Point-by-point response to the reviews

Dear Referee 1,

We would like to thank you for your professional and constructive comments concerning our manuscript entitled " Understanding rockfalls along the national road G318 in China: from source area identification to hazard probability simulation ". These comments are all valuable and helpful for revising and improving our manuscript. The main corrections in the manuscript and point-by -point responses are as following (the page number and line number in this refer to the revised manuscript)

Technical corrections:

General comment on the use of parentheses: Clearly a matter of writing style, however, IMHO the excessive use of parentheses hinders the reading flow. Personal guidance is: if it's important, rephrase it into the written sentences, if it does not merit being included in the text, remove it. The authors might check their use of parentheses with this in mind, or discard it as the referee's spleen. Does not hold for introduction of acronyms, of course.

Response: Thanks for your kind advice. In order to make it easier for readers to read, we have revised the extra parentheses in the paper.

Figure font sizes: Revise the font size and general sizing of heavily loaded figures.

Response: Thanks for your kind advice. We have modified the font size of all the figures.

Abstract:

l14: kinemics - kinematics, but is it really?

l20: results agree with measurements, fit well the acquired field data, etc., but they don't show fitness

l21: size scenarios usually are linked to recurrence periods. The de-coupling from size scenarios to recurrence period does not make sense.

Response: Thanks for your kind suggestion. In Line14, we modified the wrong words expressing kinemics to kinematics. The fitness is discussed from two aspects in this research: rockfall source area and trajectory probability. The fitness of rockfall source area is assessed in Table 4 and the comparison of different models is taken in Table 5.The fitness of trajectory is assessed in section 4.3. Fig.11 and Fig.12 shows the fitness by comparing historical and simulation results. It was indicated that the average difference value of run out is 1.97 m with an evaluation error ratio 3.66% (Line 372-373). Thank you for the suggestion that the fitness should be shown more detail in the abstract. The explanation about size scenario and return period scenarios can be found from the response "4.2 Temporal and size probability of rockfall sources"

1. Introduction

l26ff: what about debris flows, avalanches, shallow landslides, etc?

Response: Thanks for your kind suggestion. We revised the sentences as: Rockfall is one of the geological hazard…..

l29: at the border between China and Nepal. l30: crosses/leads through mountainous areas.

Response: Thanks for your kind suggestion. We revised the sentences as; …… and finally ending at the border between China and Nepal. More than 70 percent crosses through mountainous areas.

l33: book cited incorrectly, plus: does it really make sense to cite a book for common knowledge such as “rockfalls usually occur in montaineous regions”?

Response: Thanks for your kind suggestion. We have deleted the redundant expression in Line 34.

L38: derived from a digitital elevation model

Response: Thanks for your kind suggestion. We revised the sentences as: …in which slope gradient map derived from digital elevation model is used.

l41: This is not the only reason for LiDAR scanning and the Fanos et al. source is clearly focused on something different (machine learning for rockfall trajectory propagation modelling)

Response: Thanks for your kind suggestion. We deleted this sentence.

l42: The conclusion from the cited work is rather, that it is no unambiguous SAT derivation possible. That terrain is an important basis, is common knowledge. Not many rockfalls occur in the planes.

Response: Thanks for your kind suggestions. We corrected inaccurate statement as: SAT value is an important basis for Rockfall hazard Assessment.

L44: rockfall susceptibility is a combination of all of those factors. It should not be opposing, but complementary assessments.

Response: Thanks for your kind suggestion. We have revised the sentences as:

The morphology-based method is simple in data-limited areas. If data available or assessment scale is large, other conditioning factors such as discontinuities and joint sets in rocks need to be supplemented (Guzzetti et al., 1998; Jaboyedoff et al., 2003; Frattini et al., 2008; Heckmann et al., 2016).

L47: sentence makes no sense. Source areas can be identified more accurately either by using empirical, statistical or deterministic methods.

Response: Thanks for your kind suggestions. We have corrected the sentence as:

Source areas can be identified more accurately either by using empirical, statistical or deterministic methods.

l49: how widely used is RHRS? And how accurate/universal are the proposed exponential function within the original RHRS publication? It is a method amongst many.

Response: RHRS method has been used by many researchers, such as Brawner et al. (1975); Pierson (1993); Budetta (2004); Li et al. (2009); Corominas et al. (2013). The Rockfall Hazard Rating System (RHRS) is a stepwise process designed to identify potentially hazardous slopes by assigning a hazard rating. In Line 49, we did not mention exponential function within the original RHRS publication. If more explanation is needed, please don’t hesitate to tell the authors.

L51: Oommen at al. (1984) should be Bouali et al (2019)

Response: Thanks for your kind suggestions. We corrected the incorrect citations by changing Oommen at al. (1984) to Bouali et al. (2017).

Bouali, E. H., Oommen, T., Vitton, S., Escobar-Wolf, R., Brooks, C.: Rockfall Hazard Rating System: Benefits of Utilizing Remote Sensing. Environmental and Engineering Geoscience. 23 (3), 165–177. https://doi.org/10.2113/gseegeosci.23.3.165, 2017.

L54-56: arguing with academic references from roughly 30 years ago, that a method is commonly used is a bit far fetched. The problematic on input data is already discussed there.

Response: Thanks for your kind suggestion. We have revised the references to Lee, 2005; Benchelha et al., 2019. The sentence is revised as:

MLRM is used to construct slope instability susceptibility models (Chung et al., 1995; Lee, 2005; Benchelha et al., 2019).

Lee S. Application and Cross-Validation of Spatial Logistic Multiple Regression for Landslide Susceptibility Analysis. Geosciences Journal, 9(1):63-71. https://link.springer.com/content/pdf/10.1007/BF02910555.pdf, 2005.

Benchelha, S., Chennaoui Aoudjehane, H., Hakdaoui, M., El Hamdouni, R., Mansouri, H., Benchelha, T., Layelmam, M., Alaoui, M., Landslide susceptibility mapping in the Commune of Oudka, Taounate Province, north Morocco; A comparative analysis of logistic regression, multivariate adaptive regression spline, and artificial neural network models. Environ. Eng. Geosci. 26(2), 185-200. https://doi.org/10.2113/EEG-2243, 2019.

L62 ff: What is 3D collapse motion? The argument, that those models require extensive field investigation and experimental parameters as opposed to Flow-R is not substantial.

The reference Jabodeyoff et al.2003 is a link to where no manual for FLOW-R is found anymore (CONEFALL and others are found there). The statement, that FLOW-R produces more realistic results with the citations of a wrong manual is a bold – if not scientifically fraudulent - claim.

Response: Thanks for your kind suggestion. The expression of 3D Rockfall Motion is not accurate, so we removed it in Line 78. It means that the three-dimensional physical model is used to evaluate the risk of rockfall. The argument that those models require extensive field investigation and experimental parameters as opposed to Flow-R is deleted. The sentence is revised as:

Among them, Flow-R is developed for regional-scale on Matlab@2016, utilizing both empirical studies and physical modeling for gravitational hazards (Horton et al., 2013).

In Line 81, the sentence we expressed was not exact, so we deleted it to avoid ambiguity. The statementthat FLOW-R produces more realistic results with the citations of a wrong manual is deleted.

L 72: What is a fragment in this case? Usually fragments are fragmented parts from a initially released rock from the release area. Of course, those rocks are also fragments from the original rock wall etc, but in the literature, fragmentation means the breaking up of a single block during its trajectory. The influence on deposition patterns etc. is a hot topic and controversially debated. Reach angle analysis, however, can not contribute, to this discussion.

Response: Thanks for your kind suggestions. The expression of 'fragment' is not accurate. In order to avoid ambiguity, we revised this sentence as:

In trajectory path simulation, the minimum reach angle or shadow angle is a key parameter controlling the influence area of rockfall.

L 80ff: What is the temporal probability? Recurrence periods? There is a great many work around scenario building in rockfall etc. The authors have a point, that a thorough link between occurrence probability and scenario probability might be a weak point of current hazard mitigation literature. Please rephrase.

Response: Thanks for your kind suggestions. Temporal probability refers to the probability of n disasters occurring within a certain period T in a certain region. More explanation can also be found in the responses for you "4.2 Temporal and size probability of rockfall sources".

"

L85-90: Please refine the English.

Response: Thanks for your kind suggestions. We reworked the English expression to make it easier for readers to understand.we revised the sentences as:

For quantitative risk assessment of rock fall hazard, we consider that the rockfall assessment including source area identification and rock fall propagation should be at the level of probability assessment. To understand the potential risk from rock falls along national highway G318 in China, this study try to assess the potential hazard probability and risky road segments, considering a given rock fall volume over certain return period.

