remote sensing

Article

Different-Classification-Scheme-Based Machine Learning Model of Building Seismic Resilience Assessment in a Mountainous Region

Haijia Wen [1,*], Xinzhi Zhou [2], Chi Zhang [1], Mingyong Liao [1] and Jiafeng Xiao [1]

1 National Joint Engineering Research Center of Geohazards Prevention in the Reservoir Areas, Key Laboratory of New Technology for Construction of Cities in Mountain Area, Ministry of Education Chongqing, School of Civil Engineering, Chongqing University, Chongqing 400045, China
2 State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100083, China
* Correspondence: jhw@cqu.edu.cn; Tel.: +86-13251321327

Abstract: This study aims to develop different-classification-scheme-based building-seismic-resilience (BSR)-mapping models using random forest (RF) and a support vector machine (SVM). Based on a field survey of earthquake-damaged buildings in Shuanghe Town, the epicenter of the Changning M 5.8 earthquake that occurred on 17 June 2019, we selected 19 influencing factors for BSR assessment to establish a database. Based on three classification schemes for the description of BSR, we developed six machine learning assessment models for BSR mapping using RF and an SVM after optimizing the hyper-parameters. The validation indicators of model performance include precision, recall, accuracy, and F1-score as determined from the test sub-dataset. The results indicate that the RF- and SVM-based BSR models achieved prediction accuracies of approximately 0.64–0.94 for different classification schemes applied to the test sub-dataset. Additionally, the precision, recall, and F1-score indicators showed satisfactory values with respect to the BSR levels with relatively large sample sizes. The RF-based models had a lower tendency for overfitting compared to the SVM-based models. The performance of the BSR models was influenced by the quantity of total datasets, the classification schemes, and imbalanced data. Overall, the RF- and SVM-based BSR models can improve the evaluation efficiency of earthquake-damaged buildings in mountainous areas.

Keywords: earthquake; building seismic resilience (BSR); machine learning model (MLM); different classification schemes

Citation: Wen, H.; Zhou, X.; Zhang, C.; Liao, M.; Xiao, J. Different-Classification-Scheme-Based Machine Learning Model of Building Seismic Resilience Assessment in a Mountainous Region. *Remote Sens.* **2023**, *15*, 2226. https://doi.org/10.3390/rs15092226

Academic Editor: Teng-To Yu

Received: 6 March 2023
Revised: 20 April 2023
Accepted: 21 April 2023
Published: 22 April 2023

1. Introduction

An earthquake is a catastrophic natural event that occurs suddenly [1]. Due to the frequency of earthquakes, greater emphasis is being placed on ensuring the seismic resilience of cities and buildings. Resilience, derived from the Latin word 'resilio' [2], is a term that emerged in the field of ecology in the 1970s to describe the ability of systems to maintain or restore their functions after a disturbance [3]. Engineering [4,5], social science [6,7], and other fields have all adopted the notion of resilience. The concept of resilience can be applied in different dimensions. In the urban and architectural fields, building resilience constitutes the ability to maintain normal function, resist damage, or recover from highly damaging effects precipitated by the natural environment (including natural disasters) and the passage of time [8]; resilience is also used to assess and quantify a building's ability to retain or recover its operations in the aftermath of catastrophic natural catastrophes such as earthquakes [9].

BSR refers to the ability of buildings to maintain and quickly restore their functions during/following an earthquake. There are many studies on the seismic resilience of cities [10–13]. You et al. [14] proposed a methodology for assessing community resilience via the seismic performance of individual buildings. Based on the rough set method, Gizzi et al. [15,16] analyzed the reliability of using earthquake-induced building damage information to reduce the risks faced by earthquake-prone towns. At present, there are few studies on BSR, and most of them are focused on aspects of structural components [17–19]. Various scholars have studied the seismic resilience of buildings with different systems. Dong et al. [20] proposed a system for evaluating the seismic resilience of steel structures while taking economic, social, and environmental factors into account. Using the concept of fuzzy set theory, several studies [21] have described a new measurement standard for the seismic resilience evaluation of bridges. For critical infrastructures such as hospitals, Shang [22] provided a methodology for measuring hospital system seismic resilience, while Hassan et al. [23] proposed a hospital seismic function and recovery process assessment framework.

At present, there are four evaluation criteria for BSR [24]. FEMA P-58 [25] is a performance-based seismic design method that was proposed by the Federal Emergency Management Agency in 2012. Its performance evaluation results consist of the probability distribution of performance indicators such as casualties, repair costs, and repair time. Almufti and Willford [26] proposed an evaluation system for determining building resilience that improves the method for calculating repair time on the basis of FEMAP-58. The United States Resiliency Council (USRC) [27] established a building performance evaluation system in 2015 [28] based on assessment results provided by FEMA P-58 and REDI. The standard for BSR assessment [29] has been implemented in China since 1 February 2021; since then, it has helped relevant industries actively improve BSR. Clearly, BSR is associated with physical conditions, such as a building's structure, damage factors, and the environment, and social factors, such as casualties, repair costs, and time. The former set is known as physical resilience, and the latter has been termed social resilience. Zhang et al. [30] also recommended dividing catastrophe resilience evaluation into physical and social resilience, with physical resilience being more significant prior to disasters. Physical resilience refers to the ability to resist disasters, while social resilience refers to the ability to reconstruct after disasters.

Most of the previous studies on physical resilience mainly considered the influence of seismicity on buildings' structural properties. However, buildings are not independent of their environments. With regard to physical BSR, it is vital to couple the effect on the building itself with other factors [8,31]. Several researchers [32] described a hybrid information fusion strategy for statistically evaluating Nepal's earthquake resilience by developing nine indicators at the geological, architectural, and societal levels. Some investigators [33] also quantified community resilience by selecting various indicators while accounting for different dimensions. Referring to the previous studies, this paper focuses on building seismic physical resilience (BSR) in mountainous areas. It constitutes a new approach to exploring the combined effect of seismic factors, geological factors, topography factors, environmental factors, and building factors on BSR.

One of the keys to achieving rapid functional recovery after earthquakes is the timely assessment of the site [34]. Traditional on-site investigations usually require trained professionals, resulting in prolonged assessment times. However, an alternative exists, which shows promise with respect to addressing the shortcomings of the traditional approach. For this purpose, Shuanghe Town was chosen as the research location, and a database containing 19 influencing factors was established, which included five dimensions: seismic factors, geological factors, topography factors, environmental factors, and building factors. Based on MLMs (RF and SVM) whose hyperparameters were tuned, we developed evaluation models to predict the BSR level in mountainous areas. Through this evaluation model, the state of BSR can be quickly judged, the cost of data collection can be reduced, and potential risks can be identified at an early stage, thereby providing an academic and practical resource for earthquake damage mitigation/prevention and building site selection planning.

2. Study Area and Earthquake-Damaged Building Inventory

2.1. Study Area

The study area is located in the center of Shuanghe Town (Figure 1). Shuanghe Town is located near the southern border of Changning County, Sichuan Province, China, with an area of 135.37 square kilometers and a total population of 18,955 people (2017). The area's highest elevation is 953.4 m, while its lowest elevation is 367 m. Therefore, it is a mountainous and hilly terrain. Its lithology corresponds to thick mudstone and thin sandstone. Cambrian, Ordovician, and Lower Silurian strata are exposed on the surfaces of this region. Changning County is located on the Yangtze platform. There are many secondary faults in the region. The epicenter is located in the Changning-Shuanghe large anticline distribution area [35], for which the geological structure is complex.

