



Economic Risk Assessment of Future Debris Flows by Machine Learning Method

Chenchen Qiu¹ · Lijun Su² · Alessandro Pasuto³ · Giulia Bossi³ · Xueyu Geng¹

Accepted: 6 February 2024 / Published online: 1 March 2024
© The Author(s) 2024

Abstract

A reliable economic risk map is critical for effective debris-flow mitigation. However, the uncertainties surrounding future scenarios in debris-flow frequency and magnitude restrict its application. To estimate the economic risks caused by future debris flows, a machine learning-based method was proposed to generate an economic risk map by multiplying a debris-flow hazard map and an economic vulnerability map. We selected the Gyirong Zangbo Basin as the study area because frequent severe debris flows impact the area every year. The debris-flow hazard map was developed through the multiplication of the annual probability of spatial impact, temporal probability, and annual susceptibility. We employed a hybrid machine learning model—certainty factor-genetic algorithm-support vector classification—to calculate susceptibilities. Simultaneously, a Poisson model was applied for temporal probabilities, while the determination of annual probability of spatial impact relied on statistical results. Additionally, four major elements at risk were selected for the generation of an economic loss map: roads, vegetation-covered land, residential buildings, and farmland. The economic loss of elements at risk was calculated based on physical vulnerabilities and their economic values. Therefore, we proposed a physical vulnerability matrix for residential buildings, factoring in impact pressure on buildings and their horizontal distance and vertical distance to debris-flow channels. In this context, an ensemble model (XGBoost) was used to predict debris-flow volumes to calculate impact pressures on buildings. The results show that residential buildings occupy 76.7% of the total economic risk, while road-covered areas contribute approximately 6.85%. Vegetation-covered land and farmland collectively represent 16.45% of the entire risk. These findings can provide a scientific support for the effective mitigation of future debris flows.

Keywords Economic risk · Future debris flows · Gyirong Zangbo Basin · Machine learning model · Physical vulnerability matrix · Southwest Tibet, China

1 Introduction

Debris flows, characterized by velocity ranging from 1 to 20 m/s, stand out as one of the most devastating landslide types, causing considerable losses to human lives and properties in mountainous areas (Hung et al. 2001; Papathoma-Köhle

et al. 2017; Laigle and Bardou 2022). The socioeconomic toll of such event is exacerbated by the rapid urbanization and the impacts of climate change witnessed in recent years (Hardwick Jones et al. 2010; Stoffel et al. 2014; Pei et al. 2023). Therefore, in order to reduce casualties and losses, adoption of hazard and susceptibility maps have been suggested by past studies (Mondal and Maiti 2013; Chen et al. 2015; Angillieri 2020). However, these tools fall short in indicating the potential distribution of damages (Remondo et al. 2005)—a crucial aspect even in post-disaster reconstruction efforts. Conducting an economic risk assessment becomes imperative to identify areas most susceptible to debris flows and quantify the potential economic losses associated with such occurrences. The economic risk assessment will allow decision makers to prioritize mitigation projects in high-risk areas where the economic impact is significant. Therefore, generating an economic risk assessment map of

✉ Xueyu Geng
Xueyu.Geng@warwick.ac.uk

¹ School of Engineering, University of Warwick,
Coventry CV4 7AL, UK

² Key Laboratory of Mountain Hazards and Earth Surface
Process, Institute of Mountain Hazards and Environment,
Chinese Academy of Sciences, Chengdu 610041, China

³ Research Institute for Geo-Hydrological Protection
(IRPI-CNR), National Research Council of Italy,
Padova 35127, Italy

debris flows is part of the mitigation strategies aiming at minimizing economic losses (Vranken et al. 2015).

Many methods have been proposed to assess the economic risks caused by debris flows. For example, Dai et al. (2002) conducted a comprehensive overview of landslide risk assessment, introducing an equation for evaluating property damage that is applicable to debris flows. This equation incorporates the annual probability of landslide ($P(H)$), the probability of spatial impact ($P(SIH)$), the vulnerability of properties ($V(PlS)$), and elements at risk (E). However, the temporal probability was not taken into consideration in the equation. This parameter signifies the likelihood of experiencing at least one hazard event within a specific time period (Guzzetti et al. 2006). Furthermore, Liu et al. (2009) calculated the property loss resulting from debris flows using numerical simulations to delineate the debris-flow influence area. Nevertheless, this study failed to assess the damage degree of buildings and infrastructure, potentially yielding unreliable results. Moreover, a run-out model was developed by Quan Luna et al. (2014) to ascertain debris-flow intensities, which were used to determine the dangerous zone. As a result, the economic risk can be calculated with the consideration of the physical damages to the buildings. However, the application of this established model is constrained by the changes in debris-flow frequency and magnitude. Therefore, there is a pressing need to derive a more dependable method to estimate economic risks from future debris flows. Our proposed model addresses the limitations highlighted in prior studies by incorporating temporal probability into risk assessment, evaluating the damage degree of buildings, and considering the model's applicability in future scenarios.

To quantify the economic risks, a supervised machine learning (ML)-based method was proposed. Machine learning is designed to find the complex and hidden relationships between input variables and output results (Khosravi et al. 2021). This is because a supervised model is well suited in this study due to the lesser amount of data and a clear labelled data for model training. Among the several supervised models, support vector machine (SVM) performs well in handling classification problem due to its high prediction accuracy and good performance in generalization. It has been proved that this model outperformed the other supervised models due to the highest accuracy when utilizing the UCI Repository of Machine Learning Databases to conducting classification analysis (Mohamed 2017). However, this method needs a good kernel function, therefore requires the optimal hyperparameters, C and gamma, for a specific classification problem. Therefore, the genetic algorithm (GA) is introduced to generate the optimal C and gamma so that we can obtain a good kernel function for classification problem and avoid overfitting (Burbidge and Buxton 2001).

Moreover, Mohamed (2017) also suggested that SVM has a high sensitivity to input data. In this case, certainty factor (CF) is employed to increase the stability of input data so as to develop a robustness model. Therefore, a hybrid machine learning model, certainty factor-genetic algorithm-support vector classification (CF-GA-SVC), was used in our study (Qiu et al. 2022). Additionally, an ensemble ML model, extreme gradient boosting (XGBoost), was used to predict the volume of future debris flows and estimate economic losses, thus generating a reliable economic risk map. XGBoost was developed based on Gradient Boosting Decision Tree (GBDT) with the involvement of a L2 regular term to avoid overfitting (Dong et al. 2022). As a result, this model has attracted a lot of attentions from various fields due to its good performance in both computational speed and prediction accuracy (Wang et al. 2022). For example, XGBoost performed better than logistic regression (LR) in the medical field (Wang et al. 2022). Additionally, XGBoost outperformed artificial neural network (ANN) and SVM in predicting groundwater levels (Osman et al. 2021). A better performance was found in utilizing XGBoost to predict concrete strength when compared to SVM and multilayer perceptron (MLP) (Nguyen et al. 2021). Overall, XGBoost is an effective machine learning model in conducting regression analysis. Therefore, this model was selected in our study to estimate the debris-flow volume of a future event.

The Gyirong Zangbo Basin in Tibet Autonomous Region, China, was selected as the study area to implement the proposed method. This area, characterized by an intricate geological and geomorphological environment, experiences frequent and severe debris flows annually.

2 Methodology

Generally speaking, risk assessment includes hazard assessment, vulnerability assessment, and definition of the elements at risk. Therefore, a conceptual equation defined by the International Association of Engineering Geology (IAEG) and Varnes (1984) was proposed:

$$R = H \times E \times V, \quad (1)$$

where R represents the risk. H is the hazard, and V represents the vulnerability. E is the selected elements at risk. Dai et al. (2002) further developed Eq. 1 for financial risk estimation by considering the product of $P(H) \times P(SIH)$ as the hazard, where $P(H)$ is the annual probability of debris-flow occurrence, which indicates the probability of an event occurring once a year, and $P(SIH)$ represents the annual probability of spatial impact. It is used to describe the probability of an occurred event that reaches an individual

property. However, this equation ignored the annual temporal probability ($P(H_t)$), which means that there is generally at least one event in a given period of time. Therefore, Vranken et al. (2015) reworked Eq. 1 to estimate the risk of property damage by utilizing $P(H)$, $P(S|H)$, and $P(H_t)$ to represent the hazard. However, the application of this equation under future scenarios was not considered. Since vulnerability is not a fixed value (Papathoma-Köhle et al. 2017), we defined a conceptual model to estimate the economic risk caused by future debris flows at a regional scale as follow:

$$Risk = P(H_s) \times P(H_t) \times P(S|H_t) \times V_e \times E \quad (2)$$

where $P(H_s)$ represents the annual spatial probability of a debris-flow event. $P(H_t)$ is the temporal probability of occurrence of one debris-flow event. $P(S|H_t)$ represents the annual probability of spatial impact, which means the probability of a debris-flow event impacting the individual property. The product of $P(H_s)$, $P(H_t)$, and $P(S|H_t)$ represents the hazard in Eq. 1. V_e is the future economic vulnerability, and E represents the elements at risk. Therefore, the economic risk was estimated by using Eq. 2.