2. Study Area

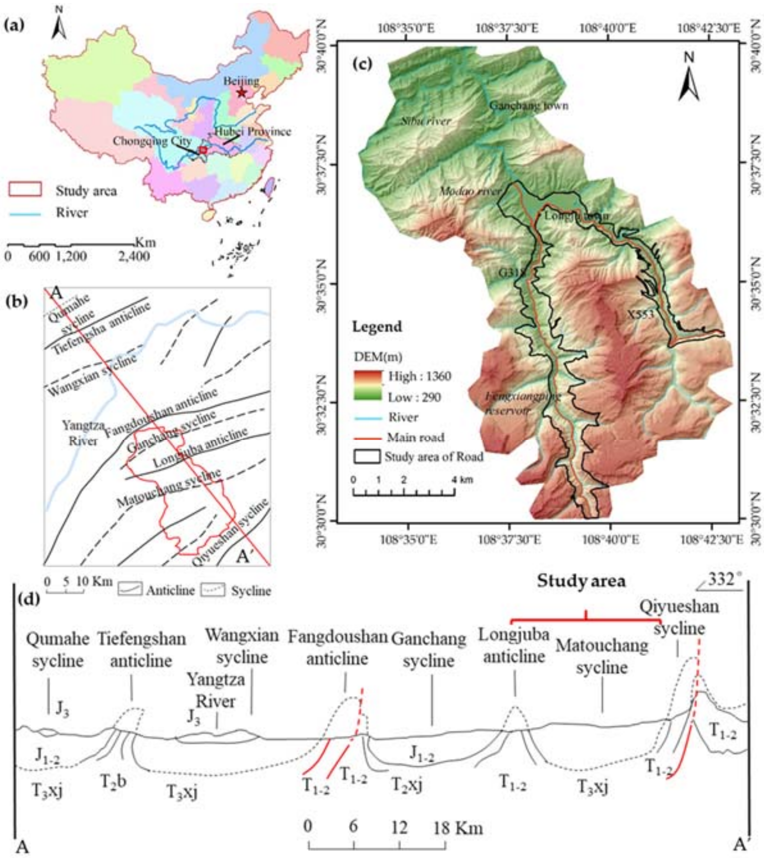
L93 Intence → Intense erosion and weathering processes

L94 600 m

Response: Thanks for your kind suggestion. We corrected the wrong words.

Figure 1: all anticlines in the figure are labelled incorrectly antivline

Response: Thanks for your kind suggestions. We have modified all the wrong words in the picture, as shown below:



L101 The lithology in the area consists mainly of purplish-red mudstone

l104 frastures   infrastructure. Are there any statistics on the events on this road and the caused dammage?

L105 Anticline   anticline

Response: Thanks for your kind suggestion. We have modified the wrong words. There are no statistical data of the events on this road but we have known in the field that there was a hazard damage to a pick-up car with 3 people inside six years ago (Fig.2d).

L107: how obvious? What do you want to tell the reader?

Response: Thanks for your kind suggestion. Sorry for the unclear and confusing statement. We have deleted the sentence.

L111: nucleus   core, near-wings?

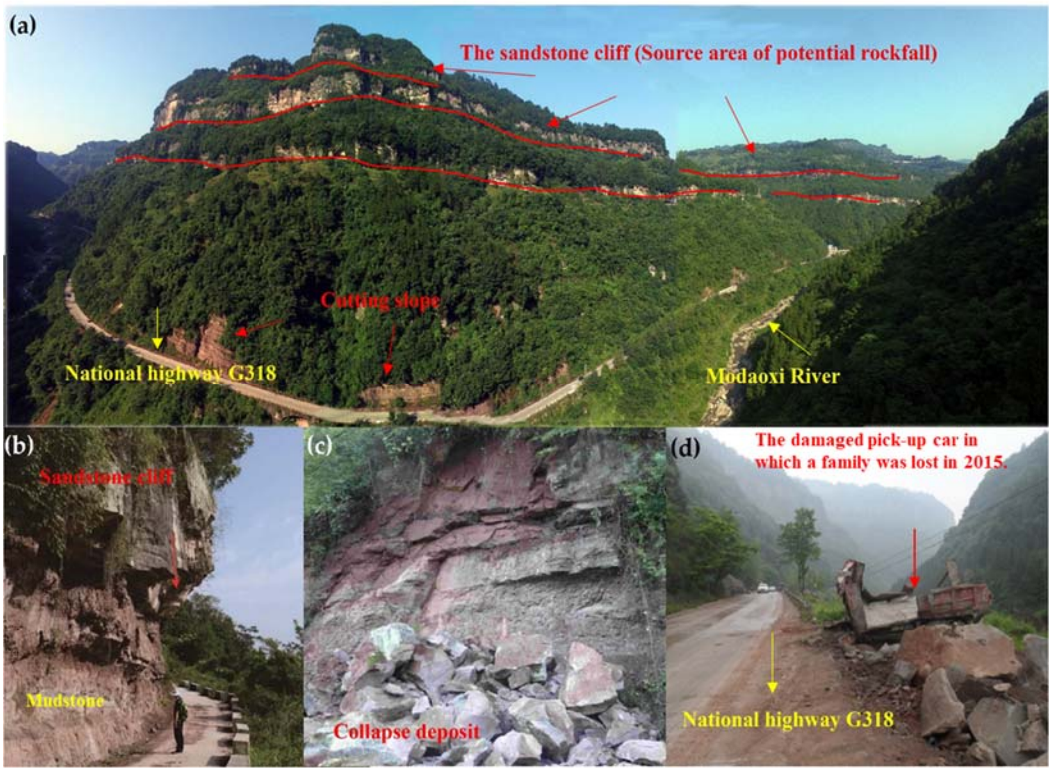
Response: Thanks for your kind suggestion. We modified the wrong word in Line 147.

L112: what is differential weathering?

Response: The effects of different weathering of lithology of cliffs or steep slopes are the cause for rockfall. Soft rocks like mudstone are easier or slower to be weathered than hard stones like sandstone. If the slope is composed by mudstone overlaid by sandstone, differential weathering will cause caves below cliffs and then rockfalls.

L117: Figure 2a shows no vehicle damage, that is Figure 2d. “The sandstone cliff collapsed” are the steep section without vegetation, not necessarily a collapse already.

Response: Thanks for your kind suggestion and sorry for the confusing labels in Figure 2. We have modified the figures and labels very carefully.



4. Methodology

Figure 3 – the presented methodology is a quite intricate interplay. A priori MLRM and RFM models work only for large data sets.

Response: Thanks for your kind suggestions. Our study area is about 21 km^2, with 108 rockfall source areas investigated in the study area. The data is sufficient for the use of MLRM and RFM models.

L144: SAT methodology according to Loye et al. shows quite a bit of DEM resolution dependency. The adaption of this procedure to a 10 m DEM is questionable.

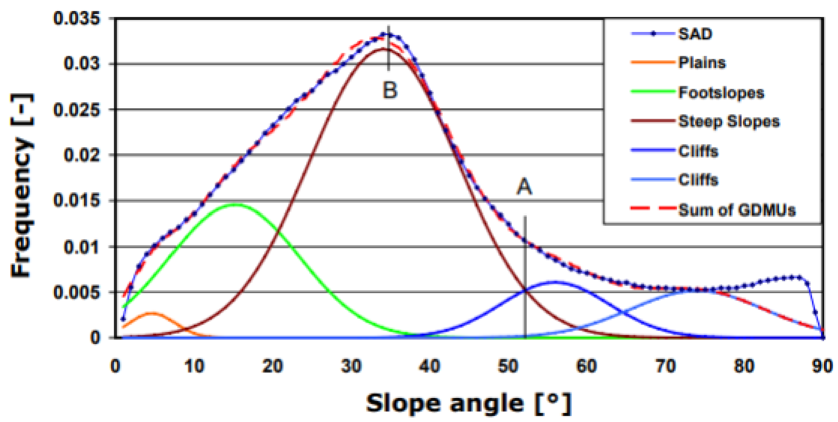
Response: Thanks for your kind suggestion. According to Figure 2 by Loye et al.(2009), there is a suggested grid size from the relationship between grid size, slope angle and slope height. As to our study area, the average slope angle is about 40 (Fig.5) and the major slope height over 5 m. The grid size 10 m is no questionable for this study.

L201: Reference should be Bak et al. (1988). Additionally, the reference deals with “Self-organized criticality”

Response: Thanks for your kind suggestion. We have corrected the reference.

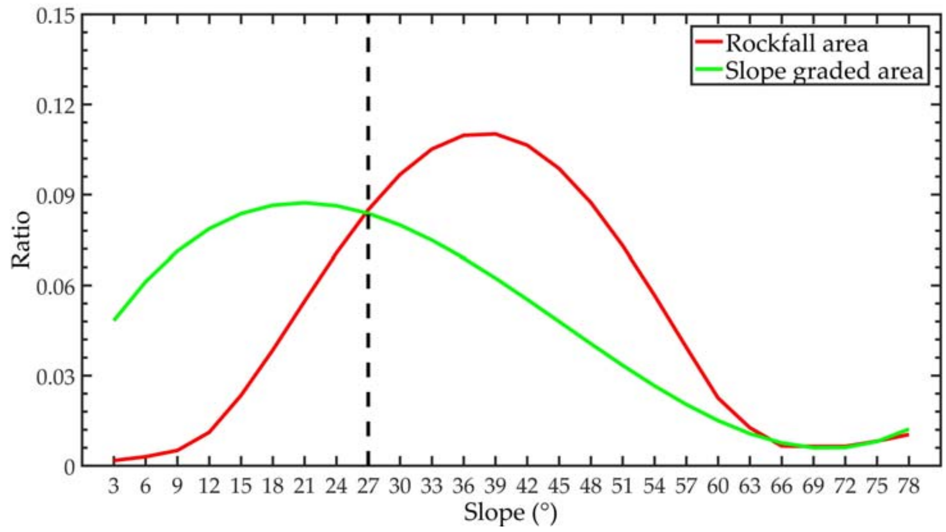
L203 DOI of source Pelletier et al. 1997 is invalid.