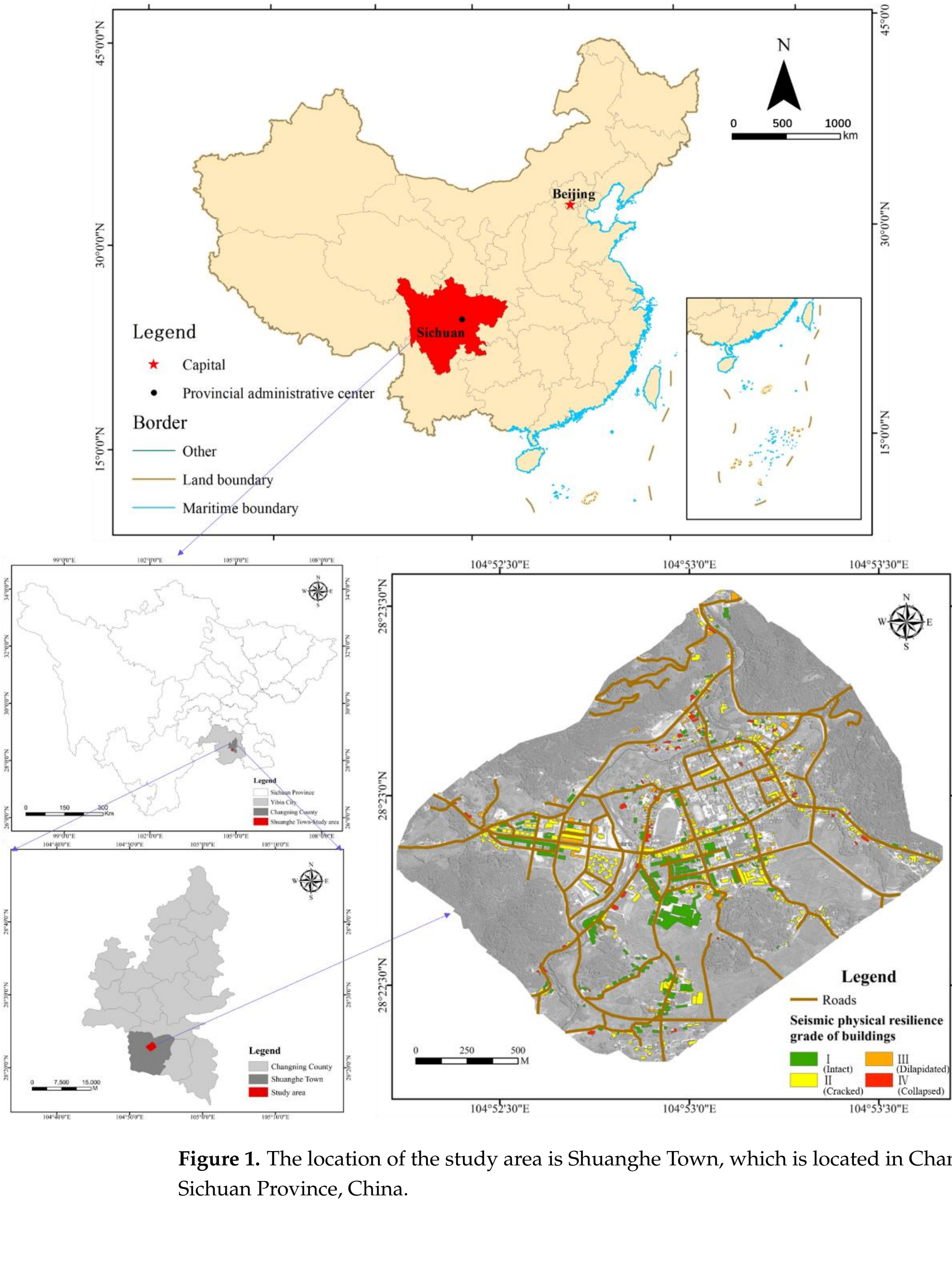


Figure 1. The location of the study area is Shuanghe Town, which is located in Changning County, Sichuan Province, China.

At 22:55 on 17 June 2019, a M 5.8 earthquake struck Changning County, Sichuan Province, with a focal depth of 6.0 km and an epicenter (28.406°N, 104.933°E) near the northeast of Shuanghe Town. The earthquake struck 8 surrounding counties, including Changning County. According to preliminary statistics, 13 people died, 220 people were injured, and the direct economic losses amounted to about USD 1.422 billion. The earthquake had a devastating effect, resulting in enormous losses of life and property, and 66% of these losses were due to a lack of seismic measures [36]. In this study, Shuanghe Town was selected as a case study area to assess BSR in a mountainous area.

2.2. Earthquake-Damaged Building Inventory and BSR Classification Schemes

Our research group conducted a post-disaster field survey at Shuanghe Town from 9–14 July 2019. Each building in Shuanghe Town was located using satellite images, and the buildings were numbered using ArcGIS. The team was divided into two groups that were assigned to visit each household in Shuanghe Town. We obtained an earthquake-damaged building inventory. A total of 855 groups of field building data were collected in three days, including with respect to building structure, number of floors, use categories, and the degree of earthquake-induced damage. At the same time, UAV aerial photography was carried out on site, and an orthophoto of Shuanghe town post-disaster was obtained.

Survey data showed that the proportions of building structure types in the study area were as follows: 82.79% brick–concrete structures, 13.00% brick–timber structures, 3.04% steel and reinforced concrete structures, and 1.17% hybrid structures. The building structures in the study area are relatively simple and easily observable, while the actual damage parameters of various building materials, foundation could not be directly obtained in the field. Using a visual technique, the buildings' damage levels were assessed. With reference to the relevant building damage grade standards in China, we focused our investigation on uneven settlement cracks and inclinations of the structures.

Using the earthquake-damaged building inventory, based on the damage degree, importance, quantity, and proportion of each dangerous component; the ability to maintain a building's function; and the reparability of a building, the BSR was divided into three classification evaluation schemes considering the initial four different earthquake damage degrees. Table 1 shows the detailed classification schemes. Table 2 shows the detailed datasets corresponding to the three classification schemes. Here, 855 buildings in the study area, whose complete information with respect to influencing factors have been obtained, were employed as 855 groups of samples. In accordance with the C4 scheme, the numbers of BSR level I, II, III, and IV are 256, 397, 125, and 77, and the corresponding percentages of the total numbers are 29.94%, 46.43%, 14.62%, and 9.01%. According to a ratio of 7:3, the whole dataset was randomly split into groups of 603 and 252 comprising training and testing sub-datasets, respectively.

Table 1. Classification evaluation schemes and description of BSR.

C2	**C3**	**C4**	**Initial Earthquake-Damaged Degrees**	**Description**	**Field Photos**	**UAV Images**
1~2	**A~C**	**I~IV**				
1: Repairable, including levels I, II, and III	A: Intact	I	Intact	The building's structure is basically intact and meets the requirements for safe use.		
	B: Damaged, including levels II and III	II	Cracked	Individual structural members are cracked or slightly inclined, the main structure is safe, or the damage can be repaired, basically meeting the requirements for safe use.		
		III	Dilapidated	Many structural members of the building are cracked or seriously inclined. Some load-bearing buildings are unsafe because they cannot be restored once damage prevents them from meeting the standards for safe usage.		
2: Non repairable	C: Collapsed	IV	Collapsed	Load-bearing structure does not meet the safety requirements as the building has collapsed.		

Table 2. Detailed datasets correspond to three classification schemes of BSR.

Classification Schemes	C2		C3			C4			
	C21	C22	C3A	C3B	C3C	C4I	C4II	C4III	C4IV
Number of Buildings	778	77	256	522	77	256	397	125	77

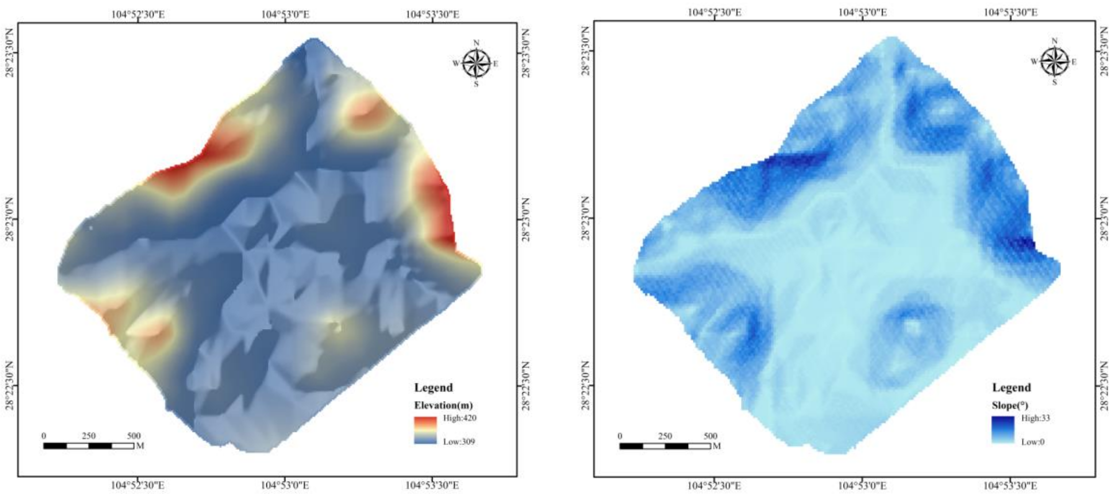
3. Data and Methods

3.1. Data

Following an earthquake, BSR is affected by the building itself and other factors, such as seismic factors, geological conditions, topography, environmental factors, and building factors. Wen et al. (2023) considered and selected five types of factors for building seismic resilience assessment [31]. By combining previous references [8,37–41], 19 factors were selected to generate the factor database for the BSR mapping of buildings with earthquake-induced damage. They are as follows: elevation, slope, aspect, curvature, plan curvature, profile curvature, micro-landform, CRDS (relative to slope structure [42]), stratum, fault wall, distance from faults, peak ground velocity (PGV), peak ground acceleration (PGA), seismic intensity, distance from rivers, distance from roads, building structure, number of floors, and building use category.