Before calculating the different components in Eq. 2, debris-flow volume prediction is the first step to be carried out. Therefore, the selection of predisposing factors holds significance in achieving reliable predictions, as these factors can exert an impact on the ultimate prediction of debris-flow volume. Huang et al. (2020) highlighted strong correlations between some geomorphic features and debris-flow volume, such as catchment area (A), channel length (L), topographic relief (R_t), and mean slope of the main channel (J). Apart from these four factors, the curvature of the main channel (C) is also selected for volume prediction based on Shi et al. (2015). The values for A , L , J , and C can be obtained through a digital elevation model (DEM) with a resolution of 30 m.¹ R_t is decided using the Focal Statistics in GIS, requiring the determination of an optimal statistical unit. However, a rectangle statistical unit was employed in this study, and it contains $n \times n$ pixels ($n = 1, 2, 3, \dots$). Therefore, in order to decide the optimal statistical unit that can reflect the characteristic of the topographical condition, we utilized the change-point model. This method can decide the slope change point in the logarithmic curve that depicts the relationship between topographic relief and the neighborhood area. This change point is therefore regarded as the optimal statistical unit. The basic steps are:

- (1) The number of pixels of a statistical unit involved in calculation is set in an ascending order ranging from 2×2 to $n \times n$;

- (2) Decide the average topographic relief (r_j) and neighborhood area (s_j) of each statistical unit using the Focal Statistics tool;
- (3) Calculate the unit relief (T_j) of each statistical unit through the ratio of r_j and s_j ($j = 1, 2, 3, \dots, n$);

$$T_j = r_j / s_j \quad (3)$$

- (4) Take the logarithm for T_j , $\ln(T_j)$, and construct a new sequence X , X is $\{x_j, j = 1, 2, 3, \dots, n\}$. Then, calculate the mean value, \bar{X} , of the sequence X .

$$\bar{X} = \sum_{j=1}^n \frac{x_j}{n} \quad (4)$$

- (5) Calculate the statistical values of S and S_j by using the following equations to define the maximum difference between S and S_j ($j = 1, 2, 3, \dots, 20$). Separate the sequence X to $X_{j1} = (x_1, x_2, \dots, x_{j-1})$ and $X_{j2} = (x_j, x_{j+1}, \dots, x_n)$.

$$S = \sum_{j=1}^n (x_j - \bar{X})^2 \quad (5)$$

$$S_j = \sum_{t=1}^{j-1} (x_t - \bar{X}_{j1})^2 + \sum_{t=j}^n (x_t - \bar{X}_{j2})^2 \quad (6)$$

where \bar{X}_{j1} is the mean value of the first part of X_{j1} , and \bar{X}_{j2} is the average value of the rest part of X_{j2} . These equations are used to calculate the expected values, $E(S - S_j)$, $j = 1, 2, 3, \dots, n$. The point that causes the maximum difference between S and S_1 is called the change point. Once the optimal statistical unit is determined, the topographic relief for each catchment is calculated using GIS.

Following the identification of the optimal statistical unit, we can calculate the topographic relief within each catchment. Finally, for a catchment, the mean value of topographic relief across different optimal units is employed to represent its topographic relief.

Subsequent to data preparation, an ensemble algorithm (XGBoost) with robust and efficient computational abilities is introduced to predict the volume of future debris-flow events. The prediction performance is evaluated using absolute bias (AB), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE), which are widely employed indices for assessing model performance.

$$AB = |y_{i,\text{pre}} - y_{i,\text{true}}| \quad (7)$$

¹ Downloaded from Geospatial Data Cloud: <https://www.gscloud.cn/>.

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_{i,\text{pre}} - y_{i,\text{true}}| \quad (8)$$

$$\text{RMSE} = \sqrt{\sum_{i=1}^m (y_{i,\text{pre}} - y_{i,\text{true}})^2 / m} \quad (9)$$

$$\text{MAPE} = \frac{100\%}{m} \sum_{i=1}^m \left| \frac{y_{i,\text{pre}} - y_{i,\text{true}}}{y_{i,\text{true}}} \right| \quad (10)$$

where $y_{i,\text{pre}}$ is the predicted debris-flow volume, and $y_{i,\text{true}}$ represents the measured debris-flow volume. m is the number of prediction values.

2.1 Debris-Flow Hazard Assessment

Debris-flow hazard assessment is fundamental for the selection of possible safe areas to locate new residential buildings and infrastructure. Hazard is calculated considering the temporal probability ($P(H_t)$), the annual probability of spatial impact to the properties ($P(S|H_t)$), and annual spatial probability ($P(H_s)$).

2.1.1 Annual Spatial Probability ($P(H_s)$) and Annual Probability of Spatial Impact ($P(S|H_t)$)

A hybrid machine learning model (CF-GA-SVC) is employed to calculate the annual spatial probability, taking into account both annual precipitation and annual average temperature. As for the annual probability of spatial impact on a property, this parameter signifies the likelihood of impact when flowing materials reach a specific property (Corominas et al. 2014). The annual probability of spatial impact may vary across different catchments due to varying occurrence frequencies for debris-flow magnitudes. For buildings and roads, the debris-flow-related annual spatial impact is derived from statistical results and interviews during the field investigations conducted by Sichuan Geological Exploration Institute from 2006 to 2018. In the case of forests and pastures, we assume a probability of damage equal to 1, implying complete destruction when a debris flow occurs. Therefore, the annual probability of spatial impact for forests and pastures is represented by the reciprocal value of the recurrence period of a debris-flow event. While the recurrence period of a debris flow is decided based on the statistical analysis of historical records.

2.1.2 Temporal Probability ($P(H_t)$)

Temporal probability is associated with the occurrence of debris-flow events of the same magnitude at least once

within a specified time period, also known as the exceedance probability (Guzzetti et al. 2006). To calculate the temporal probability, the Poisson model proposed by Crovelli and Coe (2008) is introduced:

$$P[H(t) \geq 1] = 1 - P[H(t) = 0] = 1 - \exp(-\lambda t) = 1 - \exp(-t/\mu) \quad (11)$$

where $H(t)$ denotes the number of debris-flow events within the time period t . μ is the time interval between two debris flows, and λ is equal to $1/\mu$. This formula illustrates that if the time period t is long enough, there must be one more debris-flow event observed in a catchment when the parameter μ is a constant. Concurrently, an increase in μ signifies a diminishing probability of debris-flow occurrence within the specified time period t .

2.2 Future Economic Loss

Economic loss plays a crucial role in quantitative risk assessment, serving as the link between debris-flow occurrence and the elements at risk (Bednarik et al. 2012). To estimate potential economic loss (V_e) arising from future debris-flow events, we employ the multiplication of the economic values of elements at risk and their corresponding physical vulnerabilities (V_p). For residential buildings, physical vulnerability (V_p) is related to the impact pressure (P_i) induced by the flowing material, and the horizontal (HD) and vertical (VD) distance between the building and the nearest debris-flow channel. P_i serves as an effective indicator reflecting the energy of the phenomena and the potential degree of damage to buildings (Jakob et al. 2012; Kang and Kim 2016). In contrast, HD and VD reflect the impact intensity since the actual damage is generally more significant if the building is closer to the debris-flow channel (Sturm et al. 2018).

As a first step, we propose a physical vulnerability matrix (Table 1) to assess damages to buildings. Then, the impact pressure on buildings is estimated based on the predicted debris-flow volumes. Finally, the economic loss can be appraised by multiplying physical vulnerability and unit price.

The physical vulnerabilities (V_p) of roads, farmland, and pastures are defined on the basis of fragility values as defined in Totschnig and Fuchs (2013). In this context, fourth-grade road is assigned a fragility value of 0.85, while the third-grade road has a value of 0.65. Moreover, it is assumed that pastures and farmland would suffer complete destruction in the event of a debris-flow occurrence.

The economic loss of an element caused by one debris-flow event is defined as follow:

$$V_e = \sum_i^m V_{ei} \quad (12)$$

Table 1 Physical vulnerability (V_p) of buildings in the Gyirong Zangbo Basin

P_i (kPa)	Building structure	$HD < 30$ m			$30 < HD < 100$ m		
		$VD < 4$	$4 < VD < 10$	$10 < VD < 15$	$VD < 4$	$4 < VD < 10$	$10 < VD < 15$
< 30	RC frame	0.3	0.2	0.1	0.2	0.1	/
	Non-RC frame	0.8	0.7	0.6	0.7	0.6	0.4
30–70	RC frame	0.6	0.5	0.4	0.5	0.4	0.2
	Non-RC frame	1	0.9	0.8	0.9	0.8	0.6
70–100	RC frame	0.7	0.6	0.5	0.6	0.5	0.3
	Non-RC frame	1	1	0.9	1	0.9	0.7
> 100	RC frame	0.8	0.7	0.6	0.7	0.6	0.4
	Non-RC frame	1	1	0.9	1	1	0.8
P_i (kPa)	Building structure	$100 < HD < 160$ m			$160 < HD < 230$ m		
		$VD < 4$	$4 < VD < 10$	$10 < VD < 15$	$VD < 4$	$4 < VD < 10$	$10 < VD < 15$
< 30	RC frame	0.1	/	/	/	/	/
	Non-RC frame	0.6	0.4	0.1	0.4	0.1	/
30–70	RC frame	0.4	0.2	/	0.2	/	/
	Non-RC frame	0.8	0.6	0.3	0.6	0.3	/
70–100	RC frame	0.5	0.3	/	0.3	/	/
	Non-RC frame	0.9	0.7	0.4	0.7	0.4	/
> 100	RC frame	0.6	0.4	0.1	0.4	0.1	/
	Non-RC frame	1	0.8	0.5	0.8	0.5	0.1

RC frame represents the reinforced concrete buildings. Non-RC frame includes masonry, wooden, and light steel frame structures

$$V_{ei} = V_{pi} \times V_{fi} \quad (13)$$

$$V_{fi} = \sum_i^m P_i \times A_i \quad (14)$$

where V_{ei} and V_{fi} represent the economic loss and the economic value of each element, respectively. V_p represents the physical vulnerability of an element. P_i is the price per km^2 , and A_i is the area of each element.

