged 17–59 years old, and (**F**) final landslide adaptive capacity.

6.5. Landslide Risk Analysis

The four component maps–landslide hazard, exposure, vulnerability, and adaptive capacity—were normalized on a scale from 0 to 1 to maintain consistency in landslide risk assessment. These maps were then combined with equal weights in the GIS environment to produce the landslide risk assessment map for Son La province. The resulting map was then divided into five categories, including very low, low, moderate, high, and very high, by applying the quantile technique in the ArcGIS Pro software (Figure 12).

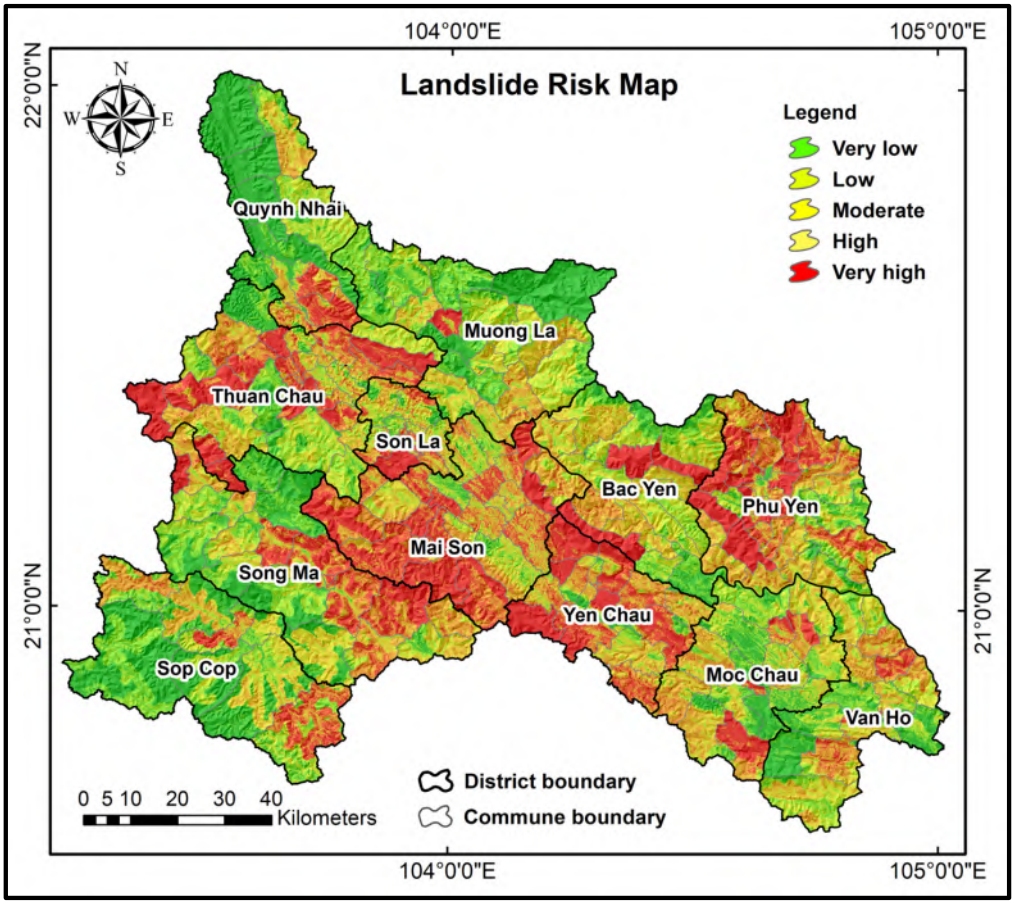

Figure 12. Landslide risk assessment map of Son La province.

The statistical analysis from the landslide risk map indicates that 5673.84 km^2 (40.21%) falls into the very high-risk and high-risk regions, 2979.64 km^2 (21.12%) falls into the moderate-risk region, and 5456.64 km^2 (38.67%) into the low-risk and very low-risk regions (Figure 13). The landslide risk map reveals that the districts facing the highest risk are predominantly located in the central and northeastern parts of Son La province, including Mai Son, Phu Yen, Thuan Chau, Yen Chau, Song Ma, and Bac Yen districts. This finding indicates that the districts with the highest risk exhibit elevated levels of landslide hazard, exposure, and vulnerability despite varying levels of adaptive capacity. These high-risk districts are frequently centred around urban and village areas, characterized by dense populations, particularly among vulnerable groups such as children under 6 years old, elderly individuals over 60 years old, and women aged 17–59. Additionally, these districts have higher demographic metrics compared to others in the province. Meanwhile, most of the areas in the Quynh Nhai and Muong La districts are recognized for their lower risk due to their minimal landslide exposure. Although Son La City has a high level of landslide exposure and vulnerability, it is located in a very low landslide susceptibility zone and has a high adaptive capacity to landslides. Additionally, due to its relatively small area, the landslide risk area of this district is not large.