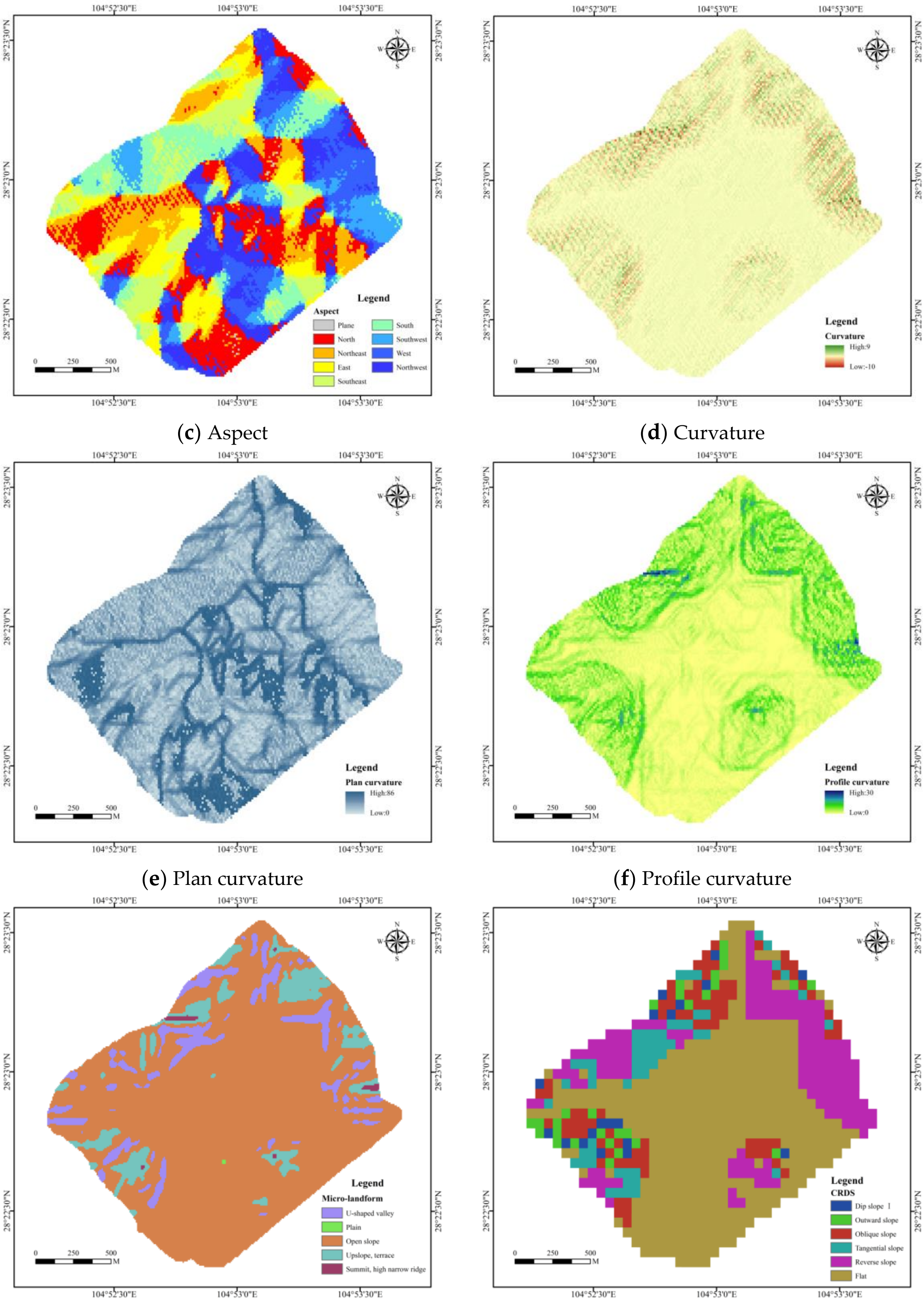
The PGV, PGA, and seismic intensity data were obtained from USGS (https://earthquake.usgs.gov/earthquakes/eventpage/us600041ry/shakemap, accessed on 18 June 2019). The stratum and fault wall data were obtained from the National Geological Archives of China (http://www.ngac.org.cn/Map/List, accessed on 19 June 2019). The road and river network data were derived from Google Earth remote sensing images. DEM can be used to process and retrieve data concerning slope, aspect, curvature, plan curvature, profile curvature, micro-landform, and CRDS via ArcGIS. In this study, we used ASTER GDEM data, corresponding to a resolution of 1 arc-second (approximately 30 m). The data can be downloaded for free from JAXA's official website and can be integrated with ArcGIS software and tools. The data on building structure, number of floors, and building category were obtained through field investigation.

Changning County, Sichuan Province, is situated at the nexus of the Huaying Mountain and Emei Mountain fault zones. This can affect the earthquake intensity with respect to the distance from the fault (Figure 2k). Meanwhile, the ground motion amplitude of the hanging wall is greater than the foot wall, which is described in Table 3. The study area, Shuanghe Town, spans 4 reverse faults (Figure 2j). By using ArcGIS to calculate Euclidean distance, the building distribution of hanging walls (above the fault plane) and footwalls (below the fault plane) can be identified.



(**a**) Elevation (**b**) Slope

Figure 2. *Cont.*

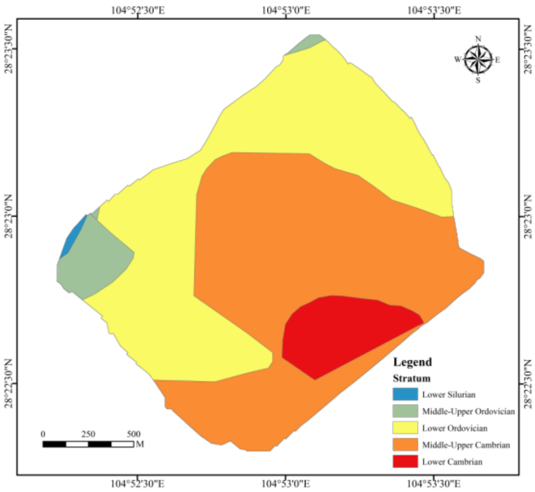


(**c**) Aspect (**d**) Curvature

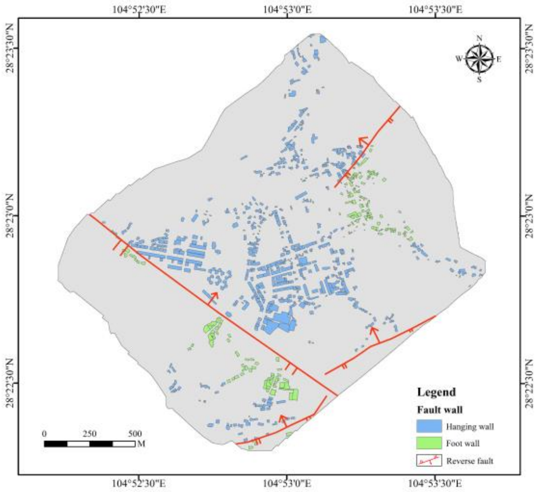
(**e**) Plan curvature (**f**) Profile curvature