3 Study Area

The Gyirong Zangbo Basin in the Tibet Autonomous Region of China was selected to test the proposed methodology and assess the economic risk related to future debris flows. This area covers 2120 km^2 , including two towns and 13 villages (Fig. 1a).

The lithology class belongs to the Mesozoic era, in which the Cretaceous (K_{2zz} and K_{1j}) mainly tended east-west (EW) direction. Cretaceous formations, primarily comprised of shale and sandstone interbedded with limestone, dominate this area. Shale, known for its fragility and erodibility, contributes to the extensive availability of sediment for transport

on the slopes of the northern section of the Gyirong Zangbo Basin. The widely distributed cracked rocks is a result of the combined effects of intense weathering and active faults (Qiu et al. 2022), potentially leading to the mobilization of sediment during heavy rainfall and the subsequent development of debris flows. In contrast, the southern part of the Gyirong Zangbo Basin exhibits a markedly different morphology, characterized by widely distributed vegetation-covered slopes. This geomorphic feature is attributed to warm air from the Indian Ocean, fostering abundant rainfall and creating an optimal environment for vegetation growth. In this southern region, forests are primarily distributed at altitudes below 2900 m. Beyond the 4100 m threshold, where glaciers are present, no forests are observed. In addition, the southern part is characterized by schist, gneiss, granulite, and migmatite (A_nZN). Both schist and gneiss belong to metamorphic rock, exhibiting a lamellar structure that results in low shear strength along the dips, leading to unstable slopes. Therefore, the expansive source areas and prevailing water conditions create an environment conducive to the initiation of debris flows in this southern region.

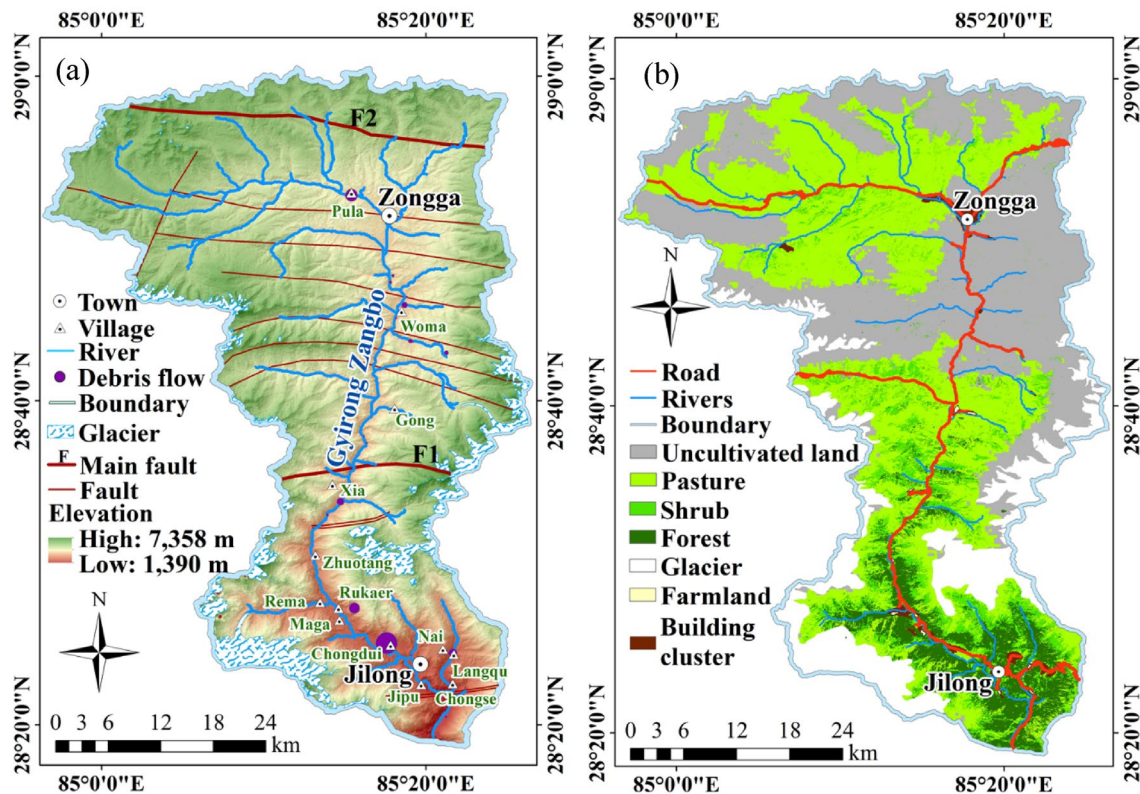


Fig. 1 **a** Location of the study area in southwestern Tibet, China; **b** Distribution of the elements at risk in the study area

4 Economic Risk Assessment

Economic risk assessment aids in quantifying the severity of potential economic losses associated with future debris flows. This section illustrates the calculation results of economic risks and provide guidance for helping risk managers, supporting decision-making process, and allocating resources effectively.

4.1 Characteristics of the Selected Elements at Risk

Four elements at risk are selected in the study area to evaluate future economic risk, including roads, residential buildings, vegetation-covered land (forest, shrub, and pasture), and farmland. The distribution characteristics of the selected elements are shown in Fig. 1b. In the study area, shrub, similar to forests and pastures, is assigned a fragility value of 1.0, implying complete destruction in the event of debris flows. Therefore, shrub, forest, pasture, and farmland are considered completely destroyed when debris flows occur. Other land cover types, such as uncultivated land and glacier, are not incorporated into the economic risk assessment

due to challenges in estimating their economic values, as they are unsuitable for farming and grazing. The widely distributed crushed rocks in the northern part cannot serve as the foundation of houses or infrastructure and therefore are not considered in such calculation. River networks are also excluded from risk assessment because there are no planned hydropower plants or fishery industries.

Generally, the roads in the Gyirong Zangbo Basin fall into two main categories according to the Standard of Reconstruction Project for County-Level and Village-Level Roads in China: county-level highways and backroads. This area encompasses two towns and 13 villages, with most buildings situated on alluvial fans or along streams. Alluvial fans are commonly chosen for residential sites in mountain environments due to their gentle slopes, smooth topography, and proximity to rivers benefiting farming and living needs (Schick et al. 1999; Marcato et al. 2012) but inevitably suffer a higher risk of debris flows. Livestock is the primary income source for the local people and a large area of the basin is pasture, which is mainly distributed in the northern part (Fig. 1b).

Table 2 Database of the prediction model in the Gyirong Zangbo Basin, including four geomorphic factors (A is catchment area, L is main channel length, R_t represents topographic relief, and J is main channel gradient) and measured volumes of debris flows (V) through field investigations

No.	A (km ²)	L (km)	R_t (m)	J (°)	V (10 ⁴ m ³)	No.	A (km ²)	L (km)	R_t (m)	J (°)	V (10 ⁴ m ³)
1	8.55	3.13	269	6.0	3.5	26	3.99	3.78	134	22.1	0.7
2	4.68	1.41	126	12.2	2.3	27	2.88	2.40	313	35.4	1.1
3	12.88	4.16	269	7.1	5.4	28	0.34	1.14	163	40.6	2
4	0.29	0.50	95	9.3	0.28	29	2.81	2.84	253	27.7	2.3
5	0.29	0.29	200	22.4	0.28	30	7.18	4.82	400	27.2	3.7
6	5.73	0.71	260	6.7	4.5	31	24.42	9.47	337	17.5	0.3
7	0.56	0.62	195	13.9	0.8	32	2.81	1.74	205	17.7	0.7
8	2.15	0.73	250	15.3	0.5	33	0.43	1.30	200	38.7	0.5
9	0.32	0.46	276	28.6	0.2	34	7.06	4.41	275	24.1	3
10	1.67	0.95	161	20.3	0.61	35	1.07	2.05	225	23.9	2.2
11	11.21	1.93	360	8.6	1.8	36	0.86	2.17	149	21.7	2.47
12	2.85	1.57	232	14.4	0.72	37	6.51	2.92	252	26.7	4.41
13	2.29	1.84	189	19.7	1.27	38	0.42	1.64	151	25.7	8.17
14	0.08	0.42	240	19.6	0.18	39	0.51	1.43	153	26.1	8.42
15	0.18	0.48	366	34.9	0.05	40	0.20	0.76	130	28.9	0.92
16	0.53	0.81	170	16.4	0.4	41	0.34	1.25	130	26.3	1.25
17	0.71	1.74	151	33.0	7.1	42	0.05	0.18	85	10.8	0.15
18	0.49	1.64	162	33.0	8.3	43	0.06	0.23	81	16.9	0.16
19	0.60	1.52	155	32.9	6.2	44	0.33	0.5	162	15.6	0.29
20	0.36	1.15	261	39.4	3.2	45	0.05	0.2	107	14.9	0.14
21	2.73	2.57	190	34.1	2	46	1.37	1.11	160	10.0	1.54
22	2.02	2.59	198	35.1	2.2	47	4.83	1.96	277	11.7	1.4
23	0.43	1.30	198	37.7	1.95	48	1.33	0.5	258	27.1	0.5
24	0.19	1.09	181	34.5	1.6	49	0.17	0.62	231	25.3	0.2
25	1.03	2.02	232	23.6	1.2						

The total number of training samples is 49

4.2 Volume Prediction and Prediction Model Assessment

The DEM used to derive parameters A , L , J , and C features a spatial resolution of 12.5 m.² As for R_p , the optimal statistical unit for calculation results in this study was determined to be 10×10 pixels. Prior to model training with the identified factors, it is imperative to conduct correlation analysis to ascertain their suitability for predicting debris-flow volume. The relationships between each factor and debris-flow volume were examined to ascertain their suitability using the data in Table 2. The analysis results revealed that the catchment area (A) factor exhibits the most robust linear relationship with volume since Person's coefficient is greater than 0.8. The length of the main channel (L) follows, presenting some correlation with volume, as well as the topographic relief (R_t) and mean slope of the main channel (J), although weaker. No correlation was observed between the mean curvature of the main channel (C) and debris-flow volume.