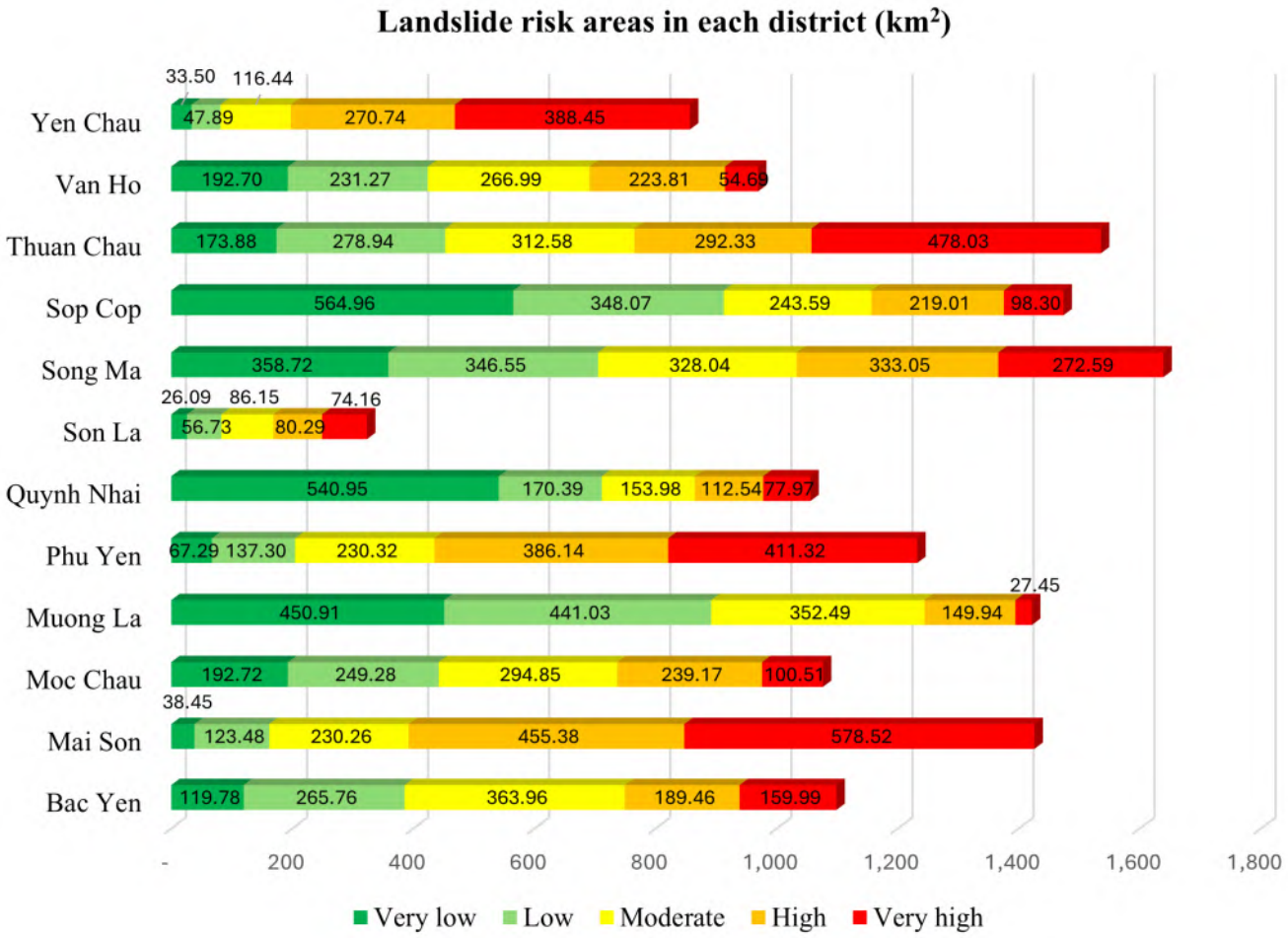

Figure 13. Statistics on landslide risk areas in each district within Son La province.

7. Discussion

Landslides are one of the most significant hazards with large-scale socio-economic and environmental impacts causing dangers to sustainable socio-economic development in mountainous areas [96]. In mountainous regions, steep terrain and unstable soil increase the likelihood of landslides, especially during heavy rainfall or seismic activity. With climate change exacerbating the severity and frequency of severe weather events, the likelihood of landslides in these areas is projected to increase, presenting further difficulties for disaster management and community resilience [97]. In Vietnam, landslide events frequently occur in mountainous regions when heavy rainfall during the rainy season triggers slope failures [38]. In addition to natural triggers, human activities such as deforestation, road construction, and unregulated mining exacerbate the landslide risk across South Asia [98]. The damage to the economy and the loss of human lives from landslides are considerably more extensive than commonly recognized [99]. The landslide risk assessment can support local authorities and communities in proactively managing and mitigating the impacts of landslides and contribute to safeguarding humans, the living environment, and natural resources [100]. The current study proposed a holistic approach to assess the landslide risk in Son La province, Vietnam, by integrating the advanced ML ensemble models and the Iyengar–Sudarshan technique.

The growing complexity and unpredictability of landslide hazards driven by climate change highlight the critical need for more advanced risk assessment methods [101]. Typically, these approaches are applied separately in creating landslide risk assessments. Some research has focused on using advanced ML hybrid models to generate landslide risk maps based on various approaches, like combining the landslide hazard map with official

population and building census data [12], integrating the best landslide susceptibility model with the estimated rainfall [22], or combining pipeline vulnerability and landslide susceptibility [23]. As far as we know, there has been no study to conduct a landslide risk assessment that considers landslide hazard, exposure, vulnerability, and adaptive capacity by integrating ML models and the Iyenga–Sudarshan technique.