(**g**) Micro-landform (**h**) CRDS

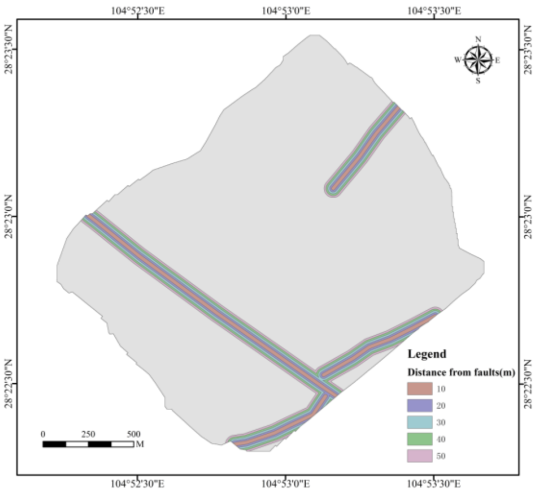
Figure 2. *Cont.*



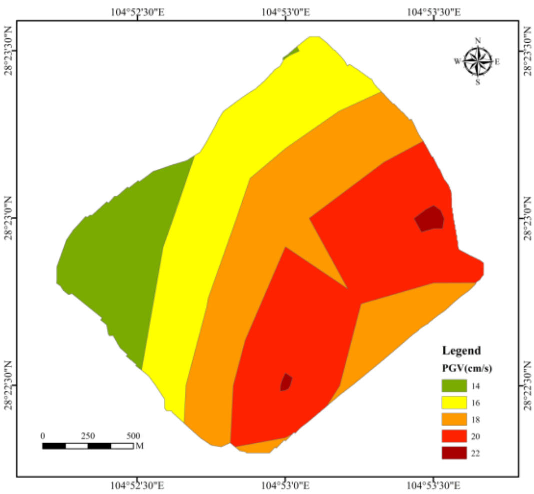
(**i**) Stratum



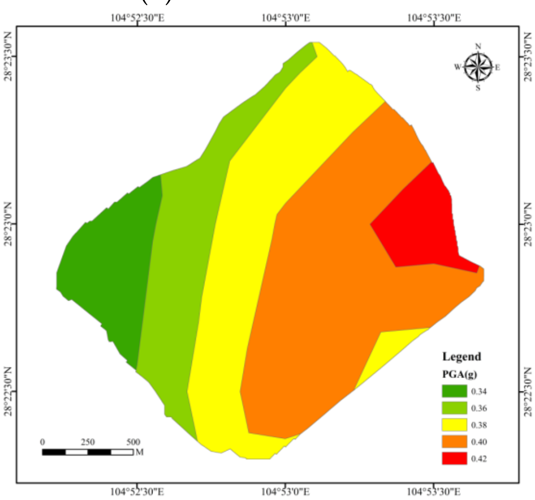
(**j**) Fault wall



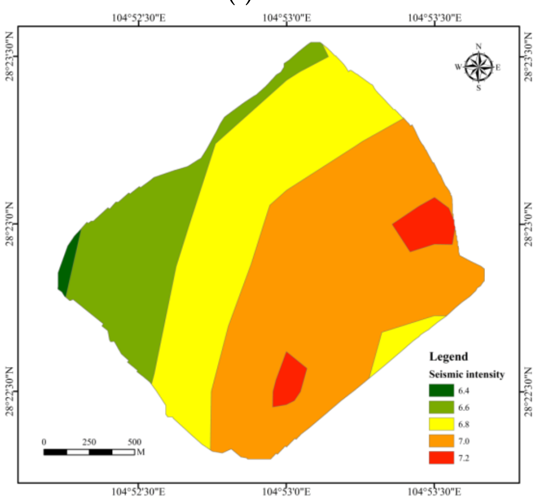
(**k**) Distance from faults



(**l**) PGV

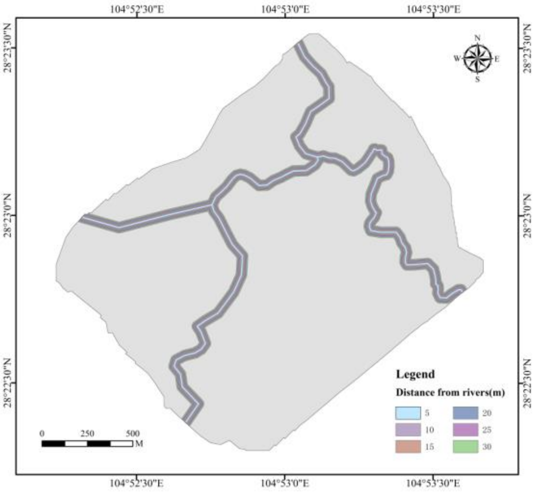


(**m**) PGA

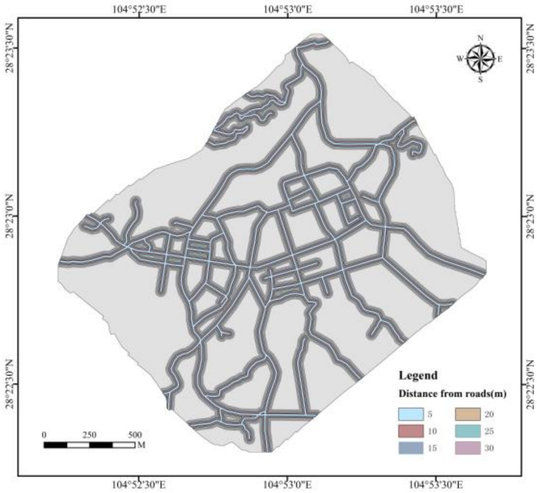


(**n**) Seismic intensity

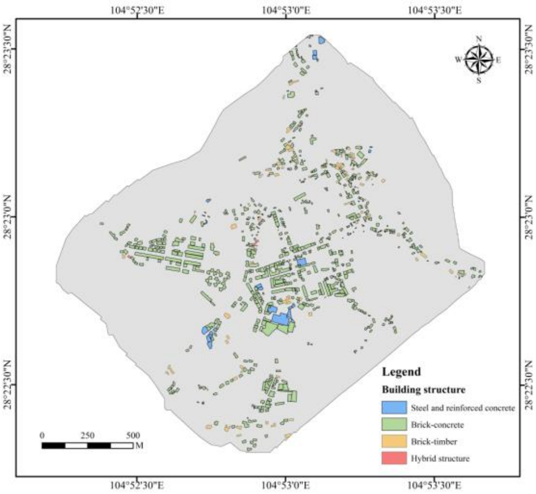
Figure 2. *Cont.*



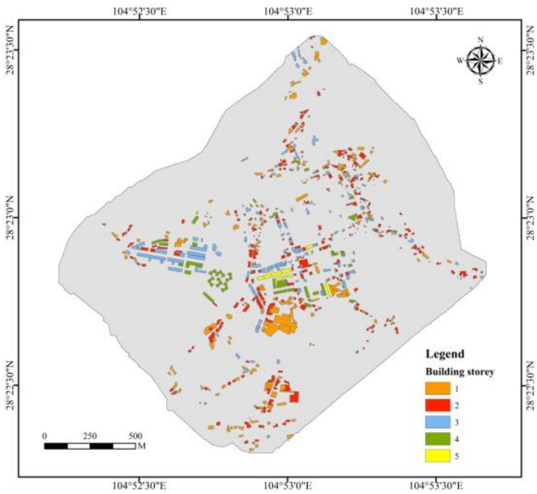
(**o**) Distance from rivers



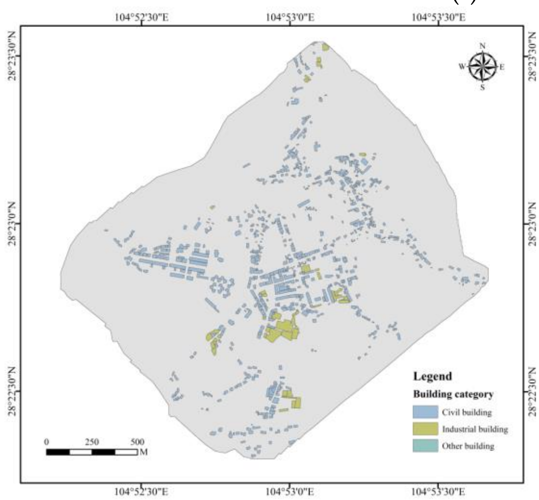
(**p**) Distance from roads



(**q**) Building structure



(**r**) Number of floors



(**s**) Building category

Figure 2. Influencing factors' thematic layers: (1) topographic factors: (**a**–**h**); (2) geological factors: (**i**–**k**); (3) seismic factors: (**l**–**n**); (4) environmental factors: (**o**,**p**); (5) building factors: (**q**–**s**).

Table 3. Categories and detailed descriptions of influencing factors.

Category	Influence Factors	Description
Topographic factors	Elevation	Elevation directly reflects the level of terrain and determines the extraction of other slope factors. Different elevations affect different land, vegetation, and climatic factors and human activities.
	Slope	Slope is the proportion in distance of horizontal to vertical height. The stress distribution in different slopes is different.
	Aspect	The slope's aspect is the direction of projection. In a mountainous environment, aspect has a significant impact on precipitation and has an impact on hydrogeology.
	Curvature	Curvature is the second-order derivative of the surface, affecting the erosion of soil via water flow.
	Plan curvature	Plan curvature refers to the change rate of surface aspect. The essence is to extract the aspect of DEM and then extract the slope of this aspect.
	Profile curvature	Profile curvature refers to the change rate of the surface slope. In essence, the idea is to extract the aspect of the DEM twice.
	Micro-landform	Micro-landform is a small undulation with complex surface on a large landform, which is mainly formed by erosion and accumulation under weathering. The strength of rock and soil in different micro-landform units are different.
Geological factors	CRDS	CRDS refers to the reclassification of both stratum dip direction and slope aspect. Different combination types have different slope stability values.
	Stratum	The formation time and weathering degree are different; the bearing capacity and lithology of stratum is different.
	Fault wall	The ground motion amplitude of the hanging wall is greater than the foot wall.
	Distance from faults	Within a specific range, the closer the fault, the looser the soil, and the more sensitive a building is to earthquake damage.
Seismic factors	PGV	The maximum absolute value of surface particle velocity during earthquake motion.
	PGA	The maximum absolute value of surface particle acceleration during earthquake motion.
	Seismic intensity	The influence of earthquake on the surface and the structural properties of buildings.
Environmental factors	Distance from rivers	Within a certain range, the closer a region is to a river, the higher the water content of soil layer; this relationship affects an area's hydrogeological conditions and foundation bearing capacity.
	Distance from roads	Within a certain range, the road construction leads to the stress redistribution of the original rocky soil. The closer an area is to a road, the greater the impact.
Building factors	Building structure	The earthquake damage degrees of different structures are different, as are their BSR levels different.
	Number of floors	Building height is different, weight is different.
	Building category	The design principles and materials of buildings with different use categories, such as industrial and civil use, are different.

Table 3 provides a detailed description of all 19 influencing factors noted above. Figure 2 shows the thematic layers of the 19 influencing factors noted above (created using ArcGIS).

3.2. Methods

Machine learning algorithms, such as RF and SVM, are increasingly being used in disaster risk assessment [43–46]. Machine learning algorithms are also commonly used to assess building vulnerability and resilience [8,47–49]. In this study, RF and SVM were combined to verify or compare the BSR evaluation results.

Figure 3 depicts a flowchart of the methodologies employed in this work. This study comprises three steps: (1) Via field investigation, combined satellite remote sensing images, a digital elevation model, and related websites, we obtained an earthquake-damaged building inventory and information relevant to BSR. A database of BSR influencing factors was established. (2) The influence factor was used as the input layer of the model, and the BSR level was used as the output layer. Then, we used RF and SVM to train the BSR evaluation models using a training sub-dataset. (3) The test sub-dataset was used to validate and evaluate the models' performance via validation indicators.

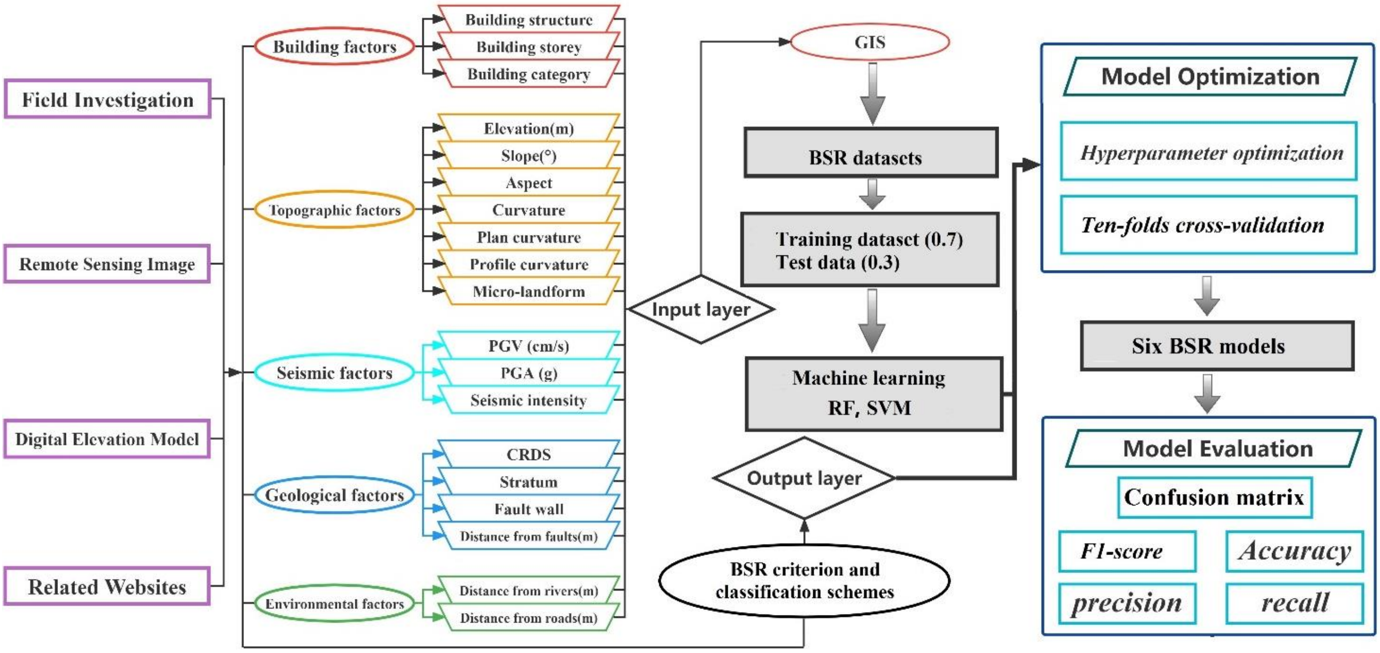


Figure 3. Flowchart of the methods used.