Therefore, C was excluded, and only A , L , R_p , and J were used to develop the debris flow volume prediction model.

Following the correlation analysis, the data were split into a training set and a testing set with a ratio of 7:3. The training set was used to develop a prediction model through tenfolds cross-validations and leave-one-out cross-validations, while the testing set served to evaluate the performance of the prediction model. Following this, the model's performance was further evaluated, demonstrating a reduction in MAPE value to 9.32%. Regarding MAE and RMSE, these two metrics are 501.4 m³ and 594.4 m³, respectively. The MB was calculated as −50.6 m³, indicating an underestimation of debris-flow volume but with a relatively small average deviation from the measured value. The prediction results are plotted against the measured values in Fig. 2a. This figure shows that the prediction results well agree with the measured values. Therefore, it can be concluded that this model performed well in predicting the volume of a future debris-flow event. In addition to assessing the overall model performance, it is equally crucial to discern the individual contribution of each geomorphic factor to the estimation of debris-flow volume. This analysis of factor importance not only

² Accessed at <https://search.asf.alaska.edu/#/>.

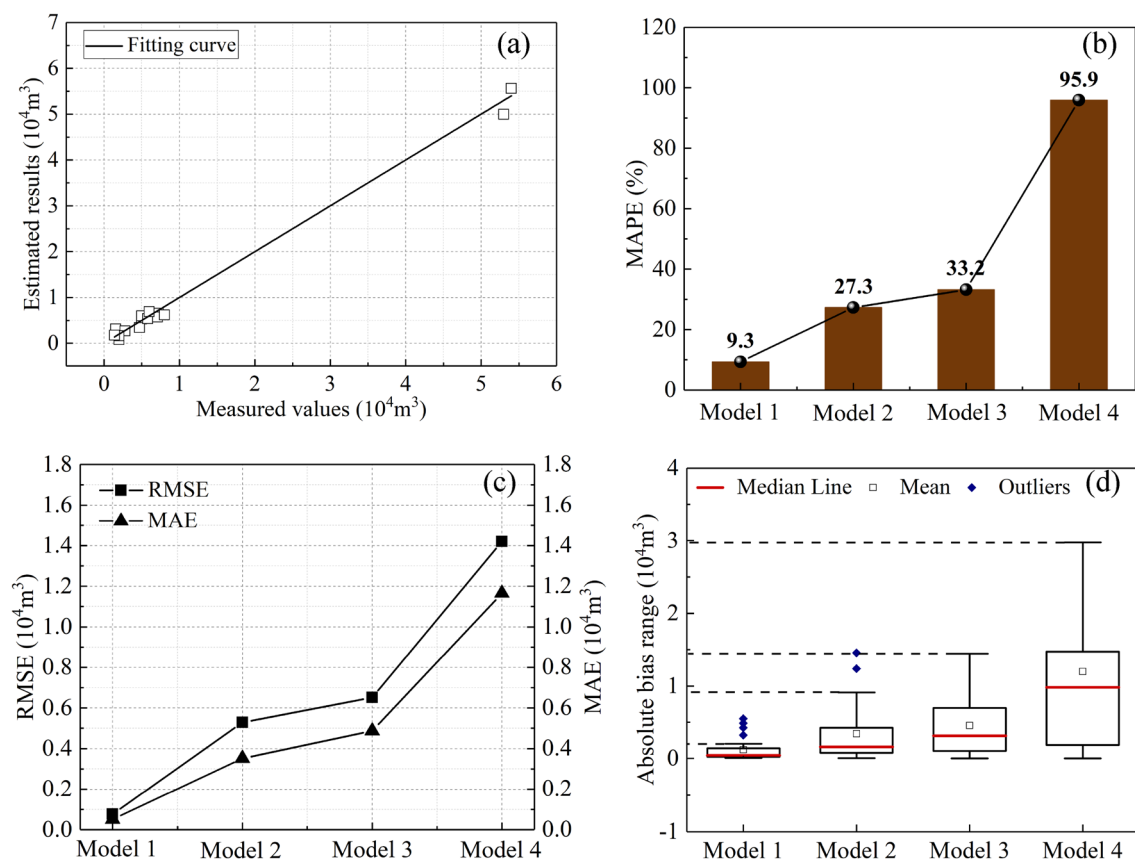


Fig. 2 **a** Scatter diagram of prediction results; **b** MAPE results of the four models; **c** RMSE and MAE results of the four models; **d** AB results of the four models (MAPE mean absolute percentage error; RMSE root mean square error; MAE mean absolute error; AB means absolute bias)

illuminates the significance of each raw factor but also provides valuable insights for utilizing the developed model in estimating debris-flow volumes. To accomplish this, we assessed the significance of each raw factor by systematically excluding one factor at a time, thereby training four models. They are Model 1 ($A + L + R_i + J$), Model 2 ($A + L + R_i$), Model 3 ($A + L$), and Model 4 (A). Subsequently, we assessed the estimation accuracy of the four models using the RMSE, MAE, AB, and MAPE indices, and the corresponding results are depicted in Fig. 2b–d. Model 1 stands out as the most effective model, whereas the estimation outcomes of Model 4 demonstrate the most significant deviation from the measured values. To conduct a more comprehensive performance comparison among the four models, we calculated the MAPE, RMSE, and MAE values (Fig. 2).

Figure 2b shows an 193.5% increase of MAPE when J was excluded from model training (Model 2), which indicates a significant decrease of prediction accuracy. Another accuracy decline of 15.8% was noted when factor R_i was omitted from the factor combination (Model 3). It can be concluded that J plays a more critical role in enhancing estimations compared to R_i . Furthermore,

a significant reduction in prediction accuracy was found when only A was used (Model 4) for debris-flow volume estimation, reaching 188.8% decline of prediction accuracy when compared to Model 3. Although catchment size may suggest the potential water storage and the maximum capacity for erodible debris materials within a debris-flow catchment, the scaling relationship between catchment areas and debris-flow volume is intricate (Gartner et al. 2008; Marchi et al. 2019; Lee et al. 2021). This complexity poses a challenge in relying on a single factor for achieving a reliable estimation of debris-flow volume. Figure 2c illustrates the variations in RMSE and MAE for the four models, showcasing an upward trend as the number of input variables decreases. The RMSE values for all four models consistently surpass the MAE values. This discrepancy arises because RMSE amplifies estimation errors, particularly those that are relatively significant. Consequently, the RMSE index is well-suited for detecting outliers, whereas MAE values typically represent the exact errors between the estimation results and raw data. Therefore, the increase in RMSE values signifies greater estimation errors and instability in estimation performance after excluding input variables (refer to Fig. 2d).

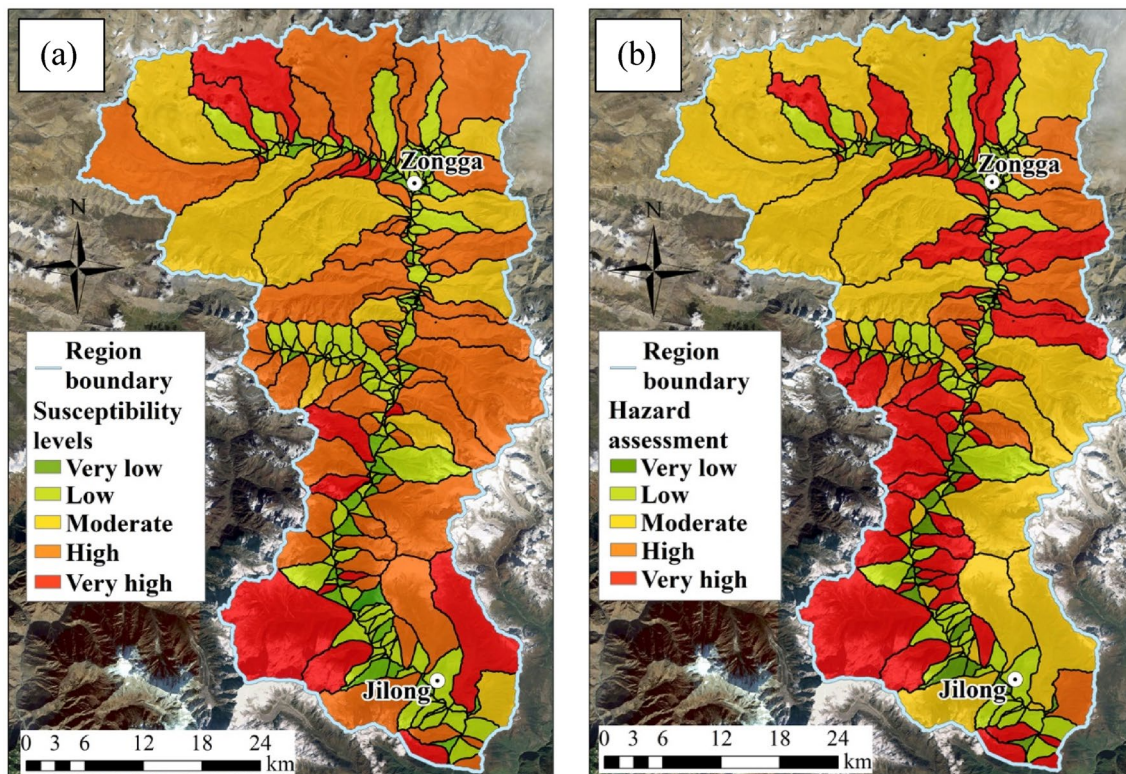


Fig. 3 **a** Susceptibility levels in the Gyirong Zangbo Basin using certainty factor-genetic algorithm-support vector classification (CF-GA-SVC); **b** Debris-flow hazard map