Response: Thanks for your kind suggestions. We have modified DOI in Pelletier et al. 1997.



Additionally, 27° is the transitions between footslopes and steep slopes. It is a rather low value in general as threshold.

Response: Thanks for your kind suggestions. The SAT we used is different from the method of considering only topographic slope data in reference Loye et al. (2009). We not only considered the topographic slope in the study area, but also historical disaster rockfall data in the study area. The relationship between slope angle and historical rockfalls in the study area is shown in the figure below. It can be seen from the green line in the following figure that the regional distribution proportion of rockfall above 27° is higher than that in Loye et al. (2009). So, 27° is a suitable threshold in our study area where the terrain is steep.



l257: In general, the Varnes et al. (1984) citation is very old, hard to retrieve and in the context of MDLP highly likely the wrong citation.

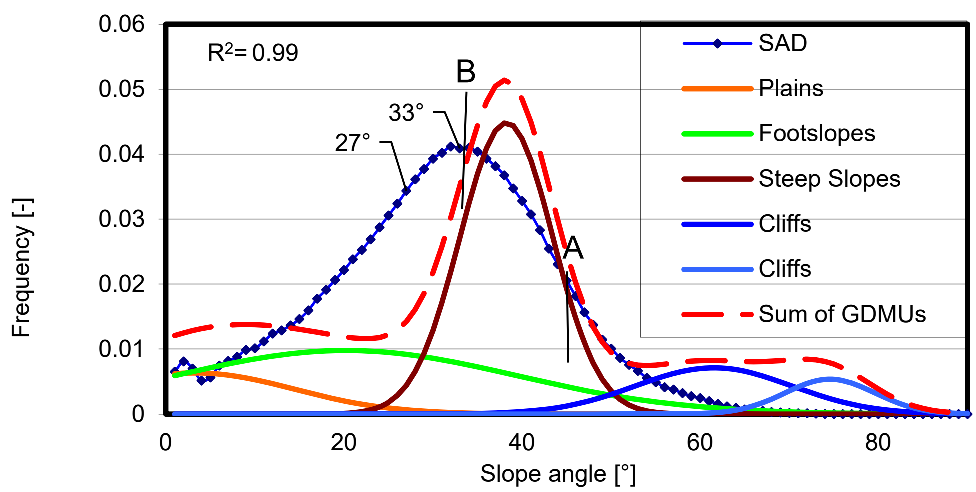
Response: Thanks for your kind suggestion. We re-checked the literature and modified the references, changing Varnes et al. (1984) to Rissanen (1978) and Vitanyi (2000).

Rissanen, J. J.: Modeling by the shortest data description, Automatica-J.IF AC, 14, 465–471, https://doi.org/10.1016/0005-1098(78)90005-5, 1978.

Vitanyi, P. M. B., Li, M.: Minimum description length induction, Bayesianism, and Kolmogorov complexity, IEEE Transactions on Information Theory, 46, no. 2, 446-464, https://doi.org/10.1109/18.825807 , 2000.

L295: Comparison with SAT model approach is not valid, as SAT model approach is done incorrectly.

Response: Thanks for your kind suggestion. According to the method by Loye et al. (2009), we made the Gaussian distribution graph of terrain slope in the study area, as shown below. It can be seen that 33° should be selected according to the way of considering only terrain. However, according to the statistical historical rockfall data, only 57.85% of the rockfall disaster is located above 33°, while 84.79% of the rockfall disaster is located above 27°. If 27° is selected, a large amount of historical data is retained, which provides a guarantee for the accurate prediction of the rockfall source area later.



4.2 Temporal and size probability of rockfall sources

Response: Thanks for your kind suggestions. Combining temporal, spatial and size probability to assess hazard probability is proved to be a scientific way according to the published papers, such as:

Wu, C.Y., Chen, S.C.: Integrating spatial, temporal, and size probabilities for the annual landslide hazard maps in the Shihmen watershed, Taiwan. Nat. Hazard. Earth Sys. 13(9), 2353-2367. https://doi.org/10.5194/nhess-13-2353-2013, 2013.

Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F., Cardinali, M.: Landslide hazard assessment in the Collazzone area, Umbria, central Italy. Nat. Hazard. Earth Sys. 6(1), 115-131. https://doi.org/1684-9981/nhess/2006-6-115, 2006.

Catani, F., Casagli, N., Ermini, L., Righini, G., Menduni, G.: Landslide hazard and risk mapping at catchment scale in the Arno River basin. Landslides. 2(4), 329-342. https://doi.org/10.1007/s10346-005-0021-0, 2005.

Brunetti, M.T., Guzzetti, F., Rossi, M.: Probability distributions of landslide volumes. Nonlinear Proc. Geoph. 16(2), 179-188. https://doi.org/10.5194/npg-16-179-2009, 2009.

Liu, B., Siu, Y.L., Mitchell, G., Xu, W.: Exceedance probability of multiple natural hazards: Risk assessment in China's Yangtze River Delta. Nat. Hazards. 69(3), 2039-2055. https://doi.org/10.1007/s11069-013-0794-8, 2013.

Melchiorre, C., Frattini, P.: Modelling probability of rainfall-induced shallow landslides in a changing climate, Otta, Central Norway. Climatic Change. 113(2), 413-436. https://doi.org/10.1007/s10584-011-0325-0, 2012.

Fu, S., Chen, L., Woldai, T., Yin, K., Gui, L., Li, D., Du, J., Zhou, C., Xu, Y., Lian, Z.: Landslide hazard probability and risk assessment at the community level: A case of western Hubei, China. Nat. Hazard. Earth Sys. 20(2), 581-601. https://doi.org/10.5194/nhess-20-581-2020, 2020.

In this research, we aims to find the quantitative probability in terms of rockfall occurrence frequency, travel distance and size, which relate to quantitative risk assessment of elements at risk. Also it is requested by local government for future management in the study area. We collected sufficient historical rockfall hazard and geological environmental data in the field. Poisson model and exponential equation are then applied to obtain temporal and size probability. Considering the risk management request, 5, 20 and 50 years return periods are requested scenarios for hazard prevention budget plan of local government. It is why we use these three return periods to assess the potential risk.

5. Discussion

l408ff: Are you altering Flow-R in order to incorporate all the promised things

Response: Yes, We have made the plan to improve the algorithm in Flow-R for better understanding the rock fall risk in our study area.

Dear Referee 2,

Thank you very much for your professional comments on our manuscript. These comments are all valuable and helpful for revising and improving our manuscript. The main corrections in the manuscript and the point-by-point responses to your comments are as following (the page number and line number in this letter refer to the revised manuscript):

This paper is the application of several existing methodologies to a given case study. I don't recommend publication of this paper for the following reasons: 1) it does not fit the requirements for the scientific paper since I did not identified scientific novelty, 2) the relevance of the complete approach is questionable.

Response: Thank you for your comments. 1) This case study is taken based on the previous studies. Meanwhile, we created an improved way to minimize the uncertainty of source area susceptibility both considering slope angle and important controlling factors of rock falls. It is proved by the study that the potential source area grids reduced from 160,823 to 4,002 with only 1.4 percent loss of historical rock fall samples. The simulation efficiency increased about 40 times, which highly reduced the burden of trajectory simulation. 2) Detailed field investigation was taken by the authors to understand the mechanism of rock fall and related risk along the road. Also, the methodology is based on previous public approaches and applied for the study area where rock fall hazard is an important risk source. The local government needs quantitative risk assessment result to guide their management work. If detailed explanation is needed, please refer to our response to the Referee 1.

Dear Referee 3,

Thank you very much for your professional comments on our manuscript. These comments are all valuable and helpful for revising and improving our manuscript. The main corrections in the manuscript and the point-by-point responses to your comments are as following (the page number and line number in this letter refer to the revised manuscript):

This paper presents a concoction of standard methods in rockfall hazard analysis and does not contain a substantial contribution to science. The reviewer can concur with the recommendations of RC2 and the assessment of RC1.

Response: Thank you for your comments. Please refer to our response to the Referee 1 and 2. If more detailed explanation is needed, please don't hesitate to tell the authors.

We tried our best to improve the manuscript. We feel great thanks for your professional review work on our manuscript, and hope that the correction and response will meet with approval.