3.2.1. Random Forest

RF uses multiple decision trees to randomly comb through different data subsets and provide classification results with the maximum number of votes as the final output [50]. RF models show strong, robust, and accurate performance with respect to dealing with complex data [51]. Wang et al. [41] used RF to propose quantitative risk assessment model of based on landslide susceptibility mapping. The model presented good prediction capacity. Zhang et al. [8] used RF to develop a building physical resilience evaluation model for mountainous areas.

3.2.2. Support Vector Machine

Due to its excellent learning capacity with respect to tackling classification challenges and minimal computing complexity, an SVM is often used as an effective method for solving classification problems in the case of small samples. To date, many in-depth disaster risk studies have been conducted using an SVM [30,33,52,53]. Zhang et al. [8] used SVM to perform building physical resilience evaluation in a mountainous area.

On the basis of the above analysis, RF and SVM were selected to develop the BSR assessment models in this study.

3.2.3. Model Evaluation Method

In analyses of prediction impact, confusion matrices are frequently used to ascertain an algorithm's performance. Each column represents the resilience level predicted by the model, and the data contained in each column indicate the number of levels. Each row represents the actual resilience level, and the data of each row represent the actual number of this level.

One of the critical steps of MLM development is to evaluate a model's prediction ability. To simplify the formulation, the confusion matrix's data corners are represented by the combination of the actual and predicted resilience level (Table 4). In this study, Nij indicates that the actual resilience level corresponds to I, while prediction corresponds to j. Then, some indicators calculated based on the confusion matrix, including accuracy, precision, and recall, can further measure the models' performance.

Table 4. Four-classification confusion matrix.

		Prediction Level (j)			
		I	**II**	**III**	**IV**
Actual level (i)	**I**	N11	N12	N13	N14
	II	N21	N22	N23	N24
	III	N31	N32	N33	N34
	IV	N41	N42	N43	N44

The accuracy (*Acc*) rate is the percentage of samples achieving correct predictions in the overall sample. It is the most fundamental, logical, and straightforward technique for determining how well a categorization model has been evaluated. However, due to the unbalanced nature of these data, the use of only accuracy rate is insufficient for evaluating model performance.

Precision (*Pre*) and recall (*Rec*) are two evaluative indicators that are widely used in statistical classification [8,54]. In this study, the number of samples is not balanced, and it is mainly the prediction effect of each BSR level of the model that is measured. *Rec* is a measure of the ability of the prediction model to select a specific level from the data sets. (1-*Rec*) denotes the missing judgment rate (MJR) of the model prediction. Pre concerns the effect of prediction, which refers to the fraction of the true values of the level projected as a given level in all the samples, revealing the model's precision and fulfilling the actual demands. (1-*Pre*) denotes the error judgment rate (EJR) of the model prediction. Both precision and recall are biased evaluative indicators, while F1-score (*F1*) is a comprehensive evaluative indicator of the previous two indicators. For multi-classification problems, the *Rec*, *Pre*, and *F1* of different BSR levels should be calculated separately.

Based on Table 4, the four evaluative indicators, *Acc*, *Pre*, *Rec*, and *F1*, are expressed as Equations (1)–(4), respectively.

$$Acc = \sum_{i=1}^{4} N_{ii} / \sum_{i=1}^{4} \sum_{j=1}^{4} N_{ij} \tag{1}$$

$$Pre_j = N_{ii} / \sum_{i=1}^{4} N_{ij} \tag{2}$$

$$Rec_i = N_{ii} / \sum_{j=1}^{4} N_{ij} \tag{3}$$

$$F1_i = 2 * Pre_i * Rec_i / (Pre_i + Rec_i) \tag{4}$$

4. Results

4.1. Optimization of Models via Hyperparameter Tuning

Hyperparameter tuning is an important factor influencing an MLM's performance [8,55]. Based on the R language, the RF model has two key hyper-parameters, namely, Mtry and ntree, which indicate the number of variables used in the binary tree at the node and the number of decision trees, respectively. By calling the RF program package, the optimal parameter mtry = 3 was selected by using for-loop iteration, and the optimal mtry was substituted into the code to observe the stability of the model error and determine the optimal ntree value (in this case, ntree = 1500).

In the SVM, the kernlab package is implemented using the R language. In the classification model, type selection C-svc and the kernel function kernel selection rbfdot are the best performing. At this stage, two key hyper-parameters are included, i.e., sigma and C, which indicate the width of the kernel function and the cost of violating the constraint, that is, the tolerance for allowing classification errors, respectively. By employing for-loop iteration using a ten-fold cross-validation method, the optimal parameter combinations were determined, namely, kpar = list (sigma = 0.300) and C = 15.

Table 5 lists the optimized key hyper-parameters of the RF and SVM.

Table 5. The results of hyper-parameters' tuning.

RF		SVM	
mtry	**ntree**	**Sigma**	**C**
3	1500	0.300	15

Based on the above optimal parameters, the ten-fold cross-validation method was used to find the optimal sample. To limit the error produced by the individual sampling techniques with respect to the model outputs and increase the accuracy of the test algorithm, the training sub-dataset was randomly divided into ten portions, nine of which were utilized as training data and one as a test dataset. Due to the unbalanced nature of the data, the weights of the four BSR levels were adjusted using the 'classwt' code after finding the optimal sample. After debugging, it was found that if the proportion of level III and IV data is increased, some level I and II data are misclassified to level III and IV, and the model's results are poor. On the contrary, if the proportion of level I and level II is increased, the classification ability of these two types will be significantly strengthened, while the classification ability of the original minority dataset will be slightly poor. Overall, the model prediction effect is improved after weight adjustment.

4.2. Results of the Different BSR Classification Schemes

Table 6 provides a detailed list of the validation indicators of the six BSR models. In summary, based on the Pre, Rec, and F1 scores, the table shows that C21, C3B, and C4II are satisfactory results for either the RF or SVM-based models. The overall model performance of the C2 scheme is superior, followed by C3 and C4. Since the C4 scheme should be more comprehensible in practice, the results of the C4 scheme are discussed in further detail below.

Table 6. Details of the validation indicators of the six BSR models.

Classification Schemes		C2		C3			C4			
		C21	**C22**	**C3A**	**C3B**	**C3C**	**C4I**	**C4II**	**C4III**	**C4IV**
RF	**Pre**	0.94	1.00	0.88	0.75	1.00	0.64	0.65	1.00	0.8
	Rec	1.00	0.38	0.58	0.98	0.13	0.62	0.85	0.33	0.50
	F1	0.97	0.55	0.7	0.85	0.22	0.63	0.74	0.50	0.62
SVM	**Pre**	0.94	0.5	0.80	0.78	1.00	0.71	0.62	0.50	0.67
	Rec	0.99	0.38	0.62	0.94	0.38	0.65	0.80	0.25	0.29
	F1	0.96	0.5	0.80	0.85	0.55	0.68	0.70	0.33	0.40

Figure 4 lists the two confusion matrices of the training sub-datasets and test sub-datasets based on the RF-based model according to the C4 scheme's results. Figure 5 lists the two confusion matrixes of the training sub-dataset and test sub-dataset for the SVM-based model according to the C4 scheme's results. The yellow highlights are the models' Acc values as determined using Equation (1). As we can see by comparison, the RF-based model's performance is relatively better than that of the SVM-based model.

Prediction

Rank	I	II	III	IV	Recall
I	226	4	0	0	98%
II	19	338	0	0	95%
III	4	13	96	0	85%
IV	3	4	0	62	90%
Precision	90%	94%	100%	100%	94%

Actual

(**a**)

Prediction

Rank	I	II	III	IV	Recall
I	16	10	0	0	62%
II	6	34	0	0	85%
III	3	4	4	1	33%
IV	0	4	0	4	50%
Precision	64%	65%	100%	80%	67.44%

Actual

(**b**)

Figure 4. Confusion matrix of prediction results achieved by RF-based model. (**a**) Training sub-dataset; (**b**) test sub-dataset.