As depicted in Fig. 2d, Model 1 demonstrates the most stable outputs, with estimation errors ranging from 0.0055×10^4 to $0.2031 \times 10^4 \text{ m}^3$. Excluding J , the maximum estimation error increases to $0.9112 \times 10^4 \text{ m}^3$, notwithstanding the presence of two abnormal values, $1.2394 \times 10^4 \text{ m}^3$ and $1.4541 \times 10^4 \text{ m}^3$, which are the primary contributors to the rise in RMSE in Fig. 2c. For Model 3, which excludes J and R_i from the model training, there is no significant decrease in the MAPE value (refer to Fig. 2b), but the maximum estimation error increases to $1.4452 \times 10^4 \text{ m}^3$. Model 4 exhibits a maximum estimation error of $7.6917 \times 10^4 \text{ m}^3$. Consequently, the developed prediction model with the incorporation of four factors performs well in estimating debris-flow volume, and 49 training samples could support the development of a reliable prediction model in the Gyirong Zangbo Basin.

4.3 Debris-Flow Hazard Assessment

Volume prediction of future debris flows can help in determining the annual probability of an event affecting the vulnerable elements and its exceedance probability within a certain time range. The calculation of susceptibility estimates the occurrence probability of this event.

4.3.1 Annual Probability of Spatial Impact to Properties and Temporal Probability

The prediction results illustrate that the volumes of future debris-flow events in this area are generally smaller than $20 \times 10^4 \text{ m}^3$. According to the Specification of Geological Investigation for Debris-Flow Stabilization (DZT0220-2006),³ debris flows are categorized as medium-scale when the volume ranges from 2×10^4 to $20 \times 10^4 \text{ m}^3$, while small-scale debris flows have a volume smaller than $2 \times 10^4 \text{ m}^3$. The field investigations indicate that small-scale debris flows occur approximately once every 2 years in this area, and the return period of medium-scale debris flows is 5 years.

For the temporal probability, a 5-year time period (t) was selected. Within this timeframe, the time intervals (μ) of small-scale and medium-scale debris flows were determined as 2 and 5 years, respectively. Therefore, the temporal probability of 0.91 was assigned for small-scale debris flows, and 0.63 for the medium-scale debris flows based on Eq. 11. This allows for the determination of the temporal probability of a

³ <http://www.gsyskc.com/uploadfiles/file/20170109/1483928553805018529.pdf> (in Chinese).

potential debris-flow event in each catchment based on the predicted debris-flow volume.

4.3.2 Annual Spatial Susceptibility and Debris-Flow Hazard Assessment Map

The causative factor combination in this area has been suggested by Qiu et al. (2022). Figure 3a displays a susceptibility map of the area with five levels, generated using the CF-GA-SVC model. The occurrence probability of debris flows ranges from 0.0028 to 0.9811. The high susceptibility level is delineated within the range of 0.7516 to 0.9533, while values between 0.9533 and 0.9811 indicate a very high susceptibility level. Apart from the high and very high susceptibility regions, the probabilities of very low and low levels are 0.0028–0.0404 and 0.0404–0.2887, respectively. The moderate level falls within the range of 0.2887–0.7516. The results highlight that the very-high susceptibility level is predominantly distributed in the southern part, attributed to the glacier's location and abundant rainfall, both capable of triggering debris flows. In recent years climate change has promoted glacier degradation and thus generated more exposed rocks and great availability of loose materials on slopes. These materials can be mobilized by water runoff or involved by mass movements, thus forming debris flows. Additionally, glacier melting, influenced by rainfall and rising temperatures, contributes to increased debris-flow occurrences. This is because a 0.5 °C increment of temperature may cause a 6.6% increase of annual streamflow based on the study of Zhang et al. (2011) when annual precipitation is a fixed value. As a result, the mobilization of the accumulated and settled materials within the catchments may also increase, which results in the increasing susceptibility. The northern part of the study area exhibits a higher susceptibility to debris flows, primarily due to widely distributed crushed rocks associated with tectonic activity. This sector features one main fault and seven secondary faults, resulting in large amounts of loose materials accumulating along slopes or at their bases. These settled materials can start to move under the effect of heavy rainfalls and glacial-melting runoff, evolving into saturated/partly saturated flows.

The hazard map resulting from the product of susceptibility, annual probability of spatial impact, and temporal probability shows values ranging from 0.0013 to 0.4415 (Fig. 3b). These values are classified into five levels—very high (0.3720–0.4415), high (0.1939–0.3720), moderate (0.0605–0.1939), low (0.0182–0.0605), and very low (0.0013–0.0182)—based on natural break point method. The catchments with very high hazard levels are mostly concentrated in the southern part.

4.4 Economic Loss Map

Using the predicted debris-flow volumes, the equations proposed by Cui et al. (2013) and Rickenmann (1999) allow for the calculation of peak discharge and debris-flow velocity. The flow depth can also be defined based on its velocity and the slope gradient of the main channel (Koch 1998; Rickenmann 1999). Consequently, the catchments with building clusters were extracted to calculate the impact pressure of future debris-flow events (Zanchetta et al. 2004) and to determine the damage degree of the buildings affected by debris flows. The potential impact pressures resulting from future debris flows are depicted in Fig. 4a. However, to assess the physical vulnerability of the residential buildings, *HD* and *VD* need to be defined. Building cluster polygons were extracted based on satellite images (Gaofen-2). The used panchromatic (black and white) images have a spatial resolution of 0.8 m. Furthermore, the Fishnet tool in GIS was used to divide these building clusters into rectangles, each representing a single building. Considering the average house type, a 500 m² threshold was applied to generate building rectangles (see building segments in Fig. 4a). Consequently, the physical vulnerability of each building was determined based on *HD* and *VD* values, along with the calculated potential impact pressure from a future debris-flow event. Apart from the residential buildings, the physical vulnerabilities of roads, farmland, and vegetation-covered land were decided based on the suggested fragility values by Totschnig and Fuchs (2013). The economic loss of each element was estimated based on the unit price in Table 3. The unit prices of buildings and roads were decided based on the meeting with the Housing and Urban–Rural Construction Agency in Gyirong during the field investigations. The unit price of vegetation-covered land was obtained through the meeting with the Forestry and Grassland Administration of Tibet, mainly including the restoration cost after the destruction by debris flows. “Proportion” in Table 3 illustrates the percentage of the area occupied by each element in relation to the overall study area.

Based on the unit price of each element, we can estimate the economic loss of each rectangle. Subsequently, the economic loss of each building segment can be calculated based on Eqs. 11, 12.

The economic losses of the catchments with building segments are presented in Fig. 4b. Since catchment was used as an analysis unit rather than a grid unit, the catchment unit can effectively reflect the hazard-inducing environment. As depicted in Fig. 4b, Zongga and Jilong do not emerge as residential areas with the most severe economic losses. This is primarily attributed to the majority of houses being constructed with reinforced concrete