Sincerely,

Lixia Chen

Understanding rockfalls along the national road G318 in China: from source area identification to hazard probability simulation

Lixia Chen[1*], Yu Zhao[1], Yuanyao Li[2], Lei Gui[3], Kunlong Yin[3], Dhruba Pikha Shrestha[4]

[1] Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan, 430074, China; lixiachen@cug.edu.cn; cugzhaoyu@cug.edu.cn

[2] Institute of Geological Survey, China University of Geosciences, Wuhan, 430074, China; liyuanyao@cug.edu.cn

[3] Faculty of Engineering, China University of Geosciences, Wuhan, 430074, China; lei.gui@cug.edu.cn; yinkl@cug.edu.cn

[4] Department of Earth Systems Analysis, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, 7500 AE Enschede, the Netherlands; d.b.p.shrestha@utwente.nl

* *Correspondence to:* lixiachen (lixiachen@cug.edu.cn);

Abstract: Rockfall hazard is frequent along the national road (G318) in west Hubei, China. To understand the distribution and potential hazard prone to road G318, this study combines the result of a 3-year engineering geological investigation, statistical modeling, and kinematics-based method to identify risky road sections. Rockfall source area cells are preliminarily identified by slope angle threshold analysis and then improved by Random Forest model and Multivariate Logistic Regression model, considering rockfall controlling factors. Temporal and size probabilities of source areas are separately calculated by Poisson distribution and power-law distribution theory. To get the reaching probabilities and potential influence area of released sources, rockfall trajectory simulation was taken by Flow-R tool. In this process, reaching angle was determined by back analysis and then validated by field investigation data. Rockfall hazard probability is finally calculated by integrating spatial, temporal, size probability, and reaching probabilities of rockfall sources. The results show that the potential source area grids reduced from 160,823 to 4,002 with only 1.4 percent loss of historical rockfall samples. The simulation efficiency increased about 40 times, which highly reduced the burden of trajectory simulation. Rockfall trajectory simulation results fit well with the field measurements, with about 96% accuracy. For the scenario of 5, 20, and 50 years return period, potential risky road sections are found out under two size scenarios (larger than 1 000 m^3, 10 000 m^3). This research helps the local government to understand the rockfalls from source area existence and potential risk to roads.

Keywords: Rockfalls; Slope angle threshold; Random Forest model; Multivariate logistic regression model; Flow-R; Reach angle

1. Introduction

Rockfall is one of geological hazard along roads in steep mountainous areas such as in the Himalayas, the Alps, in the rocky mountains, in the Andes, etc. Also in China, rockfall is a common problem in mountainous areas. The national highway G318 is the longest motorway (approx. 5476 km) in China, starting from Shanghai and passing through major cities such as Wuhan, Chongqing, Lhasa and finally ending in Kodari, at the border between China and Nepal. More than 70 percent crosses through mountainous areas. Because of the special geomorphological and geological set up, the road section(approx. 1302 km) in Hubei-Chongqing has been exposed to frequent slope failures causing property damages and disruptions of traffic. In 2016, a family in a pick-up van was lost because of a small volume but sudden rockfall along the G318. Such kind of small size but high frequency and intensity (e.g.velocity, energy) rockfalls are common in China, which can lead to human casualties and property loss. To protect the people commuting on the roads, we have to understand where the rockfall source area is and its hazard level. Once we know this, then suitable mitigation measures can be implemented.

In terms of source area identification, a large number of research results are available. Common methods for identifying the source areas can be divided into two main types: geomorphic and geological. The geomorphic approach uses the slope angle threshold (SAT) method to identify rockfall source area in which slope gradient map derived from a digital elevation model (DEM) (Jaboyedoff et al., 2003; Loye et al., 2009; Žabota et al., 2019; Liu et al., 2020). The existing research results show that the critical SAT values vary from rockfall types and study areas (e.g. $>60^\circ$ in Wieczorek et al.,1998, and Guzzetti et al.,2003; $>45^\circ$ in Jaboyedoff and Labiouse, 2003; $>37^\circ$ in Frattini et al., 2008; $>48^\circ$in Matasci, Jaboyedoff et al., 2015). So, SAT value is an important basis for rockfall hazard assessment. The morphology-based method is simple in data-limited areas. If data available or assessment scale is large, other conditioning factors such as discontinuities and joint sets in rocks need to be supplemented (Guzzetti et al., 1999; Jaboyedoff et al., 2003; Frattini et al., 2008; Heckmann et al., 2016).

Source areas can be identified more accurately either by using empirical, statistical or deterministic methods. An empirical expert evaluating system has been developed to access hazard susceptibility or probability, such as Rockfall Hazard Rating System (RHRS). It is a widely used method to identify the riskiest slopes on highways or coastal roads by many researchers, such as Brawner et al. (1975); Pierson (1993); Budetta (2004); Li et al. (2009); Corominas et al. (2013). The traditional RHRS approach is field-based: observations are made by a field crew who convert observations into slope ratings (preliminary and detailed) (Bouali et al. 2017). But it takes a lot of manpower and resources. The system is gradually optimized by using optical remote sensing data from satellites or unmanned aerial vehicles (Bouali et al. 2017). Statistical methods, such as the Random Forest Model (RFM) and Multivariate Logistic Regression Model (MLRM) are applied in Geographic Information System, especially for large or small scale areas. RFM, a machine learning algorithm based on the concept of classification trees, is able to classify landslide hazard susceptibility (Chen et al., 2014; Messenzehl et al., 2017). MLRM is used to construct slope instability susceptibility models (Chung et al., 1995; Lee, 2005; Benchelha et al., 2019). RFM can achieve higher accuracy with the same data. However, different models result in different source area locations. Thus, it is important to know which model performs better in the area of interest.

Besides rockfall source area, we also need to know rock mass trajectory paths with resulting intensity (e.g. velocity or kinetic energy) and the area it can affect. To simulate the trajectories and energy, several 2D or 3D tools or software were developed for regional-scale or site-specific rock slopes, such as CADMA by Azzoni et al. (1995); CONEFALL by Jaboyedoff et al. (2003), Flow-R by Horton et al. (2013), STONE by Guzzetti et al., (2002), RAMMS by Leine et al. (2014), DDA by Zheng et al. (2014), Rockyfor3D (Dorren L.K.A., 2016). Among them, Flow-R is developed for regional-scale on Matlab@2016, utilizing both empirical studies and physical modeling for gravitational hazards (Horton et al., 2013). It is now widely applied and has achieved good results in different countries, for example, Michoud et al. (2012) simulated the road collapse disaster in the Swiss Alps; Blahut (2010) simulated the area affected by debris flows in Tirano, Italy. Michoud, Derron et al. (2012) reported that Flow-R software provided helpful results of rock block propagations for hazard mapping and risk assessment at a regional scale in the Swiss Alps. Losasso, Dorren et al. (2016) used Flow-R to evaluate rockfall propagation extent and run-out distance in the Basilicata region, southern Italy. Flow-R can also be used to simulate other natural disasters, such as avalanches, debris flows, and floods (Horton et al., 2013).

In trajectory path simulation, the minimum reach angle is a key parameter controlling the influence area of rockfall. Reach angle is suggested by Shreve in 1968. Many researchers have used it to assess the propagation of rockfallsby making statistics on the relationship between reach angle and rock size (Losasso et al., 2017; Kanari et al., 2019; Marchelli et al., 2019; Mitchell et al., 2020). Kanari (2019) and Marchelli (2019) used the relationship between rock fall size and slope to analyze the collapsing movement and rock fall fragmentation. However, reach angle and related trajectory path modeling are not always calibrated probably due to data unavailability or budget constraints.

For quantitative risk assessment of rockfall hazard, we consider that the rockfall assessment including source area identification and rockfall propagation should be at the level of probability assessment. To understand the potential risk from rockfalls along national highway G318 in China, this study try to assess the potential hazard probability and risky road segments, considering a given rockfall volume over certain return period.

2. Study area

The research section of the national highway G318 is located in west Hubei, about 310 km Northeast of Chongqing, China (Figure 1). It covers about 21.19 km^2. Intense erosion and weathering processes created a cliffy topography in the southern part with elevation ranging from 600 m to 2000 m above sea level. Geological units in the study area (Figure1b, 1d) are mainly developed in the Middle Jurassic stratum, except Triassic limestone partly covering at the anticline. Due to the wide syncline, the stratum is mainly horizontal or gently deep in the area.