Prediction

Rank	I	II	III	IV	Recall
I	226	4	0	0	98%
II	4	353	0	0	99%
III	1	1	111	0	98%
IV	0	0	0	69	100%
Precision	98%	99%	100%	100%	99%

Actual

(**a**)

Prediction

Rank	I	II	III	IV	Recall
I	17	7	2	0	65%
II	6	32	1	1	80%
III	1	8	3	0	25%
IV	0	5	0	2	29%
Precision	71%	62%	50%	67%	63.53%

Actual

(**b**)

Figure 5. Confusion matrix of prediction results achieved by SVM-based model. (**a**) Training sub-dataset; (**b**) test sub-dataset.

As we can see in Figures 4 and 5, the RF-based model's prediction accuracy values are 0.94 and 0.67 for the training and test sub-datasets, respectively. The SVM-based model's prediction accuracies are 0.99 and 0.64 for the training and test sub-datasets, respectively. The SVM-based model has a more obvious degree of overfitting than the RF-based model.

4.3. Model Performance

4.3.1. Accuracy Analysis

Based on the integration of the data in Table 6 into Equation (1), we can obtain the Acc values (Figure 6) of the six BSR models.

According to the results of all the different classification schemes, the accuracy values of the RF-based and SVM-based models are 0.67 and 0.64 for the C4 scheme, respectively. The accuracy values of the RF-based and SVM-based models are 0.77 and 0.78 for the C3 scheme, respectively. Furthermore, the accuracy values of the RF-based and SVM-based models are 0.94 and 0.93 for the C2 scheme, respectively. (Figure 6). The results show that the BSR models' prediction accuracy noticeably increases with a decreasing classification number.

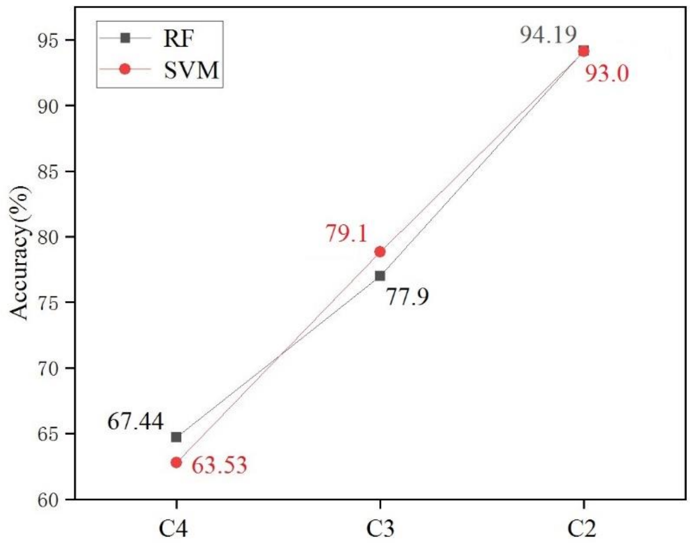


Figure 6. The six BSR models' accuracy with respect to the different classification schemes according to the test sub-dataset.

Based on Figure 6, we can further discuss the BSR models' accuracy. The accuracy of the BSR models in mountainous areas is mainly affected by the small overall sample size, the classification schemes employed, and the unbalanced nature of the data.

(1) The small overall sample size: Under the limitations of few investigators, a short research period, and occasional poor communication, the team collected 855 buildings of Shuanghe Town. The overall samples collected are relatively small, which specifically impacts the prediction performance of the MLM.
(2) Classification scheme: The classification algorithm was used initially to study the binary classification problem. This study concerns different BSR classification schemes of two to four levels. The C3 and C4 schemes are not a simple accumulation of the data from the binary classification, which leads to misclassification or indivisibility and affects the prediction accuracy for BSR to a certain extent.
(3) Unbalanced nature of the data: When data are highly unbalanced, the performance of an MLM will decrease due to the C4 scheme, the proportions of buildings with four BSR levels (I, II, III, and IV) in the total sample are unequal. In RF, the characteristics of a small number of categories will be regarded as noise, which is usually ignored. In an SVM, the classification hyperplane tends to correspond to level III and level IV minority classification datasets, so some support vectors are divided into level II and level I majority classification datasets. Due to the unbalanced nature of the sample dataset, the predicted results based on the two different MLMs (RF and SVM) all show that the results obtained by the classification algorithm should be biased towards a large number of classifications.

4.3.2. Precision Analysis

With regard to the C4 scheme for determining BSR, based on Figures 4 and 5, the accuracy and recall values of the models are shown in Figure 7.

(1) For the RF-based model, the precision values are 0.64, 0.65, 1.00, and 0.80 for BSR levels I, II, III, and IV, respectively (Figure 7a). For the SVM-based model, the precision values are 0.71, 0.62, 0.50, and 0.67 for levels I, II, III, and IV, respectively (Figure 7b). As a whole, the precision values of the RF-based model are higher than those of the SVM-based model. From among the 1-*Pre* values, the highest number of error judgments were made in reference to level I, followed by level II.
(2) For the RF-based model, the recall values are 0.62, 0.85, 0.33, and 0.50 for BSR level I, II, III, and IV, respectively (Figure 7c); for the SVM-based model, the precision values

are 0.65, 0.80, 0.25, and 0.29 for level I, II, III, and IV, respectively (Figure 7d). As a whole, the recall values of the RF-based model are higher than those of the SVM-based model. From among the 1-*Rec* values, the highest number of missing judgments were made in reference to level III, followed by level IV.

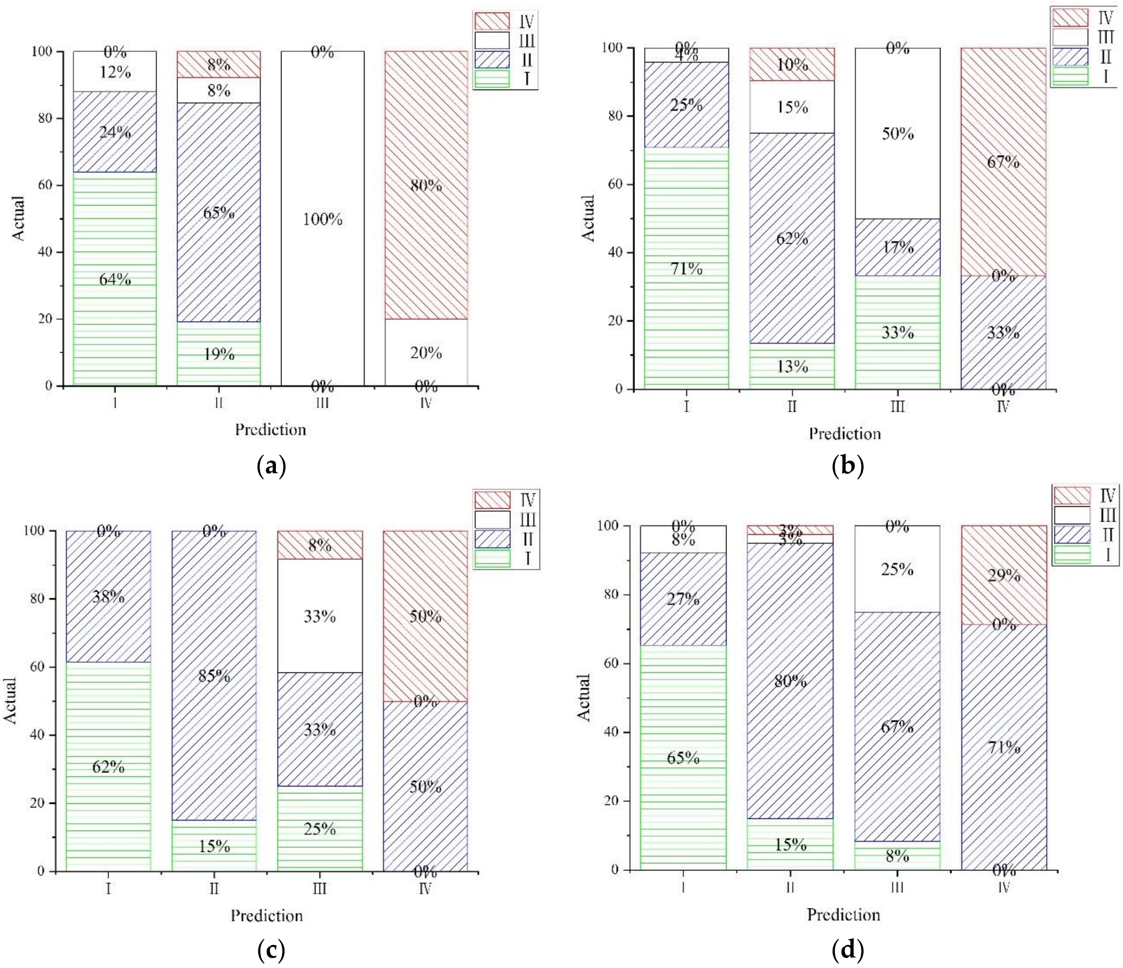


Figure 7. Prediction precision and recall of the C4 scheme-based models. (**a**) RF—Precision, (**b**) SVM—Precision, (**c**) RF—Recall, (**d**) SVM—Recall.

5. Discussion

5.1. Analysis of BSR Models' Performance Differences

For either the RF-based or SVM-based models, the precision of model prediction is the highest within the same column for any BSR level. As mentioned in Zhang et al.'s study (2022) [8] concerning three-classification models for mountainous building resilience assessment, in general, RF-based models exhibit slightly higher accuracy and lower overfitting, thus demonstrating their relatively better performance compared to SVM-based BSR models. These models can be considered to yield correct classification and accurate results for most buildings, indicating their good reference value for BSR assessment.