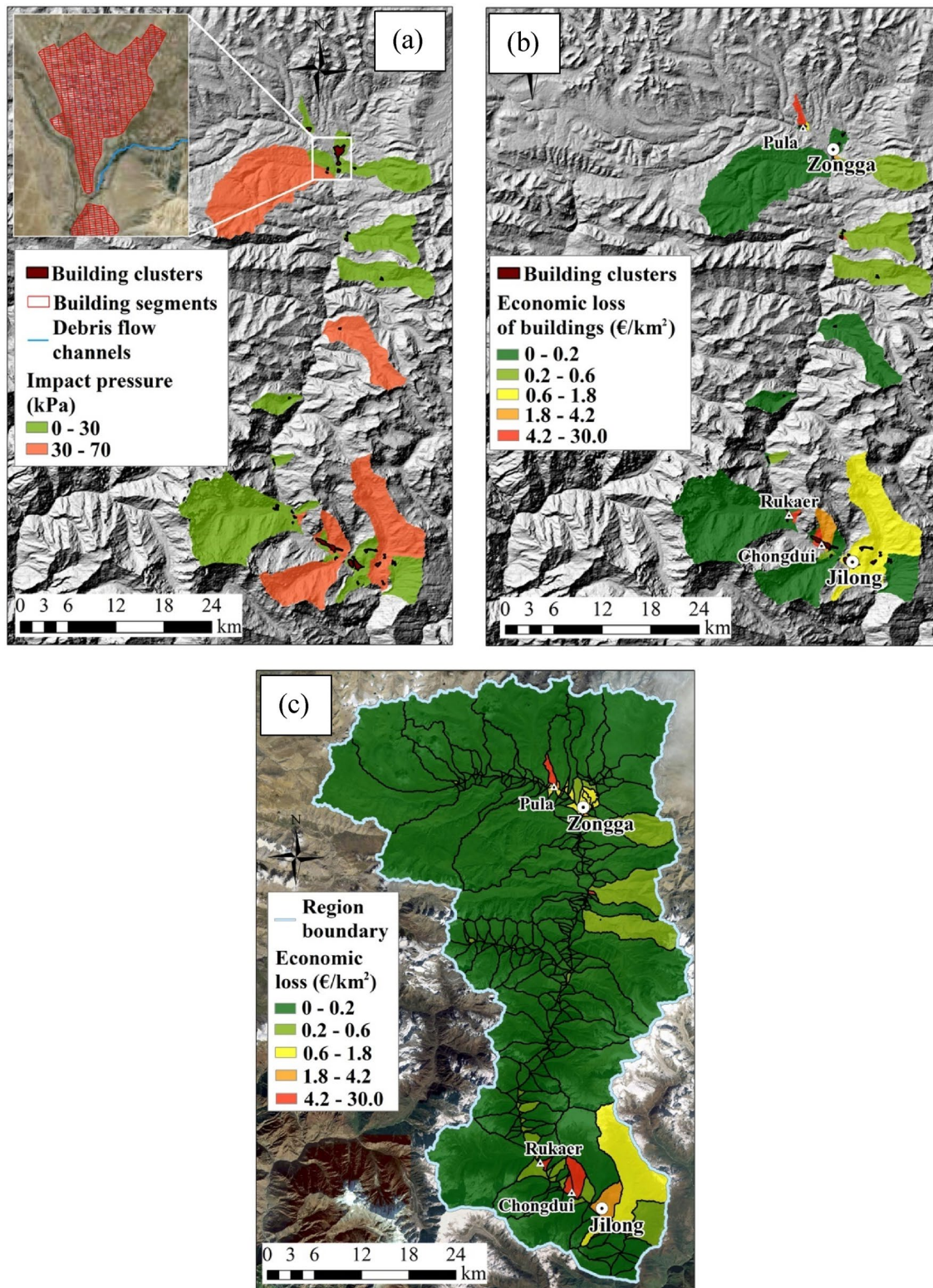
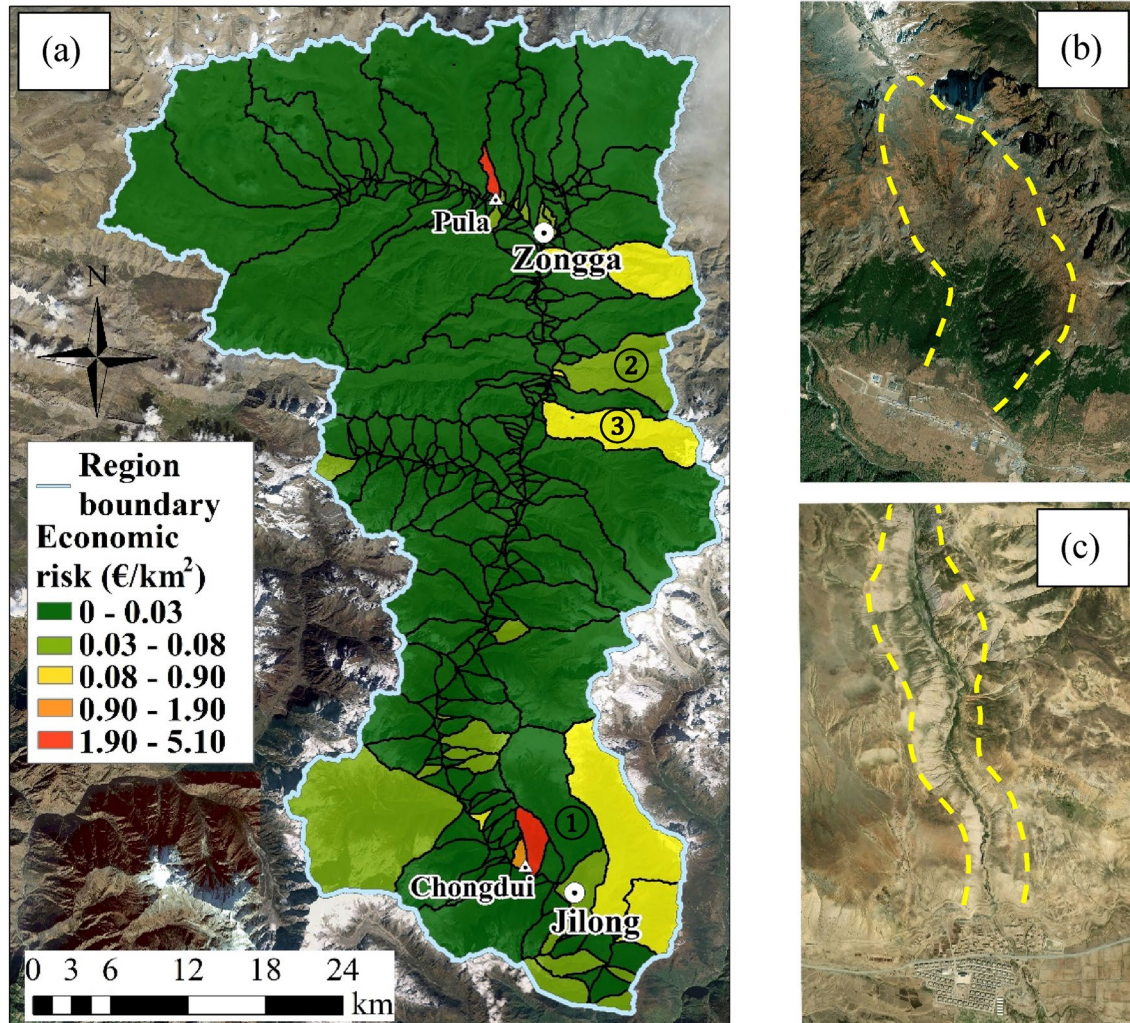


Fig. 4 **a** Impact pressures caused by future debris flows to buildings, and the economic vulnerabilities of **b** the extracted catchments and **c** all the catchments in the Gyirong Zangbo Basin

Table 3 Unit price of the elements at risk in the Gyirong Zangbo Basin

Elements	Categories	Area and length	Proportion (%)	Unit price	Value based on
Roads	Third-grade road	186.2 km	13.8	83,460 €/km	Repair cost
	Fourth-grade road	50.8 km	3.8	46,320 €/km	Repair cost
Buildings	Residential building	6.0 km ²	0.5	417.140 €/m ²	Average house price
Vegetation-covered land	Pasture	764.3 km ²	56.7	0.019 €/m ²	Official data
	Shrub	154.7 km ²	11.5	0.089 €/m ²	Official data
	Forest	165.9 km ²	12.3	0.208 €/m ²	Official data
Other types of lands	Farmland	20.9 km ²	1.5	2.090 €/m ²	Production value loss

**Fig. 5** **a** Potential economic risk caused by future debris flows in the Gyirong Zangbo Basin; **b, c** Catchments with very high risk level (1.90–5.10 €/km²), where the yellow dashed line depicts the catchment boundary

(RC frames), with the structural strength to withstand the impact of future debris flows. Similar distribution characteristics were found when all the elements at risk were considered (Fig. 4c). However, the involvement of other elements at risk can also increase the economic losses. For

example, the vulnerability in Chongdui reaches the highest level (4.2–30 €/km²), but the economic loss of buildings in this area only ranges from 1.8 to 4.2 €/km² (Fig. 4b). Additionally, an increase of 0.74 €/km² and 1.08 €/km² are found in Jilong and Zongga, respectively. Nevertheless,

the overall levels of economic loss in these two areas exhibit no significant change when compared to Fig. 4b. The areas with the highest economic loss (4.2–30.0 €/km²) are mainly situated in the residential sites, such as Zongga, Jilong, Chongdui, and Pula. Farmland is distributed closely to residential buildings, further contributing to the escalation of economic loss of the catchments. Therefore, it can be stated that residential buildings are responsible for the principal economic losses.

5 Economic Risk Map

The overall economic risk is presented in Fig. 5. The economic risk levels of 0.9–5.1 €/km² cover an area of 13.3 km², with the high risk level (0.9–1.9 €/km²) encompassing 19.5% of the territory. The largest area is covered by the very low risk level class (0–0.03 €/km²), with 1722.4 km² and comprising 261 catchments. The low risk level (0.03–0.08 €/km²) covers an area of 200.2 km² with 22 catchments. The distribution characteristics of risk levels in Fig. 5a align with the historical debris-flow distribution in Fig. 1a, as expected. Local inhabitants prefer settling on alluvial fans to meet the demands of planting and farming, as revealed by field investigations (Fig. 5b, c). This choice allows people to reside close to rivers, benefiting cattle and sheep husbandry. However, the potential economic risks stemming from future debris flows may lead to substantial property losses, particularly in residential areas with non-reinforced concrete frames. Therefore, it can be concluded that socioeconomic development can promote the accumulation of valuable elements, resulting in high and very high economic risks due to future debris flows.

Furthermore, we unveiled the contributions of the four major elements to the total economic risk. As depicted in Fig. 2, pastures cover 56.7% of the total area, with building clusters representing the smallest percentage at 0.5%. However, pastures account for the smallest proportion of the total economic risk at 2.49%, while building clusters are responsible for the largest share, constituting 76.7% of the entire economic risk. Forests contribute approximately 0.72 €/km², exceeding shrubs at 0.33 €/km². Roads, covering almost 0.2% of the total area, contribute about 4.36% of the total economic risk by third-grade roads, while fourth-grade roads are responsible for 2.49% of the total economic risk. Farmland, covering an area of 20.9 km², poses an economic risk of 0.75 €/km².