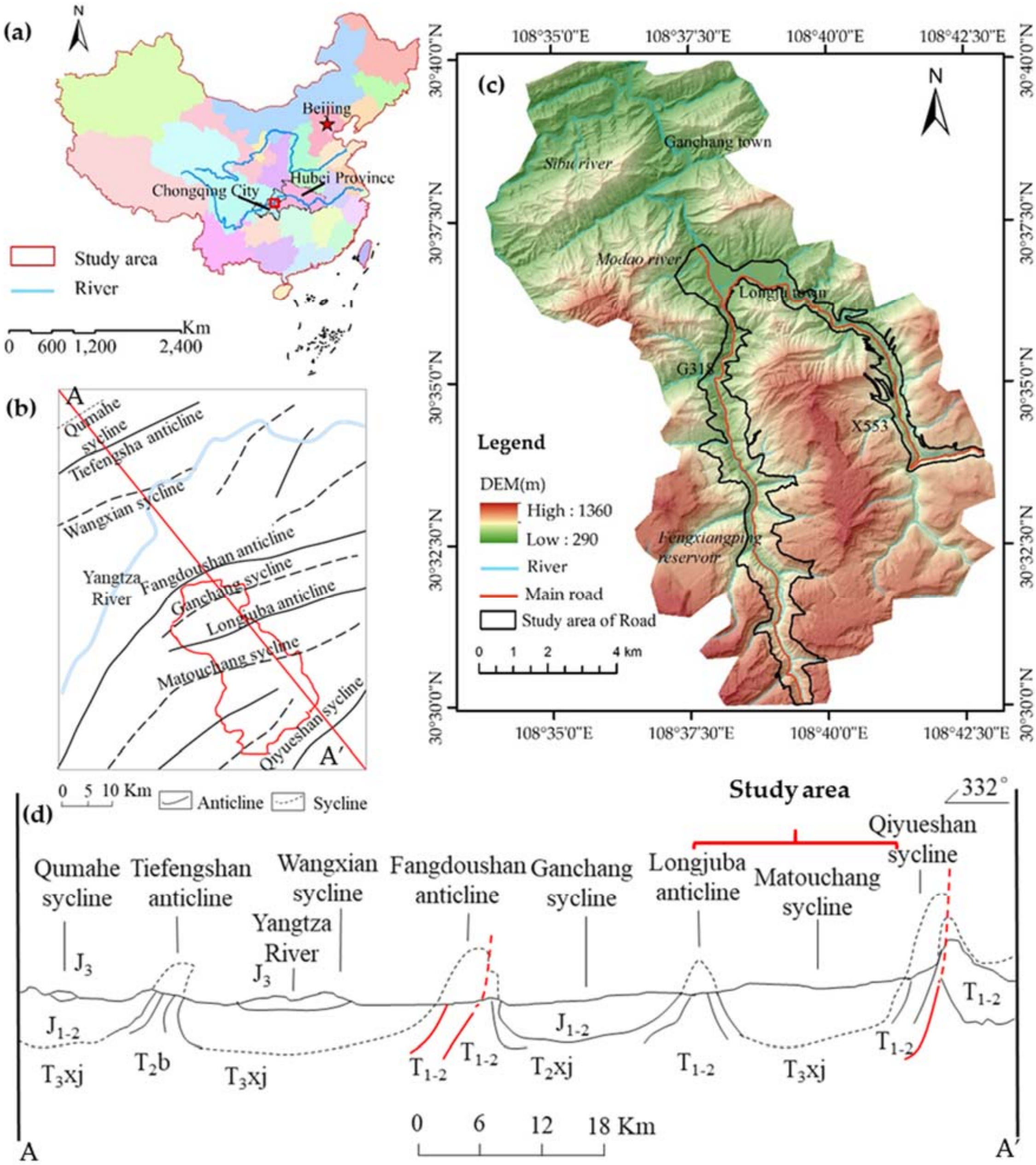


Figure1. (a) Location of the study area in the Qiyaoshan mountain ranges, China, (b) map showing the geological structure of the study area; (c) Hillshade map of the study area; (d) geological structure profile of the cross-section AA'.

The lithology in the area consists mainly of purplish-red mudstone, with sandstone and shell stone as interlayer, which has been affected by physical weathering so that most rockfalls have occurred at these sections. National (G318) and provincial (X553) roads are the main traffic ways, with a shape as an inverse Y across the area. Rockfalls occur frequently in the rainy season causing damage to infrastructure as well as human casualties.

The north of Longju town is located in the anticline of Fangdoushan and Jianchang syncline. The central part is the anticline of Longju town and the syncline of Matouchang. Jiannan anticline and Jianzhuxi syncline are located in the south. The rock strata in the core part of jianchang syncline are compressed and lithology is dense. The Matouchang syncline is narrow and steep in the northwest and broad and gentle in the southeast, so it is near the horizontal strata in the study area.

Rockfall is the main type of geological hazard in the area, especially in the Jurassic red bed (Middle Jurassic lithology) at the axial zone core or near-wings of the Matouchang syncline. Two sets of discontinuities control the rock quality and stability, combining with the stratum layer face. Due to these controlling rock structures, structural plane in sandstone and silty stone increases the probability of rockfalls.

In the recent 10 years, urbanization of the Longjuba area in the Three Gorges dam area has been promoted by the government. Accordingly, various construction works and reconstruction of transportation facilities have increased. In addition, due to the construction of a new highway in the area, which involved cutting and filling the slopes, the Longjuba area is becoming more and more hazardous, especially along the G318 (Fig.2a), it can be seen that the highway collapse causes vehicle damage (Fig.2d). According to the records of collapse disaster database in the study area, the historical time of collapse in the study area is from 1984 to 2015.

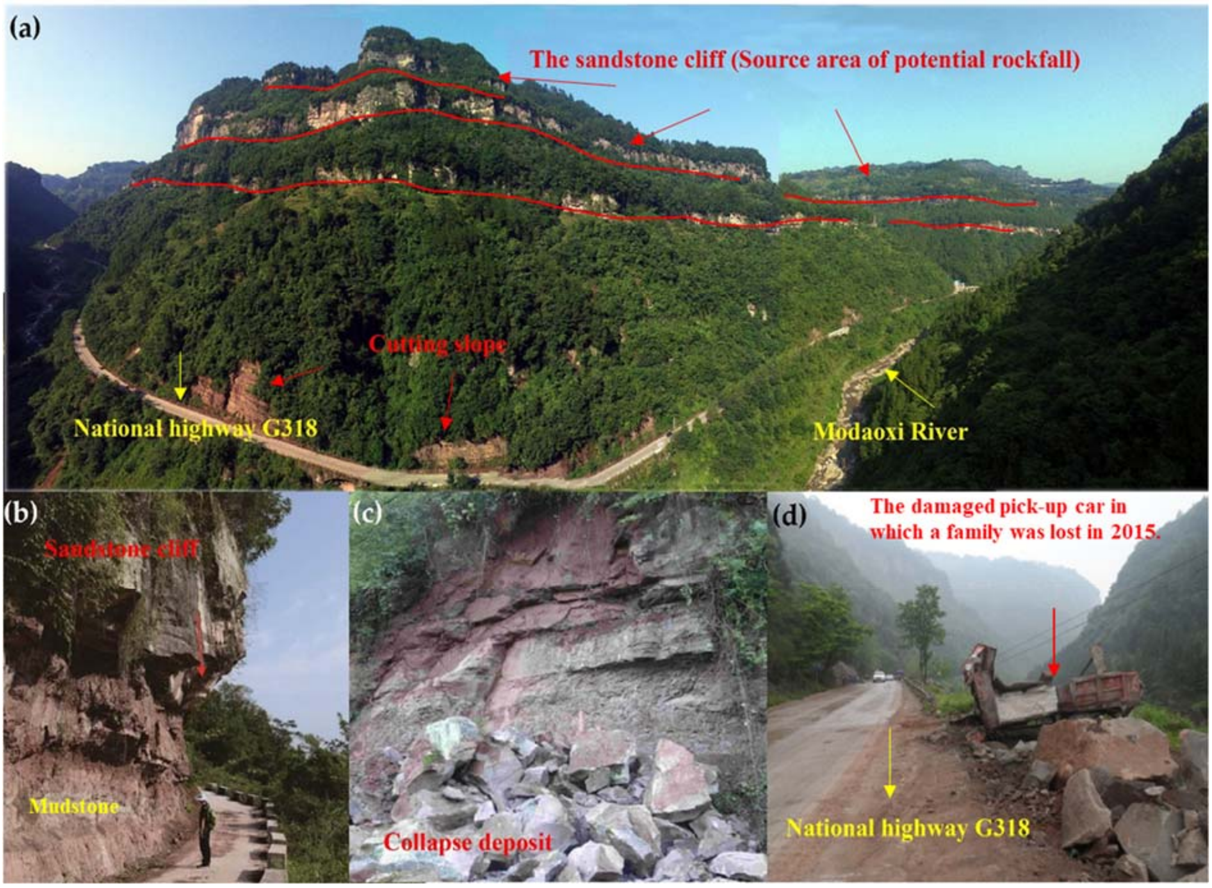


Figure 2. Geological setting and natural hazards in the study area: (a) The cliff inter-bedding of sandstone and mudstone along national highway G318 (UAVs image acquired in July 2016); (b) Sandstone cliff overlying mudstone along national highway G318; (c) Falling down of the rock sources on national highway G318; (d) Damaged pick-up car and rockfall fragmentation.