Based on the results presented in Table 6 and Figures 4, 5 and 7, we observed that the RF and SVM models exhibited different performance with respect to various datasets. This indicates that we should select different models for different datasets to achieve optimal predictive performance. To gain further insights into the differences in performance between these models, we used "1-*Rec*" to denote the missing judgment rate (MJR) of the model's prediction and "1-*Pre*" to denote the error judgment rate (EJR) of the model's prediction. Overall, the MJR and EJR of the RF- and SVM-based models exhibited a generally consistent trend, but there were noticeable differences. For example, at the C3C level, the RF-based model's MJR was lower than that of the SVM-based model, while at the C4IV level, the RF-based model's MJR was higher than that of the SVM-based model. Additionally, the RF-based models' EJR was higher than the SVM-based models' EJR for most of the BSR levels, but at the C4I level, the RF-based model's EJR was lower than that of the SVM-based model.

Based on an analysis of Figure 6 and the previous results, it can be observed that the C2 scheme has the highest accuracy (*Acc*) value, while the C4 scheme has a comparatively lower Acc value. This implies that as the number of classes increases, the overall accuracy of the models tends to decrease. Additionally, the recall values for the C22, C3C, and C4IV levels of the BSR dataset are relatively low. This can be attributed to the fact that the proportion of data in these levels in the sample is significantly lower compared to the other categories, resulting in an imbalance in the resilience levels across the three classification schemes. This finding is consistent with our previous research [8,31], highlighting that imbalanced data related to resilience levels can negatively impact model performance.

5.2. BSR Model Recommendation and Performance Improvement

Although the overall performance of the models using the C2 scheme is better than that of the C4 scheme with respect to identifying BSR levels, it is worth noting that the C4 scheme is more refined and aligns better with the practical requirements of BSR assessment. Therefore, we recommended prioritizing the improvement of a model's performance with respect to the four-classification scheme in BSR assessment.

To further enhance the performance of the unbalanced-data-based MLM multi-classification models, several strategies can be implemented. Firstly, obtaining more BSR data from other earthquake-damaged buildings can be used to train the MLM while mitigating the impact of unbalanced data. It is believed that with a sufficiently large volume of data, the performance of the unbalanced-data-based models could reach an acceptable level. Secondly, in this study, certain influential factors, such as the building foundation type, foundation burial depth, and building service life, which are crucial for determining BSR, were difficult to obtain via low data coverage; thus, they were omitted. However, the exclusion of these key factors may result in decreased model performance. Therefore, we suggest the further incorporation of relevant influential factors that are pivotal to BSR, while considering the feasibility of data acquisition. Lastly, from a methodological perspective, hybrid optimization techniques or ensemble-learning models, such as hybrid optimization based on the Synthetic Minority Over-sampling Technique (SMOTE) and hyperparameter tuning, can be employed to improve the performance of unbalanced-data-based MLM models for BSR assessment [31]. Additionally, ensemble learning can be considered to combine the strengths of various models and enhance the efficiency of earthquake damage assessment for buildings in mountainous areas. During the iterative optimization process, it was identified that the generalization ability for particles was not strong. Therefore, complementary hybrid optimization methods can be used to optimize models and further enhance their performance and accuracy. We believe that the application of these strategies will help us more substantially address the issue of earthquake- induced damage on buildings in mountainous areas and provide more precise and reliable support for the seismic evaluation of buildings.

6. Conclusions

In this study, which concerned earthquake-damaged buildings in Shuanghe Town, by considering the combined effects of seismic, geological, topographical, environmental, and building factors on the BSR in the region's mountainous area, we identified 19 factors influencing BSR and established a corresponding BSR database. Based on the training sub-dataset, we developed six assessment BSR models by combining two MLMs and three classification schemes. The important findings and recommendations of this study are as follows:

(1) For machine learning models, the performance of identifying whether or not buildings have suffered earthquake-induced damage is better than that for identifying different degrees of earthquake-induced damage.
(2) The C2 scheme has the highest Acc value, while the C4 scheme's performance is more dependent on practicality with regard to BSR assessment. That is to say, the BSR assessment models' performance was affected by the small number of total samples, the classification scheme used, and the unbalanced nature of the data.
(3) The RF-based model, with a slightly higher Acc and lower overfitting, offers better performance than that of the SVM-based BSR model, which can provide a reference for BSR assessment in mountainous areas.

The main focus of this study is the assessment of the physical BSR in mountain areas. If the relevant data regarding the influencing factors of social resilience are collected later, which can also be processed by the BSR model, the proposed models can be extended, and their performance can be further improved by flexibly adding more BSR training data and relevant BSR factors of different dimensions.

Author Contributions: Conceptualization, H.W.; Data curation, M.L.; Formal analysis, X.Z. and J.X.; Funding acquisition, H.W.; Investigation, M.L.; Methodology, H.W.; Project administration, H.W.; Software, C.Z.; Supervision, H.W.; Validation, H.W. and J.X.; Visualization, C.Z.; Writing—original draft, X.Z.; Writing—review and editing, H.W. All authors have read and agreed to the published version of the manuscript.

Funding: Natural Science Foundation of Chongqing, Award Number: CSTB2022NSCQ-MSX0594 | Recipient: Haijia Wen. Chongqing Science and Technology Commission, Award Number: cstc2018jscx-msybX0310 | Recipient: Haijia Wen.

Data Availability Statement: Data are contained within the article. The corresponding author can provide the necessary model upon request.

Acknowledgments: Thanks to all those who helped with this work, including the field earthquake-damaged building survey conducted by Yu Lin, Yang Wu, and Hui Zhang and some preliminary work on UAV image analysis conducted by Yue Wang and Jiahui Xu.

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

References

1. Kahandawa, K.A.R.V.; Domingo, N.D.; Park, K.S.; Uma, S.R. Earthquake damage estimation systems: Literature review. *Procedia Eng.* **2018**, *212*, 622–628. [CrossRef]
2. Cimellaro, G.P.; Reinhorn, A.M.; Bruneau, M. Framework for analytical quantification of disaster resilience. *Eng. Struct.* **2010**, *32*, 3639–3649. [CrossRef]
3. Holling, C.S. Resilience and stability of ecological systems. *Annu. Rev. Ecol. Syst.* **1973**, *4*, 1–23. [CrossRef]
4. Wildavsky, A.B. Searching for Safety. *J. Risk Insur.* **1988**, *57*, 564. [CrossRef]
5. Zhou, Y.; Lu, X. State-of-the-art on rocking and self-centering structures. *J. Build. Struct.* **2011**, *32*, 1–10. [CrossRef]
6. Adger, W.N. Social and ecological resilience: Are they related? *Prog. Hum. Geogr.* **2000**, *24*, 347–364. [CrossRef]
7. Folke, C. Resilience: The emergence of a perspective for social–ecological systems analyses. *Glob. Environ. Chang.* **2006**, *16*, 253–267. [CrossRef]
8. Zhang, C.; Wen, H.-J.; Liao, M.-Y.; Lin, Y.; Wu, Y.; Zhang, H. Study on Machine Learning Models for Building Resilience Evaluation in Mountainous Area: A Case Study of Banan District, Chongqing, China. *Sensors* **2022**, *22*, 1163. [CrossRef]
9. Li, X.; Yu, H.; Liu, P. Resilience-based seismic performance concept, assessment methodology and engineering application. *Build. Struct.* **2018**, *48*, 1–7. [CrossRef]
10. Bozza, A.; Asprone, D.; Manfredi, G. Developing an integrated framework to quantify resilience of urban systems against disasters. *Nat. Hazards* **2015**, *78*, 1729–1748. [CrossRef]
11. Gian Paolo Cimellaro, A.M.A.; Renschler, C.; Andrei, M.; Reinhorn, F.A.; Arendt, A.L. PEOPLES: A Framework for Evaluating Resilience. *J. Struct. Eng.* **2016**, *142*, 4016063. [CrossRef]
12. Lu, X.; Liao, W.; Fang, D.; Lin, K.; Tian, Y.; Zhang, C.; Zheng, Z.; Zhao, P. Quantification of disaster resilience in civil engineering: A review. *J. Saf. Sci. Resil.* **2020**, *1*, 19–30. [CrossRef]
13. Marasco, S.; Cardoni, A.; Zamani Noori, A.; Kammouh, O.; Domaneschi, M.; Cimellaro, G.P. Integrated platform to assess seismic resilience at the community level. *Sustain. Cities Soc.* **2021**, *64*, 102506. [CrossRef]