To validate the economic risk results, several catchments identified during the field investigations conducted by Sichuan Geological Exploration Institute are examined. As shown in Fig. 5a, a debris-flow event occurred in catchment 1, with a volume smaller than 2×10^4 m³, classified as a small-scale debris flow. Therefore, the annual probability

of spatial impact and temporal probability are 0.55 and 0.91, respectively, and the susceptibility of this catchment is 0.9423. The field investigation estimated an economic loss of 3.85×10^6 €. In this case, the calculated economic risk is 0.01 €/km², aligning with the classified risk level in Fig. 5a. For catchment 2 in Fig. 5a, a property loss of 3.96×10^6 € was estimated during the field investigations. The loss was caused by a small-scale debris flow (estimated volume is 1.7×10^4 m³). The susceptibility of this catchment is 0.74. Therefore, the economic risk is 0.04 €/km², falling within the range of 0.03–0.08. Moreover, a total of six debris flows were observed on the same day in catchment 3, most of which are hill-slope debris flows causing damages to residential areas and roads. Therefore, these events are all small-scale but claimed a total property loss of 10.2×10^6 €. The susceptibility in this catchment is 0.94. Consequently, the economic risk is estimated at 0.13 €/km², placing it in the moderate risk level (0.08–0.90 €/km²).

These results can indicate the effectiveness and reliability of the proposed method in estimating economic loss risk related to future debris flows and therefore provide guidance for decision makers about site selection of future infrastructure and countermeasure construction.

6 Discussion and Limitations

The integration of a hybrid CF-GA-SVC model and an ensemble XGBoost model serves to enhance the reliability of the presented results. To be more specific, the hybrid ML model (CF-GA-SVC) demonstrates an effective improvement in the prediction accuracy of susceptibility compared to traditional approaches (Vranken et al. 2015; Fu et al. 2020) and the conventional ML models (Staley et al. 2017). This is because the complexity of the debris-flow event occurrence process poses a significant challenge for susceptibility assessment using simple empirical methods. Therefore, generating a reliable debris-flow hazard map through such approaches may be unrealistic. In contrast, the introduction of ML methods, with their robust ability to handle complex spatial heterogeneity problems, facilitates the establishment of more reliable and precise relationships among numerous variables. Furthermore, the hybridization of different individual ML models enhances the robustness and accuracy of each single ML model (Ardabili et al. 2020). This hybrid ML method offers a scientific, advanced, and systematic means to describe the triggering conditions of debris flows.

Additionally, a proper vulnerability evaluation is fundamental for estimating economic risk. Past studies repeatedly focused on physical vulnerability using different approaches (Papathoma-Köhle et al. 2017) but further studies may be needed. If we can predict the

volume of future debris-flow events, the impact pressure on buildings can be defined. To achieve this, we introduced an ensemble ML model (XGBoost), which is composed of a series of regressors to extract the relationship between geomorphological conditions and possible debris-flow volume. The subsequent step involves defining a factor combination essential for analyzing impacts on residential buildings, thereby supporting the formulation of the economic risk map. We are aware that impact pressure, along with *HD* and *VD*, cannot fully encompass the complex effects of a debris flow on the buildings and therefore we proposed this synthetic procedure that can provide an average-based assessment of the damage degree that can occur in a residential area based on the construction typology. However, sample size may emerge as an uncertainty to impact the performance of models in estimating debris-flow volume. Therefore, we incorporated 32 more historical debris-flow events in Sichuan, China into model training to test the reliability of the developed model. The performance of this model with 81 debris-flow samples was further assessed, with MAPE reaching 9.64%, which is slightly higher than the MAPE value when 49 samples were used for model development. This result indicates that the predictions are accurate in relative terms, which is a favorable aspect of its performance. The MAE and RMSE are 560.7 m³ and 776.3 m³, respectively. The MB is -90 m³, which indicates an underestimation of the debris-flow volume but relatively a small average deviation from the measured value. Therefore, the incorporation of more training samples may result in a slight decrease in prediction accuracy without impacting significantly on the prediction efficiency, which demonstrates the reliability of the developed model based on 49 training samples. We are also aware that 49 or even 81 samples are still not enough if a wider application of this model is expected. This is because the completion of this task needs continuous input of debris-flow data in different areas and regions globally, which may not be fully completed in this study. However, the model developed in this study can support the effective risk assessment in the Gyirong Zangbo Basin and even the Himalayan regions due to the similar topographic and meteorological conditions, and we will also keep working on this model to increase its robustness so that we can use the model to achieve reliable predictions at a national scale or continental scale. Overall, the incorporation of the 32 more historical debris flows can increase the robustness of the developed prediction model but cannot change the prediction results to a large extent. The predicted economic vulnerabilities and risks of the catchments show slight changes without causing changes to the classified vulnerability and risk levels.

The economic analysis presented thus far serves to map risk levels in various catchments, offering effective

guidance and scientific support for decision makers. By quantifying potential risks, decision makers can gain a better understanding of future challenges, enabling them to prioritize actions and optimize resource allocation in high-risk zones. This approach facilitates long-term urban planning, policy development, and the formulation of adaptation strategies to effectively reduce and manage identified risks. Moreover, the preparedness and emergency response system would be implemented accordingly. Despite these strengths, some limitations persist, suggesting room for improvement in the proposed methodology's performance. The database of debris-flow occurrences needs further enrichment to refine the volume prediction model. Another potential area for enhancement lies in augmenting the physical vulnerability assessment with new data, considering additional building characteristics such as shape and the number of windows. Nevertheless, all these limitations cannot alter the fact that the proposed ML-based method represents a new tool for generating a map of economic risk caused by future debris-flow events. It also signifies a practical method to deliver accurate and reliable warnings to local residents about the risks posed by debris flows.

7 Conclusion

In this study, a machine learning-based method was proposed with the integration of a hybrid and ensemble machine learning model to produce a map of economic risk due to future debris-flow events. This map was derived through the assessment of debris-flow hazards and the calculation of economic losses associated with the elements at risk. Completing the debris-flow assessment involves multiplying the annual spatial probability of debris-flow occurrence, the annual probability of spatial impact on properties, and the temporal probability. During this process, we used a hybrid machine learning model (in this case certainty factor-genetic algorithm-support vector classification) to calculate debris flow susceptibility. This hybrid model integrated topographical, ecological, geological, and meteorological factors into model development. Apart from the debris-flow assessment, the economic losses sustained by residential buildings were analyzed, since this element comprises the majority of the total economic risk. In this case, an ensemble machine learning model, XGBoost, was employed to predict the final volume of future debris-flow events because it calculates impact pressure on residential buildings. To provide reliable physical vulnerabilities for buildings at risk, we analyzed the horizontal and vertical distance values of buildings and proposed a physical vulnerability matrix based on such analysis. Finally, the multiplication of physical vulnerability and unit price of properties led

to the evaluation of possible economic losses. To test the efficiency and feasibility of this method in estimating the economic risk, the Gyirong Basin in southwestern Tibet was selected as the study site. The results revealed that residential buildings account for 76.7% of the total economic risk, followed by farmland and forests. These calculated results align with the actual distribution of debris flows based on field investigations, demonstrating the method's accuracy and applicability. Our findings suggest that this method is suitable for the regional assessment of economic risks caused by future debris flows in mountainous areas. The proposed methodology can thus provide an effective guidance and scientific support for decision makers in the fields of risk understanding, resource prioritization, definition of mitigation and adaptation strategies, long-term planning, and emergency response, to prevent and mitigate future debris-flow risk in the Gyirong Zangbo Basin.