In this study, we proposed an integrated framework combining ML models for landslide susceptibility with the Iyenga–Sudarshan technique in landslide risk assessments. Landslide inventory data are crucial in susceptibility mapping since they supply a holistic database of historical landslide events [50]. A set of 1771 landslide points was gathered using information from the Vietnam Institute of Geosciences and Mineral Resources website, field surveys, and Google Earth image analysis to provide a foundation for landslide risk prediction. At the same time, fifteen landslide-influencing factors were selected based on available databases that could be collected, including elevation, slope, slope direction, curvature, terrain roughness, the Terrain Wetness Index (TWI), rainfall, stream density, road density, distance to the road, distance to the river, hydrogeology, geology, geomorphology, and land cover. All landslide-influencing factors were converted to the raster format with a spatial resolution of 30 m using the ArcGIS Pro environment. The landslide inventory points and fifteen landslide-affecting factors were used as input data for ML models to build landslide susceptibility maps. Six ML ensemble models comprised Decorate-UB, Dagging-UB, Bagging-UB, MultiScheme-UB, Cascade Generalization-UB, and MultiBoostAB-UB and were applied to build landslide susceptibility maps. The landslide susceptibility map that demonstrated the highest accuracy through validation was selected as the landslide hazard map. The landslide exposure, vulnerability, and adaptive capacity maps were developed by gathering a range of socio-economic and physical indicators. These indicators representing landslide exposure, vulnerability, and adaptive capacity for each commune in the study area were collected from various sources, including the 2022 statistical yearbooks of the 11 districts within the province and the National Data Portal website (https://data.gov.vn/SitePages/Index.aspx#/index, accessed on 1 October 2024). These indicators' weights were established using the Iyengar–Sudarshan method, which led to the creation of maps showing their exposure, vulnerability, and ability to adapt to landslides. Finally, the landslide hazard, exposure, vulnerability, and adaptive capacity maps with their respective weights were integrated for spatial analysis to develop a landslide risk map for the study area.

The frequency and severity of landslide events are sometimes exacerbated by human activities such as road construction, deforestation, agricultural expansion, and construction activities, leading to increasing landslide risk [102]. Meanwhile, a landslide risk approach built on the combination of landslide hazard, exposure, vulnerability, and adaptive capacity provides a holistic approach to landslide risk, accounting for not just the landslide hazard itself but also how exposure, vulnerability, and adaptive capacity influence overall landslide risk. The landslide risk framework proposed in this study offers an effective platform for disaster mitigation management efforts by integrating hazards, exposure, vulnerability, and adaptive capacity, allowing for more accurate and targeted risk evaluations. Furthermore, this approach can be applied to areas of various scales facing different disaster risk types.

8. Conclusions

This research presents a comprehensive framework that combines cutting-edge ML models with the Iyengar and Sudarshan approach to evaluate landslide risk in Son La province, a mountainous region in northwestern Vietnam. This study developed six advanced hybrid ML models—Decorate UB (DCUB), Dagging UB (DGUB), Bagging UB (BGUB), MultiScheme UB (MSUB), Cascade Generalization UB (CGUB), and MultiBoostAB UB (MBUB)—all leveraging UltraBoost (UB) as the foundational classifier to generate landslide susceptibility maps. The landslide hazard map was identified by selecting the most accurate susceptibility map through cross-validation and AUC analysis. The Iyengar–Sudarshan method was employed to assign appropriate weights to landslide exposure,

vulnerability, and adaptive capacity indicators. Finally, these four essential components—landslide hazard, exposure, vulnerability, and adaptive capacity—were integrated into a GIS platform, creating a precise and thorough landslide risk assessment map. The obtained result is a highly detailed and reliable landslide risk assessment map, offering a powerful tool for disaster preparedness and mitigation for mountainous provinces in Vietnam. Although this theoretical framework can be applied to different regions in Vietnam, the computational process is quite complex. It requires an in-depth understanding of machine learning techniques and time-consuming calculations. Future research will focus on monitoring and evaluating landslide developments in real time to comprehensively assess the level of impact of landslide risks on the socio-economic aspects in Son La province. In reality, these obtained results will provide important information in planning targeted development, allocating resources, protecting vulnerable communities, and investing in technical infrastructure for priority areas, such as the central and northeastern regions of Son La province, including Mai Son, Phu Yen, Thuan Chau, Yen Chau, Song Ma, and Bac Yen districts.