3. Methodology

Rockfall hazard probability assessment was carried out following the flow chart shown in Figure 3. Firstly, the study area was screened to extract the source area of collapse by using slope and topographic factors. On the basis of slope factors' analysis, Slope

SAT analysis is used to reduce the screening range of rockfall source areas. The most important inducing factors for collapse development were extracted from the obtained basic data to identify the source area of collapse. Multivariate logistic regression model and the random forest model were used and their results were compared by using the value of Area Under the ROC Curve (AUC) to predict and identify the collapse source area. The spatial probability for the source area determination is simulated using Flow-R. The temporal and size probabilities were assessed using a historical rockfall distribution pattern. The methodology is described in detail as follows:

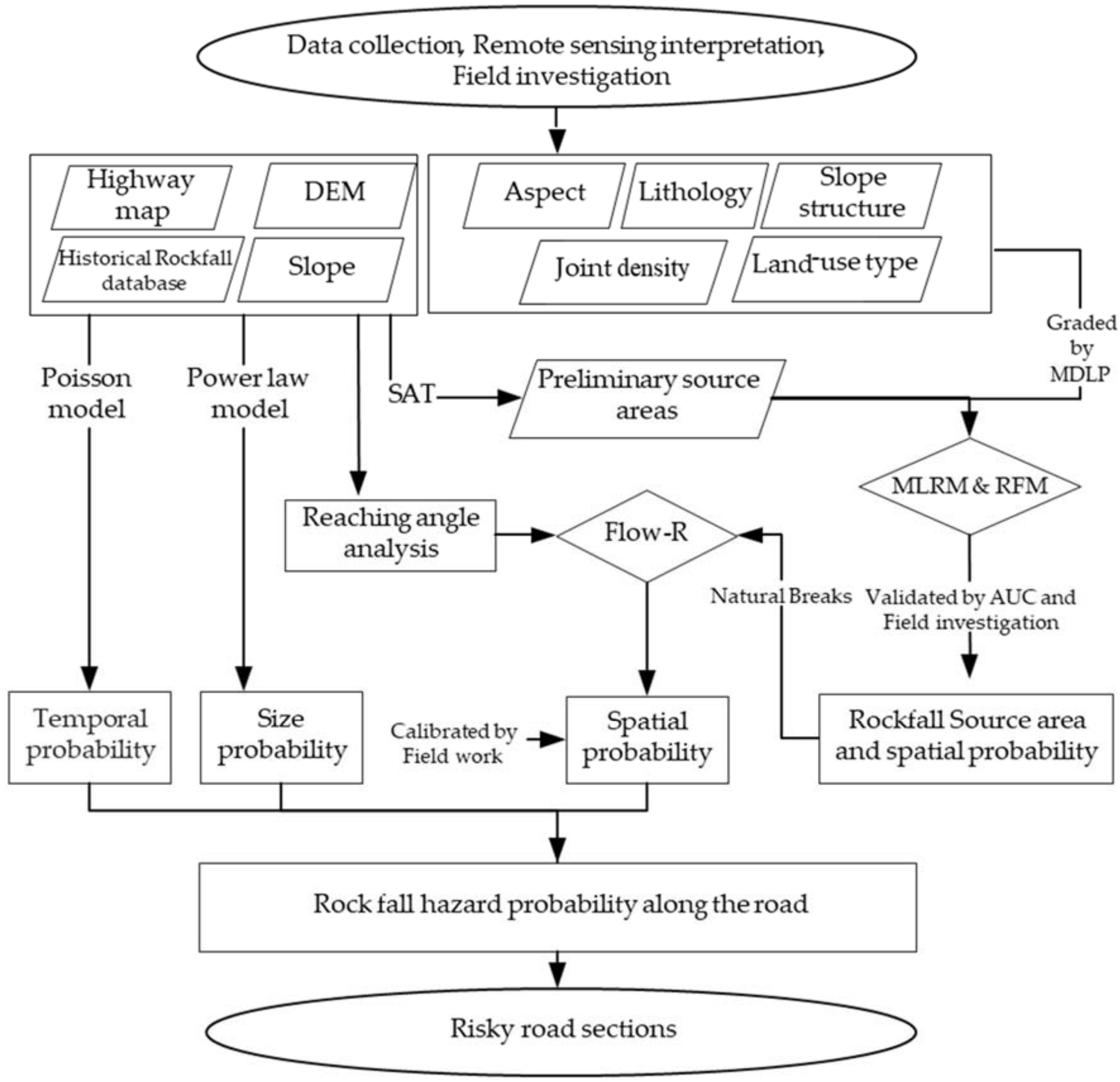


Figure 3. Workflow of rockfall hazard probability assessment and risky road section identification

3.1 Data collection

The historical rockfall inventory map was generated by data collection (data from Wuhan Geological Survey Center, China Geological Survey Bureau), three years field investigation (2014-2016), and remote sensing image (Gaofen-1 data) interpretation. Among them, 20 were interpreted by remote sensing. In total 108 rockfall locations were identified, covering 31 years from 1984 to 2015. Among them, 31 rockfalls have precise volume data, which was later used for source area identification, temporal probability, and size probability analysis. Especially, the transportation characteristics (e.g. run out) are available for 37 rockfalls, which were used for calibrating rockfall reaching probability simulation. There is no record of repeated disasters at all historical collapse sites.

Besides the rockfall inventory data, other datasets were collected as follows:

● A 10m resolution DEM was generated from GaoFen-1 remote sensing data (resolution: 1m, image time: 2015.03.30), from which slope, elevation, aspect, roughness, curvature, and solar radiation were generated using ArcGIS.

● A Geological map (1:10 000) was used to extract geological spatial layers such as lithology, faults, and slope structure map. The slope structure map was generated using the standard and stratigraphic altitude advocated by Cruden (1991).

● The joint density data was gathered in the field in 2015. Joint sets were measured at 108 rockfall source areas.

● The land-use map was generated from the GaoFen-1 remote sensing data by applying the Spectral Angle Mapper Classification method in ENVI software.

Specific data used are shown in Table 1.

Table 1. Source of data

data	sources	resolution
DEM	Gaofen-1 data (1m, image acquisition: 2015.03.30)	10 m
Geological structure	Geological map	1:50 000
Lithology	Geological map	1:50 000
Joint density data	Field survey	1:10 000
Rockfall inventory data	Historical Data Collection, Field survey data and Gaofen-1 data	/
Remote sensing image	Gaofen-1 data	1 m

3.2 Spatial probability of rockfall sources

Rockfall sources are preconditions of rockfall hazards and risks. We need to determine the potential rocky slopes which have the possibility to be unstable. In this study, three steps are recommended.

Calculate the preliminary rockfall source area

Firstly, we need to select the preliminary rockfall areas. In order to make a fine quantitative analysis of the collapse source area, we need to digitize and resample the study area. According to the scope of the study area and the scale of the collapse, the size of the grid is determined comprehensively. The preliminary source identification area of the collapse is constrained by the SAT method (Loye et al., 2009) in sequence. The SAT method can separate and remove the rockfall traveling area and accumulation area. SAT method determines the slope threshold based on the relationship between the number of historical collapses and the slope. The units with slopes steeper than the threshold are identified as preliminary rockfall source areas.

Secondly, rockfall conditioning factors in preliminary source areas are extracted and processed. The formation of collapse is controlled by topography, physical and chemical weathering, human engineering disturbance, and other factors. Therefore, we selected some factors that have the most serious impact on rock collapse in the study area. In addition to slope degree, the determination of fine compounds source area is also constrained by slope aspect, elevation, lithology, slope structure (spatial position between formation occurrence and slope face), joint density, land-use type, etc. Among these factors, slope, aspect, elevation, joint density, and distance to roads are continuity factors. We use the minimal description length principle (MDLP) to classify these continuity factors to improve the model prediction ability. MDLP is a method of discretizing continuous attributes, which has less manual intervention and better quantitative effect (Rissanen, 1978) than the methods such as equal frequency, equal width, and artificial definition.

Calculate the initial rockfall source area

Then, the susceptibility of the preliminary rockfall source area is assessed and compared by MLRM and RFM. In MLRM, the dependent variable is a dichotomic variable, with an absence-presence value of a certain characteristic. In this study, this variable is historical rockfalls. The RFM is a mining method based on statistical learning theory. It uses the idea of bagging to select a number of training samples and the establishment of a decision tree. The output category is obtained by various categories of the voting output tree. The main advantages of RFM are random sampling and features, avoiding overfitting, and improving the accuracy and stability of the model. RFM has achieved good results in the field of early warning of geological disasters (Chen et al., 2014; Provost et al., 2017). To reflect the importance of each variable, the Mean Decrease Gini (MDG) index was used. The higher the MDG index is, the more important the predictor (Liaw et al., 2012). Before model prediction, rockfall source area and non-rockfall source area samples are prepared. Rockfall source areas are identified as historical hazards. Non-rockfall source areas are randomly selected at least 500 meters away from rockfall source areas. We use 70% data of each group to generate a training dataset for model building and the remaining 30% for model testing. Using these samples and the conditioning factors, rockfall susceptibility is modeled by MLRM and RFM. The performance of the two models was evaluated.

Obtain the final rockfall source area

Finally, we classify the susceptibility value into five levels (very low, low, moderate, high, and very high) by the Natural Breaks method. In this method, breaks are classified as large as possible between groups and as small as possible within groups. The units with the highest class on the susceptibility map by the model with better performance are further finalized as rock fall source areas.

3.3 Temporal probability of rockfall sources

The temporal probability of rockfalls is evaluated by assuming that rockfalls are independent random events in the time domain (Crovelli et al., 2000; Fu et al., 2019). In this study, the Poisson model is adopted for constructing temporal probability. It is the exceedance probability of rockfall occurrence during a given period as follows:

$$P_t = 1 - e^{-t/RI}, RI = T / N \quad (1)$$

Where t is the return period, e.g., 5, 20, and 50 years; the recurrence interval (RI) is the historical mean recurrence interval for each rockfall source unit; T is the temporal interval of the rockfall database; N is the number of historical rockfalls recorded in each unit. Considering the possibility of missing rockfall points in the database, the units without historical records but having the highest class of spatial probability in the source area susceptibility map are set as historical rockfall units.