Acknowledgments This work was financially supported by the Key Laboratory of Mountain Hazards and Earth Surface Processes, Chinese Academy of Sciences; the European Union's Horizon 2020 research and innovation program Marie Skłodowska-Curie Actions Research and Innovation Staff Exchange (RISE) under grant agreement (Grant No. 778360); the National Natural Science Foundation of China (Grant No. 51978533); and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA20030301).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Angillieri, M.Y.E. 2020. Debris flow susceptibility mapping using frequency ratio and seed cells, in a portion of a mountain international route, Dry Central Andes of Argentina. *Catena* 189: Article 104504.
- Ardabili, S., M. Amir, and A.R. Várkonyi-Kóczy. 2020. Advances in machine learning modeling reviewing hybrid and ensemble methods. In *Engineering for sustainable future: Selected papers of the 18th International Conference on Global Research and Education Inter-Academia-2019*, ed. A.R. Várkonyi-Kóczy, 215–227. Cham: Springer.
- Bednarik, M., I. Yilmaz, and M. Marschalko. 2012. Landslide hazard and risk assessment: A case study from the Hlohovec-Sereď landslide area in south-west Slovakia. *Natural Hazards* 64(1): 547–575.
- Burbidge, R., and B. Buxton. 2001. An introduction to support vector machines for data mining. *Keynote papers, young OR12*: 3–15 <https://svms.org/tutorials/BurbidgeBuxton2001.pdf>. Accessed 10 Feb 2024.
- Chen, X.Z., H. Chen, Y. You, and J.F. Liu. 2015. Susceptibility assessment of debris flows using the analytic hierarchy process method—A case study in Subao River valley, China. *Journal of Rock Mechanics and Geotechnical Engineering* 7(4): 404–410.
- Corominas, J., C. van Westen, P. Frattini, L. Cascini, J.P. Malet, S. Fotopoulou, F. Catani, and M. Van Den Eeckhaut et al. 2014. Recommendations for the quantitative analysis of landslide risk. *Bulletin of Engineering Geology and the Environment* 73: 209–263.
- Crovelli, R.A., and J.A. Coe. 2008. *Probabilistic methodology for estimation of number and economic loss (cost) of future landslides in the San Francisco Bay Region, California*. Reston, VA: U.S. Geological Survey.
- Cui, P., L.Z. Xiang, and Q. Zou. 2013. Risk assessment of highways affected by debris flows in Wenchuan Earthquake area. *Journal of Mountain Science* 10: 173–189.
- Dai, F.C., C.F. Lee, and Y.Y. Ngai. 2002. Landslide risk assessment and management: An overview. *Engineering Geology* 64: 65–87.
- Dong, J.W., Y. Chen, B.Y. Yao, X. Zhang, and N.F. Zeng. 2022. A neural network boosting regression model based on XGBoost. *Applied Soft Computing* 125: Article 109067.
- Fu, S., L. Chen, T. Woldai, K.L. Yin, L. Gui, D.Y. Li, J. Du, and C. Zhou, et al. 2020. Landslide hazard probability and risk assessment at the community level: A case of western Hubei, China. *Natural Hazards and Earth System Sciences* 20(2): 581–601.
- Gartner, J.E., S.H. Cannon, P.M. Santi, and V.G. Dewolfe. 2008. Empirical models to predict the volumes of debris flows generated by recently burned basins in the western US. *Geomorphology* 96(3–4): 339–354.
- Guzzetti, F., M. Galli, P. Reichenbach, F. Ardizzone, and M. Cardinali. 2006. Landslide hazard assessment in the Collazzone area, Umbria, central Italy. *Natural Hazards and Earth System Sciences* 6(1): 115–131.
- Hardwick Jones, R., S. Westra, and A. Sharma. 2010. Observed relationships between extreme sub-daily precipitation, surface temperature, and relative humidity. *Geophysical Research Letters*. <https://doi.org/10.1029/2010GL045081>.
- Huang, J., T.C. Hales, R.Q. Huang, N.P. Ju, Q. Li, and Y. Huang. 2020. A hybrid machine-learning model to estimate potential debris-flow volumes. *Geomorphology* 367: Article 107333.
- Hungr, O., S.G. Evans, M.M. Bovis, and J.N. Hutchinson. 2001. A review of the classification of landslides of the flow type. *Environmental & Engineering GeoScience* 7(3): 221–238.
- Jakob, M., D. Stein, and M. Ulmi. 2012. Vulnerability of buildings to debris flow impact. *Natural Hazards* 60(2): 241–261.
- Kang, H.S., and Y.T. Kim. 2016. The physical vulnerability of different types of building structure to debris flow events. *Natural Hazards* 80(3): 1475–1493.
- Khosravi, K., Z.S. Khozani, and L. Mao. 2021. A comparison between advanced hybrid machine learning algorithms and empirical equations applied to abutment scour depth prediction. *Journal of Hydrology* 596: Article 126100.
- Koch, T. 1998. Testing various constitutive equations for debris flow modelling. *IAHS-AISH Publication* 248: 249–257.
- Laigle, D., and E. Bardou. 2022. Mass-movement types and processes: Flow-like mass movements, debris flows and earth flows. *Treatise on Geomorphology (2nd edn.)* 5: 61–84.
- Lee, D.H., E. Cheon, H.H. Lim, S.K. Choi, Y.T. Kim, and S.R. Lee. 2021. An artificial neural network model to predict debris-flow volumes caused by extreme rainfall in the central region of South Korea. *Engineering Geology* 281: Article 105979.
- Liu, K.F., H.C. Li, and Y.C. Hsu. 2009. Debris flow hazard assessment with numerical simulation. *Natural Hazards* 49(1): 137–161.

- Marcato, G., G. Bossi, F. Rivelli, and L. Borgatti. 2012. Debris flood hazard documentation and mitigation on the Tilcara alluvial fan (Quebrada de Humahuaca, Jujuy Province, North-West Argentina). *Natural Hazards and Earth System Sciences* 12(6): 1873–1882.
- Marchi, L., M.T. Brunetti, M. Cavalli, and S. Crema. 2019. Debris-flow volumes in northeastern Italy: Relationship with drainage area and size probability. *Earth Surface Processes and Landforms* 44(4): 933–943.
- Mohamed, A.E. 2017. Comparative study of four supervised machine learning techniques for classification. *International Journal of Applied Science and Technology* 7(2): 5–18.
- Mondal, S., and R. Maiti. 2013. Integrating the analytical hierarchy process (AHP) and the frequency ratio (FR) model in landslide susceptibility mapping of Shiv-khola watershed, Darjeeling Himalaya. *International Journal of Disaster Risk Science* 4(4): 200–212.
- Nguyen, H., T. Vu, T.P. Vo, and H.T. Thai. 2021. Efficient machine learning models for prediction of concrete strengths. *Construction and Building Materials* 266: Article 120950.
- Osman, A.I.A., A.N. Ahmed, M.F. Chow, Y.F. Huang, and A. El-Shafie. 2021. Extreme gradient boosting (XGBoost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Engineering Journal* 12(2): 1545–1556.
- Papathoma-Köhle, M., B. Gems, M. Sturm, and S. Fuchs. 2017. Matrices, curves and indicators: A review of approaches to assess physical vulnerability to debris flows. *Earth-Science Reviews* 171: 272–288.
- Pei, Y.Q., H.J. Qiu, D.D. Yang, Z.J. Liu, S.Y. Ma, J.Y. Li, M.M. Cao, and W. Wufuer. 2023. Increasing landslide activity in the Taxkorgan River Basin (eastern Pamirs Plateau, China) driven by climate change. *Catena* 223: Article 106911.
- Qiu, C.C., L.J. Su, Q. Zou, and X.Y. Geng. 2022. A hybrid machine-learning model to map glacier-related debris flow susceptibility along Gyirong Zangbo watershed under the changing climate. *Science of the Total Environment* 818: Article 151752.
- Quan Luna, B., J. Blahut, C. Camera, C. van Westen, T. Apuani, V. Jetten, and S. Sterlacchini. 2014. Physically based dynamic run-out modelling for quantitative debris flow risk assessment: A case study in Tresenda, northern Italy. *Environmental Earth Sciences* 72(3): 645–661.
- Remondo, J., J. Bonachea, and A. Cendrero. 2005. A statistical approach to landslide risk modelling at basin scale: From landslide susceptibility to quantitative risk assessment. *Landslides* 2(4): 321–328.
- Rickenmann, D. 1999. Empirical relationships for debris flows. *Natural Hazards* 19(1): 47–77.
- Schick, A.P., T. Grodek, and M.G. Wolman. 1999. Hydrologic processes and geomorphic constraints on urbanization of alluvial fan slopes. *Geomorphology* 31(1–4): 325–335.
- Shi, M.Y., J.P. Chen, D.Y. Sun, and C. Cao. 2015. Hazard assessment of debris flows based on the catastrophe progression method: A case study from the Wudongde Dam site. *International Journal of Heat and Technology* 33(1–4): 217–220.
- Staley, D.M., J.A. Negri, J.W. Kean, J.L. Laber, A.C. Tillery, and A.M. Youberg. 2017. Prediction of spatially explicit rainfall intensity-duration thresholds for post-fire debris-flow generation in the western United States. *Geomorphology* 278: 149–162.
- Stoffel, M., T. Mendlik, M. Schneuwly-Bollschweiler, and A. Gobiet. 2014. Possible impacts of climate change on debris-flow activity in the Swiss Alps. *Climatic Change* 122: 141–155.
- Sturm, M., B. Gems, F. Keller, B. Mazzorana, S. Fuchs, M. Papathoma-Köhle, and M. Aufleger. 2018. Understanding impact dynamics on buildings caused by fluvial sediment transport. *Geomorphology* 321: 45–59.
- Totschnig, R., and S. Fuchs. 2013. Mountain torrents: Quantifying vulnerability and assessing uncertainties. *Engineering Geology* 155: 31–44.
- Varnes, D.J. 1984. *Landslide hazard zonation: A review of principles and practice*. Washington, DC: The National Academies of Sciences, Engineering, and Medicine.
- Vranken, L., G. Vantilt, M. Van Den Eeckhaut, L. Vandekerckhove, and J. Poesen. 2015. Landslide risk assessment in a densely populated hilly area. *Landslides* 12(4): 787–798.
- Wang, R.R., L.P. Wang, J. Zhang, M. He, and J.G. Xu. 2022. XGBoost machine learning algorithm performed better than regression models in predicting mortality of moderate-to-severe traumatic brain injury. *World Neurosurgery* 163: e617–e622.
- Zanchetta, G., R. Sulpizio, M.T. Pareschi, F.M. Leoni, and R. Santacroce. 2004. Characteristics of May 5–6, 1998 volcanoclastic debris flows in the Sarno area (Campania, southern Italy): Relationships to structural damage and hazard zonation. *Journal of Volcanology and Geothermal Research* 133(1–4): 377–393.
- Zhang, M., Q. Ren, X. Wei, J. Wang, X. Yang, and Z. Jiang. 2011. Climate change, glacier melting and streamflow in the Niyang River Basin, southeast Tibet, China. *Ecohydrology* 4(2): 288–298.