Author Contributions: Conceptualization, T.X.T., S.L., H.H., Q.D.B. and C.L.; Methodology, T.X.T., S.L. and C.L.; Software, T.X.T., Q.D.B., L.Q.N., D.Q.N. and C.-T.T.; Validation, T.X.T., H.H. and Q.D.B.; Formal analysis, T.X.T., H.H., Q.D.B., L.Q.N., D.Q.N. and C.-T.T.; Investigation, T.X.T., H.H., L.Q.N., D.Q.N. and C.-T.T.; Resources, T.X.T., H.H., Q.D.B. and C.L.; Data curation, T.X.T., H.H., Q.D.B. and C.L.; Writing—original draft, T.X.T., Q.D.B. and C.-T.T.; Writing—review & editing, T.X.T. and Q.D.B.; Visualization, T.X.T., Q.D.B. and C.L.; Supervision, S.L. and C.L.; Project administration, H.H.; Funding acquisition, H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by the Hanoi University of Civil Engineering (HUCE) under grant number 13-NNC-ĐHXDHN.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Hussain, M.A.; Zhang, S.; Muneer, M.; Moawwez, M.A.; Kamran, M.; Ahmed, E. Assessing and mapping spatial variation characteristics of natural hazards in Pakistan. *Land* **2022**, *12*, 140. [CrossRef]
2. Froude, M.J.; Petley, D.N. Global fatal landslide occurrence from 2004 to 2016. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 2161–2181. [CrossRef]
3. Kirschbaum, D.; Stanley, T.; Zhou, Y. Spatial and temporal analysis of a global landslide catalog. *Geomorphology* **2015**, *249*, 4–15. [CrossRef]
4. Barman, J.; Soren, D.D.L.; Biswas, B. Landslide susceptibility evaluation and analysis: A review on articles published during 2000 to 2020. In *Monitoring and Managing Multi-Hazards: A Multidisciplinary Approach*; Springer: Cham, Switzerland, 2022; pp. 211–220.
5. Lu, W.; Xiao, Z.; Chen, Y.; Sun, J.; Chen, F. Spatiotemporal Characteristics and Rainfall Thresholds of Geological Landslide Disasters in ASEAN Countries. *Atmosphere* **2024**, *15*, 599. [CrossRef]
6. Kjekstad, O.; Highland, L. Economic and social impacts of landslides. In *Landslides–Disaster Risk Reduction*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 573–587. [CrossRef]
7. Jaedicke, C.; Van Den Eeckhaut, M.; Nadim, F.; Hervás, J.; Kalsnes, B.; Vangelsten, B.V.; Smith, J.T.; Tofani, V.; Ciurean, R.; Winter, M.G. Identification of landslide hazard and risk 'hotspots' in Europe. *Bull. Eng. Geol. Environ.* **2014**, *73*, 325–339. [CrossRef]
8. Huang, R.; Li, W. Formation, distribution and risk control of landslides in China. *J. Rock Mech. Geotech. Eng.* **2011**, *3*, 97–116. [CrossRef]
9. Tiwari, B.; Ajmera, B.; Dhital, S. Characteristics of moderate-to large-scale landslides triggered by the M w 7.8 2015 Gorkha earthquake and its aftershocks. *Landslides* **2017**, *14*, 1297–1318. [CrossRef]
10. Bozzolan, E.; Holcombe, E.A.; Pianosi, F.; Marchesini, I.; Alvioli, M.; Wagener, T. A mechanistic approach to include climate change and unplanned urban sprawl in landslide susceptibility maps. *Sci. Total Environ.* **2023**, *858*, 159412. [CrossRef]