3.4 Size probability of rockfall sources

Rockfall size probability is calculated by analyzing the relationship between rockfall volume and cumulative frequency. Bak et al. (1988) proposed that there is a certain power index relationship between rockfall volume and its frequency, which has been verified in many regions (Pelletier et al., 1997; Malamud et al., 2004). This study follows the formula proposed by Malamud (2004) to fit the size probability.

$$P\left(V;\rho,a,s\right)=\frac{1}{a\Gamma\left(\rho\right)}\left[\frac{a}{V-s}\right]^{\rho+1} * \exp\left(-\frac{a}{V-s}\right) \quad (2)$$

Where P_v is size probability; V is rockfall volume; ρ is parameter primarily controlling power-law decay for medium and large values in three-parameter inverse-gamma probability distribution; a is parameter primarily controlling location of maximum probability in three-parameter inverse-gamma probability distribution; s is parameter primarily controlling exponential rollover for small values in three-parameter inverse-gamma probability distribution; $\Gamma(\rho)$ is the gamma function of ρ.

3.5 Reaching probability of rock fragments to roads

The probability of rock fragments from rocky sources is simulated by using the Flow-R software, which can assess rockfall hazards with probabilistic trajectory paths at a regional scale. Rockfall source areas introduced in Section 3.2 are input data in Flow-R. Besides, the rockfall trajectory path is determined by important input data, the reach angle. It is the arctangent of the line which connects the rockfall source area with the most distant boulder (Figure 4, Eq.3). The simulation assumes that the falling blocks stop at the point of intersection of the above-mentioned line with the topography where the energy is 0 (Copons et al., 2009).

$$\theta = \arctan (H/L) \quad (3)$$

Where θ is the reach angle, which is from the vertical drop H and the horizontal component of the travel distance L. The longer the travel distance is, the lower the reach angle value will be.

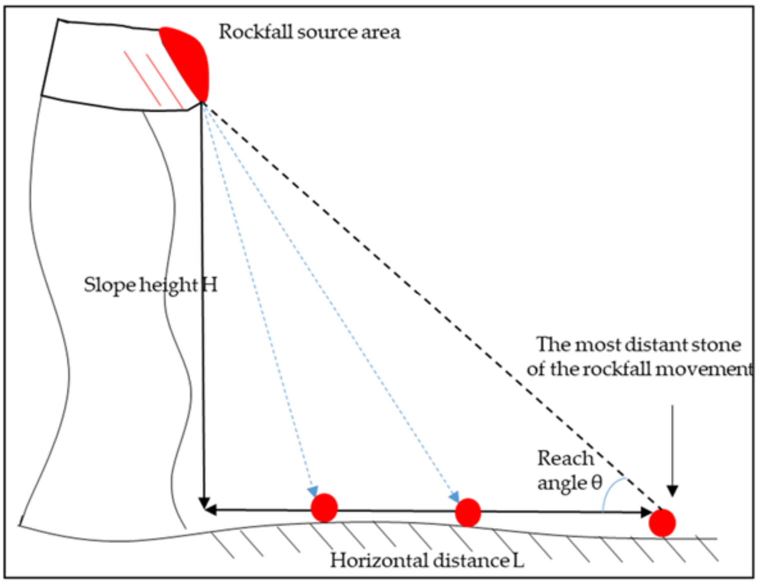


Figure 4. Reach angle diagram

3.6 Rockfall hazard probability assessment

The purpose of rockfall hazard assessment in this study is to know the possibility of rockfall fragments reaching the road with a certain magnitude under a certain return period. We multiply four probabilities to assess the hazard level (Eq.4). By overlaying the hazard probability map with the highway map, risky road sections can be identified finally.

$$H = P_s \times P_t \times P_v \times P_r \tag{4}$$

Where H is rockfall hazard probability; P_s is the spatial probability of rockfall sources introduced in Section 3.3.1, P_t is the temporal probability of rockfall sources, P_v is size probability of rockfall sources; P_r is reaching probability of rockfall sources to roads.

3.7 Validation

The generated susceptibility maps by MLRM and RFM were validated by using the Receiver Operating Characteristics (ROC) curve (Cruden et al., 1991; Zezere et al., 2017) and expert re-evaluation on for typical slopes in the field. The larger the value of AUC is, the more effective the evaluation result is. The performance of the two models was measured and compared. The units with the highest class susceptibility by the better model will be determined as rockfall source areas. The temporal probability of rockfall sources validation is not easy to be taken due to limited data on hazard occurrence time. Whereas the size probability of rockfall sources is validated by calculating the R-squared value of the exceeding probability distribution curve.

Parameters for rockfall reaching probabilities are firstly calibrated and determined by repeating trials in Flow-R on two historical rockfall events with detailed run-out measurements. Then the selected parameters are further validated by simulating the other 35 of the 108 historical rockfalls with accumulation area information. Pictures of historical rockfalls by UVA are also used to verify the accuracy of runout distance and reach angle.

4. Results

4.1 Rockfall source area determination

Using the SAT method, the slope angle threshold for the study area is determined as 27° (Figure 5). This value is smaller than the research result by other researchers (e.g. Wieczorek et al., 1998, and Guzzetti, Reichenbach, et al., 2003; Jaboyedoff and Labiouse; Frattini et al., 2008; Matasci, Jaboyedoff, et al., 2015). It underlines that the areas with different topography or geology characteristics do not have the same SAT value. Based on this SAT value, 763 847 cells with a slope greater than 27 ° are selected as preliminary rockfall sources from the total area with 1 443 012 cells.

Figure 5 shows the relationship between slope factor and collapse disaster with 3 ° step size. In Figure 5, the rockfall area ratio equals the source area of rockfall within a certain slope degree range divided by the total rockfall areas; graded area ratio equals to the slope within certain slope degree divided by the total study area. When the rockfall area ratio is in the interval of [27°, 78°], the rockfall area ratio begins to be greater than the slope graded area ratio. About 80% of rockfalls are concentrated in this section. After 27°, the rockfall area ratio begins to be greater than the slope graded area ratio. Therefore, the area with a slope greater than 27 ° is selected as the preliminary rockfall source.

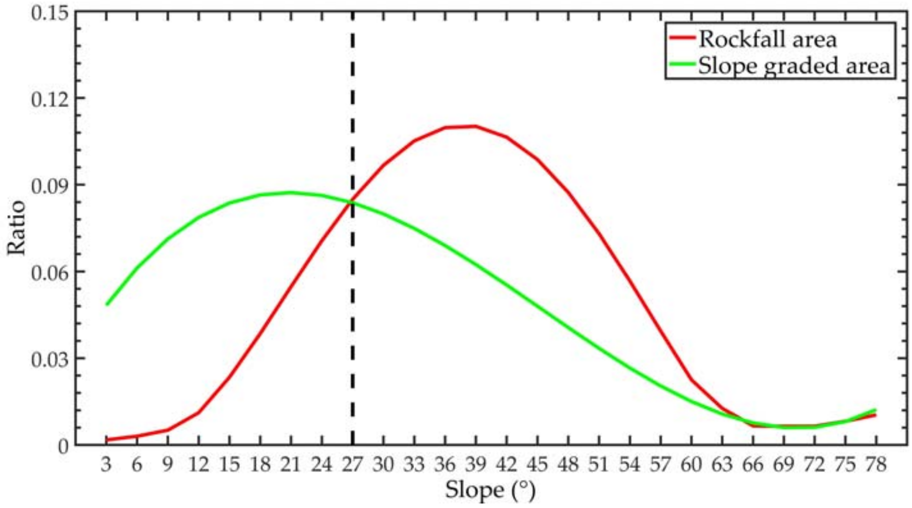


Figure 5. Relationship between rockfall distribution frequency and slope angle

The preliminary rockfall sources are further classified by considering eight conditioning factors, such as slope, aspect, elevation, slope structure, lithology, joint density, distance to road and land-use type etc. Table 2 lists out these factors. The numerical factors, such as slope degree, are reclassified by using MDLP (Rissanen, 1978; Vitanyi, 2000).