14. You, T.; Wang, W.; Chen, Y. A framework to link community long-term resilience goals to seismic performance of individual buildings using network-based recovery modeling method. *Soil Dyn. Earthq. Eng.* **2021**, *147*, 106788. [CrossRef]
15. Gizzi, F.; Masini, N.; Potenza, M.R.; Zotta, C.; Tilio, L.; Danese, M.; Murgante, B. Assessing Macroseismic Data Reliability through Rough Set Theory: Application on Vulture Area (Basilicata, Southern Italy). In *Advances in Intelligent Decision Technologies, Smart Innovation, Systems and Technologies*; Phillips-Wren, G., Jain, L.C., Nakamatsu, K., Howlett, R.J., Eds.; Springer: Berlin, Germany, 2010; Volume 4, VII, pp. 279–288. [CrossRef]
16. Gizzi, F.; Masini, N.; Potenza, M.R.; Zotta, C.; Tilio, L.; Danese, M.; Murgante, B. Assessing Macroseismic Data Reliability through Rough Set Theory: The Case of Rapolla (Basilicata, Southern Italy). In *Lecture Notes in Computer Science*; Springer: Berlin, Germany, 2010; Volume 6016, pp. 320–330, ISSN 0302-9743. [CrossRef]
17. Dong, Y.; Wang, X.; Yang, S.; Zhou, X.; Long, S.; Lv, L. Total probability seismic performance evaluation based on next generation performance-based design. *World Earthq. Eng.* **2020**, *36*, 12–20.
18. Lu, X.; Wu, D.; Zhou, Y. State-of-the-art of earthquake resilient structures [State-of-the-art of earthquake resilient structures]. *J. Build. Struct.* **2019**, *40*, 1–15. [CrossRef]
19. Porter, K.A.; Kiremidjian, A.S.; Legrue, J.S. Assembly-Based Vulnerability of Buildings and Its Use in Performance Evaluation. *Earthq. Spectra* **2001**, *17*, 291–312. [CrossRef]
20. Dong, Y.; Frangopol, D.M. Performance-based seismic assessment of conventional and base-isolated steel buildings including environmental impact and resilience. *Earthq. Eng. Struct. Dyn.* **2006**, *45*, 739–756. [CrossRef]
21. Andrić, J.M.; Lu, D. Fuzzy methods for prediction of seismic resilience of bridges. *Int. J. Disaster Risk Reduct.* **2017**, *22*, 458–468. [CrossRef]
22. Shang, Q.; Wang, T.; Li, J. A Quantitative Framework to Evaluate the Seismic Resilience of Hospital Systems. *J. Earthq. Eng. JEE* **2020**, *26*, 3364–3388. [CrossRef]
23. Hassan, E.M.; Mahmoud, H. Full functionality and recovery assessment framework for a hospital subjected to a scenario earthquake event. *Eng. Struct.* **2019**, *188*, 165–177. [CrossRef]
24. Ren, J.; Pan, P.; Wang, T.; Zhou, Y.; Wang, H.; Shan, M. Interpretation of GB/T 38591—2020 'Standard for seismic resilience assessment of buildings'. *J. Build. Struct.* **2021**, *42*, 48–56. [CrossRef]
25. Agency, F.; Hamburger, R.; Bachman, R.; Heintz, J.; Whittaker, A. *Seismic Performance Assessment of Buildings*; FEMA: Washington, DC, USA, 2012; Volume Methodology, p. 1.
26. Almufti, I.; Willford, M. REDi Rating System: Resilience-based Earthquake Design Initiative for the Next Generation of Buildings. *Arup Co* **2013**. [CrossRef]
27. Reis, E.; Mayes, R.L. The U.S. Resiliency Council (USRC) and the Building Rating System. *Improv. Seism. Perform. Exist. Build. Other Struct.* **2015**, *54*, 764. [CrossRef]
28. Xiao, Y.; Zhou, Y.; Wu, H.; Pan, P.; Wang, T. Comparative study on GB/T 38591—2020 'Standard for seismic resilience assessment of buildings' and relevant international standards. *J. Build. Struct.* **2021**, *42*, 194–202. [CrossRef]
29. *GB/T 38591-2020*; Standard for Seismic Resilience Assessment of Buildings. State Administration for Market Regulation: Beijing, China, 2020.
30. Zhang, X.; Song, J.; Peng, J.; Wu, J. Landslides-oriented urban disaster resilience assessment—A case study in ShenZhen, China. *Sci. Total Environ.* **2019**, *661*, 95–106. [CrossRef]
31. Wen, H.J.; Wu, J.N.; Zhang, C.; Zhou, X.Z.; Liao, M.Y.; Xu, J.H. Hybrid optimized RF model of seismic resilience of buildings in mountainous region based on hyperparameter tuning and SMOTE. *J. Build. Eng.* **2023**, *62*, 106488. [CrossRef]
32. Chen, W.; Zhang, L. Resilience assessment of regional areas against earthquakes using multi-source information fusion. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107833. [CrossRef]
33. Al Rifat, S.A.; Liu, W. Measuring community disaster resilience in the conterminous coastal United States. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 469. [CrossRef]
34. Lu, X.; Xu, Y.; Tian, Y.; Cetiner, B.; Taciroglu, E. A deep learning approach to rapid regional post-event seismic damage assessment using time-frequency distributions of ground motions. *Earthq. Eng. Struct. Dyn.* **2021**, *50*, 1612–1627. [CrossRef]
35. Li, D.; Zhan, Y.; Ding, Z.; Gao, J.; Wu, P.; Meng, L.; Sun, X.; Zhang, X. Upper crustal velocity and seismogenic environment of the Changning MS6. 0 earthquake region in Sichuan, China. *Chin. J. Geophys.* **2021**, *64*, 18–35. [CrossRef]
36. Wang, W.; Chen, H.; Jiang, X.; Ma, L.; Xu, Y.; Qu, M. Lessons from the casualties caused by the Changning Ms 6.0 earthquake in China. *Nat. Hazards Res.* **2021**, *1*, 81–87. [CrossRef]
37. Zhang, J.; Ma, X.; Zhang, J.; Sun, D.; Zhou, X.; Mi, C.; Wen, H. Insights into geospatial heterogeneity of landslide susceptibility based on the SHAP-XGBoost model. *J. Environ. Manag.* **2023**, *332*, 117357. [CrossRef] [PubMed]
38. Narjabadifam, P.; Hoseinpour, R.; Noori, M.; Altabey, W. Practical seismic resilience evaluation and crisis management planning through GIS-based vulnerability assessment of buildings. *Earthq. Eng. Eng. Vib.* **2021**, *20*, 25–37. [CrossRef]
39. Wen, H.-J.; Xie, P.; Xiao, P.; Hu, D.-P. Rapid susceptibility mapping of earthquake-triggered slope geohazards in Lushan County by combining romote sensing and the AHP model developed for the Wenchuan earthquake. *Bull. Eng. Geol. Environ.* **2017**, *76*, 909–921. [CrossRef]
40. Wen, H.; Wu, S.; Wang, G.; Xie, P.; Li, Y.; Zhang, C. A Vulnerability Assessment Method of Buildings Affected by Earthquake Based on Data Mining No. CN201811259565.6). Available online: https://d.wanfangdata.com.cn/patent/ChJQYXRlbnROZXdTMjAyMTAxMDkSEENOMjAxODExMjU5NTY1LjYaCDRuYXJubm5j (accessed on 19 June 2019).

41. Wang, Y.; Wen, H.; Sun, D.; Li, Y. Quantitative Assessment of Landslide Risk Based on Susceptibility Mapping Using Random Forest and GeoDetector. *Remote Sens.* **2021**, *13*, 2625. [CrossRef]
42. 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* **2020**, *105*, 1255–1279. [CrossRef]
43. Deng, W.; Zhou, J. Approach for feature weighted support vector machine and its application in flood disaster evaluation. *Disaster Adv.* **2013**, *6*, 51–58.
44. Liu, D.; Fan, Z.; Fu, Q.; Li, M.; Faiz, M.-A.; Ali, S.; Li, T.; Zhang, L.; Khan, M.-I. Random forest regression evaluation model of regional flood disaster resilience based on the whale optimization algorithm. *J. Clean. Prod.* **2020**, *250*, 119468. [CrossRef]
45. Motta, M.; de Castro Neto, M.; Sarmento, P. A mixed approach for urban flood prediction using Machine Learning and GIS. *Int. J. Disaster Risk Reduct.* **2021**, *56*, 102154. [CrossRef]
46. Stumpf, A.; Kerle, N. Object-oriented mapping of landslides using Random Forests. *Remote Sens. Environ.* **2011**, *115*, 2564–2577. [CrossRef]
47. Liuzzi, M.; Aravena Pelizari, P.; Geiß, C.; Masi, A.; Tramutoli, V.; Taubenböck, H. A transferable remote sensing approach to classify building structural types for seismic risk analyses: The case of Val d'Agri area (Italy). *Bull. Earthq. Eng.* **2019**, *17*, 4825–4853. [CrossRef]
48. Mangalathu, S.; Sun, H.; Nweke, C.C.; Yi, Z.; Burton, H.V. Classifying earthquake damage to buildings using machine learning. *Earthq. Spectra* **2020**, *36*, 183–208. [CrossRef]
49. Mangalathu, S.; Burton, H.V. Deep learning-based classification of earthquake-impacted buildings using textual damage descriptions. *Int. J. Disaster Risk Reduct.* **2019**, *36*, 101111. [CrossRef]
50. Taalab, K.; Cheng, T.; Zhang, Y. Mapping landslide susceptibility and types using Random Forest. *Big Earth Data* **2018**, *2*, 159–178. [CrossRef]
51. Zhou, X.-Z.; Wen, H.-J.; Zhang, Y.-L.; Xu, J.-H.; Zhang, W.-G. Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. *Geosci. Front.* **2021**, *12*, 101211. [CrossRef]
52. El-Tawil, S.; Ibrahim, A.; Eltawil, A. Stick-Slip Classification Based on Machine Learning Techniques for Building Damage Assessment. *J. Earthq. Eng. JEE* **2021**, *26*, 5848–5865. [CrossRef]
53. Mahmoudi, S.N.; Chouinard, L. Seismic fragility assessment of highway bridges using support vector machines. *Bull. Earthq. Eng.* **2016**, *14*, 1571–1587. [CrossRef]
54. Deng, X.; Liu, Q.; Deng, Y.; Mahadevan, S. An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Inf. Sci.* **2016**, *340–341*, 250–261. [CrossRef]
55. Sun, D.-L.; Shi, S.-X.; Wen, H.-J.; Xue, J.-H.; Zhou, X.-Z.; Wu, J.-P. A Hybrid Optimization Method of Factor Screening Predicated on GeoDetector and Random Forest for Landslide Susceptibility Mapping. *Geomorphology* **2021**, *363*, 107623. [CrossRef]