11. Van Westen, C.J.; Castellanos, E.; Kuriakose, S.L. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Eng. Geol.* **2008**, *102*, 112–131. [CrossRef]
12. Novellino, A.; Cesarano, M.; Cappelletti, P.; Di Martire, D.; Di Napoli, M.; Ramondini, M.; Sowter, A.; Calcaterra, D. Slow-moving landslide risk assessment combining Machine Learning and InSAR techniques. *Catena* **2021**, *203*, 105317. [CrossRef]
13. Sun, D.; Wen, H.; Zhang, Y.; Xue, M. An optimal sample selection-based logistic regression model of slope physical resistance against rainfall-induced landslide. *Nat. Hazards* **2021**, *105*, 1255–1279. [CrossRef]
14. Corominas, J.; van Westen, C.; Frattini, P.; Cascini, L.; Malet, J.-P.; Fotopoulou, S.; Catani, F.; Van Den Eeckhaut, M.; Mavrouli, O.; Agliardi, F. Recommendations for the quantitative analysis of landslide risk. *Bull. Eng. Geol. Environ.* **2014**, *73*, 209–263. [CrossRef]
15. Tahmasebi, P.; Kamrava, S.; Bai, T.; Sahimi, M. Machine learning in geo-and environmental sciences: From small to large scale. *Adv. Water Resour.* **2020**, *142*, 103619. [CrossRef]
16. Aslam, B.; Zafar, A.; Khalil, U. Development of integrated deep learning and machine learning algorithm for the assessment of landslide hazard potential. *Soft Comput.* **2021**, *25*, 13493–13512. [CrossRef]
17. Mohan, A.; Singh, A.K.; Kumar, B.; Dwivedi, R. Review on remote sensing methods for landslide detection using machine and deep learning. *Trans. Emerg. Telecommun. Technol.* **2021**, *32*, e3998. [CrossRef]
18. Efeoğlu, E. Use of Machine Learning Algorithms in Location Determination for Safe Construction. *Int. J. Appl. Methods Electron. Comput.* **2023**, *11*, 197–202. [CrossRef]
19. Moustafa, A.F.; Cary, T.W.; Sultan, L.R.; Schultz, S.M.; Conant, E.F.; Venkatesh, S.S.; Sehgal, C.M. Color doppler ultrasound improves machine learning diagnosis of breast cancer. *Diagnostics* **2020**, *10*, 631. [CrossRef]
20. Syam, V.; Safal, S.; Bhutia, O.; Singh, A.K.; Giri, D.; Bhandari, S.S.; Panigrahi, R. A non-invasive method for prediction of neurodegenerative diseases using gait signal features. *Procedia Comput. Sci.* **2023**, *218*, 1529–1541. [CrossRef]
21. Costache, R.; Arabameri, A.; Costache, I.; Crăciun, A.; Pham, B.T. New machine learning ensemble for flood susceptibility estimation. *Water Resour. Manag.* **2022**, *36*, 4765–4783. [CrossRef]
22. Mallick, J.; Alqadhi, S.; Talukdar, S.; AlSubih, M.; Ahmed, M.; Khan, R.A.; Kahla, N.B.; Abutayeh, S.M. Risk assessment of resources exposed to rainfall induced landslide with the development of GIS and RS based ensemble metaheuristic machine learning algorithms. *Sustainability* **2021**, *13*, 457. [CrossRef]
23. Wen, H.; Liu, L.; Zhang, J.; Hu, J.; Huang, X. A hybrid machine learning model for landslide-oriented risk assessment of long-distance pipelines. *J. Environ. Manag.* **2023**, *342*, 118177. [CrossRef]
24. Alcántara-Ayala, I. Integrated landslide disaster risk management (ILDRiM): The challenge to avoid the construction of new disaster risk. *Environ. Hazards* **2021**, *20*, 323–344. [CrossRef]
25. Feyissa, G.; Zeleke, G.; Gebremariam, E.; Bewket, W. GIS based quantification and mapping of climate change vulnerability hotspots in Addis Ababa. *Geoenviron. Disasters* **2018**, *5*, 14. [CrossRef]
26. Mekonen, A.A.; Berlie, A.B. Rural households' livelihood vulnerability to climate variability and extremes: A livelihood zone-based approach in the Northeastern Highlands of Ethiopia. *Ecol. Process.* **2021**, *10*, 55. [CrossRef]
27. Lien, M.K. Vulnerability Assessment of Climate Change on Sea Level Rise Impacts on Some Economic Sectors in Binh Dinh Province, Vietnam. *Am. J. Clim. Chang.* **2019**, *8*, 302–324. [CrossRef]
28. Laitonjam, N.; Singh, R.; Feroze, S. Vulnerability to climate change: Review of conceptual framework. *Econ. Aff.* **2018**, *63*, 473–479. [CrossRef]
29. Majumder, S.; Roy, S.; Bose, A.; Chowdhury, I.R. Multiscale GIS based-model to assess urban social vulnerability and associated risk: Evidence from 146 urban centers of Eastern India. *Sustain. Cities Soc.* **2023**, *96*, 104692. [CrossRef]
30. Iyengar, N.S.; Sudarshan, P. A method of classifying regions from multivariate data. *Econ. Political Wkly.* **1982**, *17*, 2047–2052.
31. Erena, S.H.; Worku, H. Urban flood vulnerability assessments: The case of Dire Dawa city, Ethiopia. *Nat. Hazards* **2019**, *97*, 495–516. [CrossRef]
32. Kumar, D.; Bhattacharjya, R.K. Estimation of Integrated Flood Vulnerability Index for the Hilly Region of Uttarakhand, India. *J. Hazard. Toxic Radioact. Waste* **2020**, *24*, 04020051. [CrossRef]
33. Jha, D.K.; Bhattacharyya, R.K.; Shyam, S.; Ratnayke, U.R. Indicator based assessment of integrated flood vulnerability index for Brunei Darussalam. *Int. J. Disaster Risk Manag.* **2020**, *2*, 47–70. [CrossRef]
34. Murthy, C.; Singh, J.; Kumar, P.; Sesha Sai, M. A composite index for drought hazard assessment using CPC rainfall time series data. *Int. J. Environ. Sci. Technol.* **2017**, *14*, 1981–1988. [CrossRef]
35. Nguyen, T.T.; Ngo, H.H.; Guo, W.; Nguyen, H.Q.; Luu, C.; Dang, K.B.; Liu, Y.; Zhang, X. New approach of water quantity vulnerability assessment using satellite images and GIS-based model: An application to a case study in Vietnam. *Sci. Total Environ.* **2020**, *737*, 139784. [CrossRef] [PubMed]
36. Chau, V.N.; Holland, J.; Cassells, S. Institutional structures underpinning flood management in Vietnam. *Int. J. Disaster Risk Reduct.* **2014**, *10*, 341–348. [CrossRef]
37. Tran, V.T.; An-Vo, D.-A.; Cockfield, G.; Mushtaq, S. Assessing livelihood vulnerability of minority ethnic groups to climate change: A case study from the northwest mountainous regions of Vietnam. *Sustainability* **2021**, *13*, 7106. [CrossRef]
38. Bui, Q.D.; Ha, H.; Khuc, D.T.; Nguyen, D.Q.; von Meding, J.; Nguyen, L.P.; Luu, C. Landslide susceptibility prediction mapping with advanced ensemble models: Son La province, Vietnam. *Nat. Hazards* **2023**, *116*, 2283–2309. [CrossRef]

39. Centre, A. Viet Nam, Flooding, Landslide and Whirlwinds in Son La Province (24 August 2021). Available online: https://reliefweb.int/report/viet-nam/viet-nam-flooding-landslide-and-whirlwinds-son-la-province-24-aug-2021 (accessed on 12 October 2024).
40. IFRC. Viet Nam, Asia-Pacific | Typhoon Yagi-Operation Update #1 (MDRVN024). Available online: https://reliefweb.int/report/viet-nam/viet-nam-asia-pacific-typhoon-yagi-operation-update-1-mdrvn024 (accessed on 12 October 2024).
41. Gallopín, G.C. Linkages between vulnerability, resilience, and adaptive capacity. *Glob. Environ. Chang.* **2006**, *16*, 293–303. [CrossRef]
42. Polsky, C.; Neff, R.; Yarnal, B. Building comparable global change vulnerability assessments: The vulnerability scoping diagram. *Glob. Environ. Chang.* **2007**, *17*, 472–485. [CrossRef]
43. Ciurean, R.L.; Schröter, D.; Glade, T. Conceptual frameworks of vulnerability assessments for natural disasters reduction. In *Approaches to Disaster Management-Examining the Implications of Hazards, Emergencies and Disasters*; Intechopen: London, UK, 2013; pp. 1–32.
44. UNDRR. *Sendai Framework for Disaster Risk Reduction 2015–2030*; UNDRR: Geneva, Switzerland, 2015.
45. Adelekan, I.; Cartwright, A.; Chow, W.; Colenbrander, S.; Dawson, R.; Garschagen, M.; Haasnoot, M.; Hashizume, M.; Klaus, I.; Krishnaswamy, J.; et al. *What the Latest Science on Impacts, Adaptation and Vulnerability Means for Cities and Urban Areas*; The Ar6 Summary For Urban Policymakers Series; Indian Institute for Human Settlement: Bengaluru, India, 2022.
46. Van Westen, C.; Van Asch, T.W.; Soeters, R. Landslide hazard and risk zonation—Why is it still so difficult? *Bull. Eng. Geol. Environ.* **2006**, *65*, 167–184. [CrossRef]
47. Pellicani, R.; Van Westen, C.J.; Spilotro, G. Assessing landslide exposure in areas with limited landslide information. *Landslides* **2014**, *11*, 463–480. [CrossRef]
48. Antronico, L.; De Pascale, F.; Coscarelli, R.; Gullà, G. Landslide risk perception, social vulnerability and community resilience: The case study of Maierato (Calabria, southern Italy). *Int. J. Disaster Risk Reduct.* **2020**, *46*, 101529. [CrossRef]
49. Shah, M.A.R.; Renaud, F.G.; Anderson, C.C.; Wild, A.; Domeneghetti, A.; Polderman, A.; Votsis, A.; Pulvirenti, B.; Basu, B.; Thomson, C.; et al. A review of hydro-meteorological hazard, vulnerability, and risk assessment frameworks and indicators in the context of nature-based solutions. *Int. J. Disaster Risk Reduct.* **2020**, *50*, 101728. [CrossRef]
50. Kavzoglu, T.; Colkesen, I.; Sahin, E.K. Machine learning techniques in landslide susceptibility mapping: A survey and a case study. In *Landslides: Theory, Practice and Modelling*; Springer: Cham, Switzerland, 2019; pp. 283–301.
51. Luino, F.; Barriendos, M.; Gizzi, F.T.; Glaser, R.; Gruetzner, C.; Palmieri, W.; Porfido, S.; Sangster, H.; Turconi, L. Historical data for natural hazard risk mitigation and land use planning. *Land* **2023**, *12*, 1777. [CrossRef]
52. McColl, S.T. Chapter 2—Landslide causes and triggers. In *Landslide Hazards, Risks, and Disasters*, 2nd ed.; Davies, T., Rosser, N., Shroder, J.F., Eds.; Elsevier: Amsterdam, The Netherlands, 2022; pp. 13–41.
53. Cevasco, A.; Pepe, G.; Brandolini, P. The influences of geological and land use settings on shallow landslides triggered by an intense rainfall event in a coastal terraced environment. *Bull. Eng. Geol. Environ.* **2014**, *73*, 859–875. [CrossRef]
54. Mario, B.; Giacomo, P.; Paola, G.; Giuseppe, P.; Marco, P.; Katia, S.; Francesco, C. Landslides and predisposing factors of the Southern Apennines, Italy. *J. Maps* **2023**, *19*, 2137065. [CrossRef]
55. Abraham, M.T.; Satyam, N.; Lokesh, R.; Pradhan, B.; Alamri, A. Factors affecting landslide susceptibility mapping: Assessing the influence of different machine learning approaches, sampling strategies and data splitting. *Land* **2021**, *10*, 989. [CrossRef]
56. Ha, H.; Bui, Q.D.; Tran, D.T.; Nguyen, D.Q.; Bui, H.X.; Luu, C. Improving the forecast performance of landslide susceptibility mapping by using ensemble gradient boosting algorithms. *Environ. Dev. Sustain.* **2024**, 1–35. [CrossRef]
57. Pourghasemi, H.R.; Moradi, H.R.; Fatemi Aghda, S.M. Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances. *Nat. Hazards* **2013**, *69*, 749–779. [CrossRef]
58. Kyriazos, T.; Poga, M. Dealing with multicollinearity in factor analysis: The problem, detections, and solutions. *Open J. Stat.* **2023**, *13*, 404–424. [CrossRef]
59. Senaviratna, N.A.M.R.; Cooray, T.M.J. Diagnosing Multicollinearity of Logistic Regression Model. *Asian J. Probab. Stat.* **2019**, *5*, 1–9. [CrossRef]
60. Raheem, M.A.; Udoh, N.S.; Gbolahan, A.T. Choosing Appropriate Regression Model in the Presence of Multicolinearity. *Open J. Stat.* **2019**, *9*, 159–168. [CrossRef]
61. Mulyanto, A.D. mVIF Package: A Tool for Detecting Multicollinearity Without Dependent Variables. *Matics J. Ilmu Komput. Dan Teknol. Inf. (J. Comput. Sci. Inf. Technol.)* **2022**, *14*, 70–73. [CrossRef]
62. Melville, P.; Mooney, R.J. Constructing diverse classifier ensembles using artificial training examples. In Proceedings of the IJCAI International Joint Conference on Artificial Intelligence, Acapulco, Mexico, 9–15 August 2003; pp. 505–510.
63. Akour, M.; Banitaan, S.; Alsghaier, H.; Al Radaideh, K. Predicting daily activities effectiveness using base-level and meta level classifiers. In Proceedings of the 2019 7th International Symposium on Digital Forensics and Security (ISDFS), Barcelos, Portugal, 10–12 June 2019; pp. 1–7.
64. Ting, K.M.; Witten, I.H. *Stacking Bagged and Dagged Models*; (Working Paper 97/09); Department of Computer Science, University of Waikato: Hamilton, New Zealand, 1997.
65. Morteza, A.; Sadipour, M.; Fard, R.S.; Taheri, S.; Ahmadi, A. A dagging-based deep learning framework for transmission line flexibility assessment. *IET Renew. Power Gener.* **2023**, *17*, 1092–1105. [CrossRef]
66. Breiman, L. Bagging predictors. *Mach Learn* **1996**, *24*, 123–140. [CrossRef]

67. Ribeiro, M.H.D.M.; dos Santos Coelho, L. Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Appl. Soft Comput.* **2020**, *86*, 105837. [CrossRef]
68. Hong, H. Assessing landslide susceptibility based on hybrid Best-first decision tree with ensemble learning model. *Ecol. Indic.* **2023**, *147*, 109968. [CrossRef]
69. Nidhi, N.; Kumar, M.; Agarwal, S. Comparative Analysis of Heterogeneous Ensemble Learning using Feature Selection Techniques for Predicting Academic Performance of Students. In Proceedings of the 2021 2nd International Conference on Computational Methods in Science & Technology (ICCMST), Mohali, India, 17–18 December 2021; pp. 212–217.
70. Ado, M.; Amitab, K.; Maji, A.K.; Jasińska, E.; Gono, R.; Leonowicz, Z.; Jasiński, M. Landslide susceptibility mapping using machine learning: A literature survey. *Remote Sens.* **2022**, *14*, 3029. [CrossRef]
71. Webb, G.I. Multiboosting: A technique for combining boosting and wagging. *Mach. Learn.* **2000**, *40*, 159–196. [CrossRef]
72. Gama, J.; Brazdil, P. Cascade generalization. *Mach. Learn.* **2000**, *41*, 315–343. [CrossRef]
73. Kavzoglu, T.; Teke, A. Predictive Performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (XGBoost) and natural gradient boosting (NGBoost). *Arab. J. Sci. Eng.* **2022**, *47*, 7367–7385. [CrossRef]
74. Orhan, O.; Bilgilioglu, S.S.; Kaya, Z.; Ozcan, A.K.; Bilgilioglu, H. Assessing and mapping landslide susceptibility using different machine learning methods. *Geocarto Int.* **2022**, *37*, 2795–2820. [CrossRef]
75. Hosenuzzaman, M.; Kibria, M.G.; Sarkar, R.; Abedin, M.A. Landslide, agricultural vulnerability, and community initiatives: A case study in South-East part of Bangladesh. In *Impact of Climate Change, Land Use and Land Cover, and Socio-Economic Dynamics on Landslides*; Springer: Singapore, 2022; pp. 123–145.
76. Modugno, S.; Johnson, S.; Borrelli, P.; Alam, E.; Bezak, N.; Balzter, H. Analysis of human exposure to landslides with a GIS multiscale approach. *Nat. Hazards* **2022**, *112*, 387–412. [CrossRef]
77. Hongtao, Y.; Ting, M. Changes in the geographical distributions of global human settlements. *J. Resour. Ecol.* **2021**, *12*, 829–839. [CrossRef]
78. Amini Hosseini, K.; Izadkhah, Y.O. From "Earthquake and safety" school drills to "safe school-resilient communities": A continuous attempt for promoting community-based disaster risk management in Iran. *Int. J. Disaster Risk Reduct.* **2020**, *45*, 101512. [CrossRef]
79. Rofiah, N.H.; Kawai, N.; Hayati, E.N. Key elements of disaster mitigation education in inclusive school setting in the Indonesian context. *Jàmbá J. Disaster Risk Stud.* **2021**, *13*, 1159. [CrossRef] [PubMed]
80. Nikoo, N.; Babaei, M.; Mohaymany, A.S. Emergency transportation network design problem: Identification and evaluation of disaster response routes. *Int. J. Disaster Risk Reduct.* **2018**, *27*, 7–20. [CrossRef]
81. Zhou, M.; Yuan, M.; Yang, G.; Mei, G. Risk analysis of road networks under the influence of landslides by considering landslide susceptibility and road vulnerability: A case study. *Nat. Hazards Res.* **2023**, *4*, 387–400. [CrossRef]
82. Pham, N.T.T.; Nong, D.; Sathyan, A.R.; Garschagen, M. Vulnerability assessment of households to flash floods and landslides in the poor upland regions of Vietnam. *Clim. Risk Manag.* **2020**, *28*, 100215. [CrossRef]
83. Ho, N.N.; Lai, P.T.; Truong, T.C.A.; Do, T.T.; Nguyen, T. The contribution of livelihood diversification activities to poverty reduction of ethnic minority households: A case study in Son La Province, Vietnam. *J. Infrastruct. Policy Dev.* **2024**, *8*, 6465. [CrossRef]
84. Salvati, P.; Petrucci, O.; Rossi, M.; Bianchi, C.; Pasqua, A.A.; Guzzetti, F. Gender, age and circumstances analysis of flood and landslide fatalities in Italy. *Sci. Total Environ.* **2018**, *610–611*, 867–879. [CrossRef]
85. Nor Diana, M.I.; Muhamad, N.; Taha, M.R.; Osman, A.; Alam, M.M. Social vulnerability assessment for landslide hazards in Malaysia: A systematic review study. *Land* **2021**, *10*, 315. [CrossRef]
86. Ciciolla, L.; Luthar, S.S. Invisible household labor and ramifications for adjustment: Mothers as captains of households. *Sex Roles* **2019**, *81*, 467–486. [CrossRef] [PubMed]
87. Pham, N.T.T.; Nong, D.; Garschagen, M. Natural hazard's effect and farmers' perception: Perspectives from flash floods and landslides in remotely mountainous regions of Vietnam. *Sci. Total Environ.* **2021**, *759*, 142656. [CrossRef] [PubMed]
88. Daramola, A.Y.; Oni, O.T.; Ogundele, O.; Adesanya, A. Adaptive capacity and coping response strategies to natural disasters: A study in Nigeria. *Int. J. Disaster Risk Reduct.* **2016**, *15*, 132–147. [CrossRef]
89. Tran, Q.A.; Le Tu, H.; Hoang Van, M. Availability and readiness of communal health services: Results from 2015 Vietnam District and Commune Health Facility Survey. *Int. J. Healthc. Manag.* **2021**, *14*, 1206–1212. [CrossRef]
90. Deng, Z.; Li, W.; Dong, W.; Sun, Z.; Kodikara, J.; Sheng, D. Multifunctional asphalt concrete pavement toward smart transport infrastructure: Design, performance and perspective. *Compos. Part B Eng.* **2023**, *265*, 110937. [CrossRef]
91. Ozkazanc, S.; Yuksel, U.D. Evaluation of Disaster Awareness and Sensitivity Level of Higher Education Students. *Procedia-Soc. Behav. Sci.* **2015**, *197*, 745–753. [CrossRef]
92. Linnenluecke, M.K.; McKnight, B. Community resilience to natural disasters: The role of disaster entrepreneurship. *J. Enterprising Communities People Places Glob. Econ.* **2017**, *11*, 166–185. [CrossRef]
93. Sun, F.; Li, H.; Cai, J.; Hu, S.; Xing, H. Examining organizational collaboration and resource flows of disaster response system based on a time-dynamic perspective. *Int. J. Disaster Risk Reduct.* **2024**, *108*, 104565. [CrossRef]
94. Perera, E.; Jayawardana, D.; Jayasinghe, P.; Ranagalage, M. Landslide vulnerability assessment based on entropy method: A case study from Kegalle district, Sri Lanka. *Model. Earth Syst. Environ.* **2019**, *5*, 1635–1649. [CrossRef]

95. Connor, R.; Hiroki, K. Development of a method for assessing flood vulnerability. *Water Sci. Technol.* **2005**, *51*, 61–67. [CrossRef] [PubMed]
96. Mestanza-Ramón, C.; Ordoñez-Alcivar, R.; Arguello-Guadalupe, C.; Carrera-Silva, K.; D'Orio, G.; Straface, S. History, socioeconomic problems and environmental impacts of gold mining in the Andean Region of Ecuador. *Int. J. Environ. Res. Public Health* **2022**, *19*, 1190. [CrossRef] [PubMed]
97. Gariano, S.L.; Guzzetti, F. Landslides in a changing climate. *Earth-Sci. Rev.* **2016**, *162*, 227–252. [CrossRef]
98. Chang, M.; Dou, X.; Zhu, X.; Ma, Y. Integrated risk assessment of landslide in karst terrains: Advancing landslides management in Beiliu City, China. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *132*, 104046. [CrossRef]
99. Alam, A.; Ahmed, B.; Sammonds, P.; Maksud Kamal, A.S.M. Applying rainfall threshold estimates and frequency ratio model for landslide hazard assessment in the coastal mountain setting of South Asia. *Nat. Hazards Res.* **2023**, *3*, 531–545. [CrossRef]
100. Alam, E.; Ray-Bennett, N.S. Disaster risk governance for district-level landslide risk management in Bangladesh. *Int. J. Disaster Risk Reduct.* **2021**, *59*, 102220. [CrossRef]
101. Zeng, T.; Gong, Q.; Wu, L.; Zhu, Y.; Yin, K.; Peduto, D. Double-index rainfall warning and probabilistic physically based model for fast-moving landslide hazard analysis in subtropical-typhoon area. *Landslides* **2024**, *21*, 753–773. [CrossRef]
102. Hoang, N.-D.; Tien Bui, D. Spatial prediction of rainfall-induced shallow landslides using gene expression programming integrated with GIS: A case study in Vietnam. *Nat. Hazards* **2018**, *92*, 1871–1887. [CrossRef]