Table 2. Rockfall source area conditioning factors with classes by using minimal description length principle

Conditioning factors	Classes	Conditioning factors	Classes
Slope (°)		Slope structure	Over-dip slope
	27 - 34		Under-dip slope
			Oblique slope
	34 - 38		Tranverse slope
	38 - 46		Anaclinal slope
		Lithology	Soft rock or soft and hard interbedded rock
	46 - 51		Soft rock with hard rock
	≥51		Hard rock with weak interbed
Aspect (°)	North	Joint density (number of joints /m)	0 - 0.68
	Flat-North-Northeast		0.68 - 0.8
	Northeast-East		0.8 - 1
	East-Southeast		1 - 1.5
	Southeast-Southwest		1.5 - 3
	Southwest-Northwest		0 - 30
	Northwest-North	Distance to road (m)	30 - 70
Elevation (m)	290 - 350		70 - 250
	350 - 430		250 - 260
	430- 470		260 - 360
	470 - 550		360 - 410
	550- 670		≥410
	670 - 1100	Land-use type	Grasslands and Open Wood
	1100 - 1140		Rock and Exposed Soil
	1140 - 1240		Water
	≥1240		Rural Settlement

Figure 6 shows susceptibility maps of rockfall source area by MLRM and RFM. In terms of the ranking of importance of the factors, distance from road and slope is the most important as shown in both the models (Figure 7). However, a big difference exists in lithology and land use. RFM ranks lithology as a relatively insignificant predictor but this factor is treated to be the third important in MLRM. As to the land-use factor, it is not effective or the least significant in the ranking in both models.

ROC curve analysis shows that the success rate of MLRM is 93%, while RFM is 5% higher (Figure 8). It indicates that RFM has a better model performance than MLRM in the study area. The prediction performance of the two models was further evaluated and compared in the field (Table 3).

Four typical slopes along G318 road were selected for validation, including steep slope with sandstone inter-bedding with mudstone, gentle rocky slope, steep slope with high vegetation cover, and gentle rocky slope with high vegetation cover. Each slope was evaluated by the experts resulting in the possibility of rockfall source potential. In the comparison results of the susceptibility of four typical slopes, the RFM results of the susceptibility of three slopes were consistent with the expert judgments. It further shows that RFM has a better evaluation effect.

Using the result from RFM, rockfall source area spatial probabilities are finally divided into four classes: [0, 0.25), [0.25, 0.60), [0.60, 0.88) and [0.88, 1]. The percentage of historical rockfalls in each class is shown in Table 4. A region with a spatial probability of [0.88, 1] from RFM was finally selected to simulate rock fragment trajectories. This allowed identifying 5349 grid cells (10×10 m) as final sources of rockfalls, about 0.53 km^2 (0.70% of the total study area). Inspection of the map of the rockfall source cells revealed a good agreement with the local morphology, and in particular with the location of the edges of the rock cliffs and with the location of the release areas of known rockfall events.

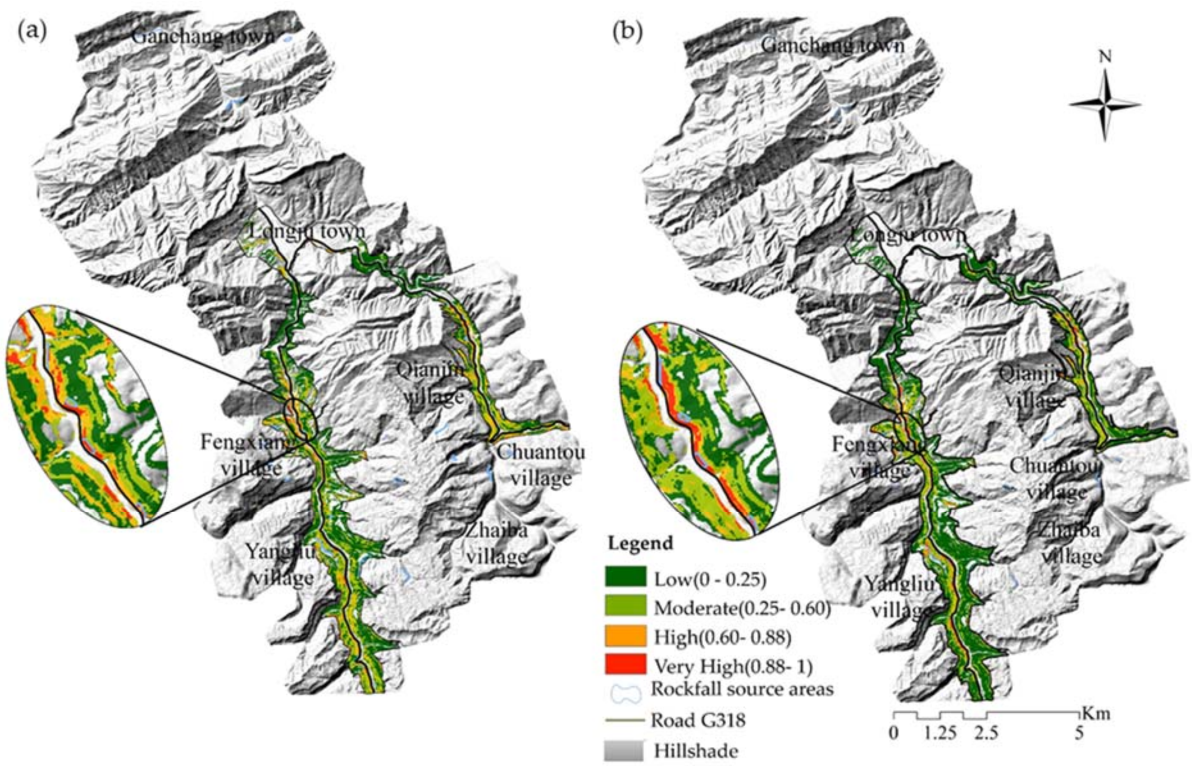


Figure 6. Spatial probability maps of rockfall sources by (a) MLRM and (b) RFM

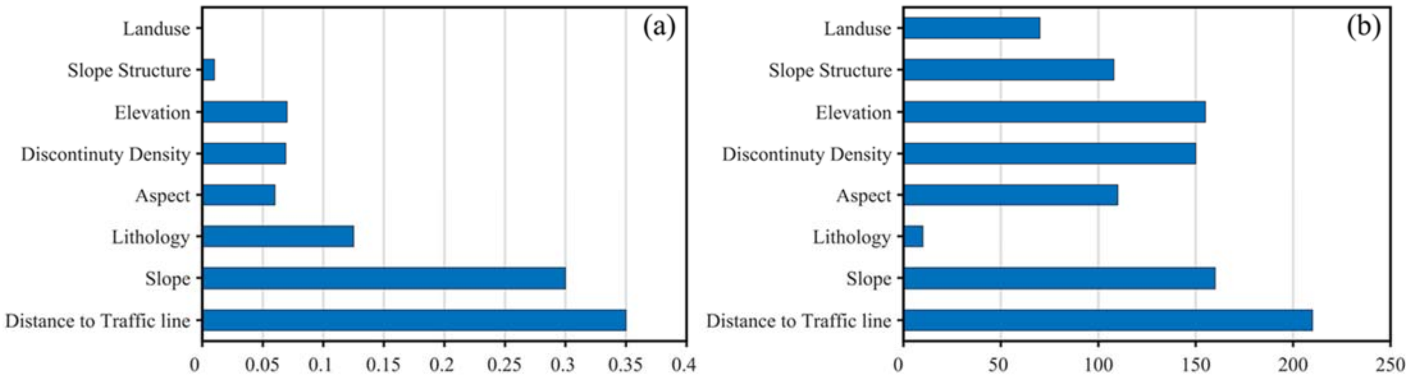


Figure 7. Ranking of predictor factors by (a) MLRM and (b) RFM.

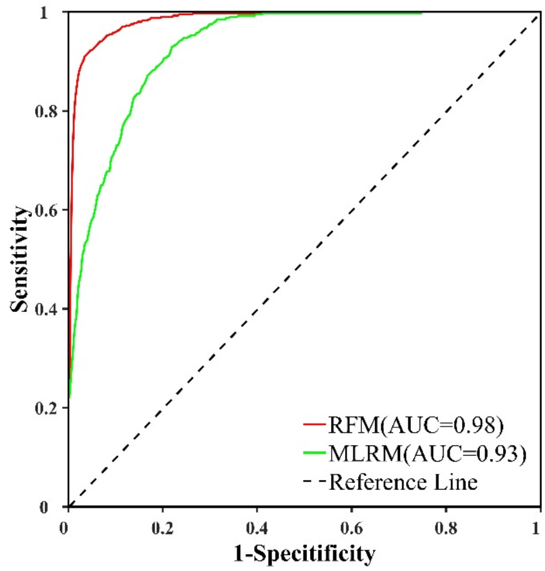


Figure 8. Accuracy comparison between MLRM and RFM

Table 3. Comparison of rockfall source area probability result from MLRM, RFM, and field evaluation

(The legend is the same as in Figure 6)

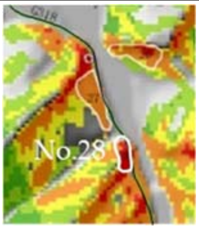
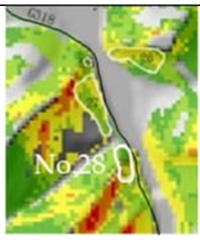
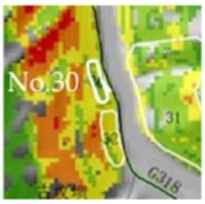
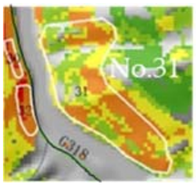
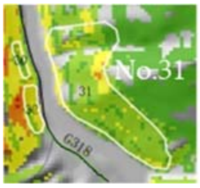
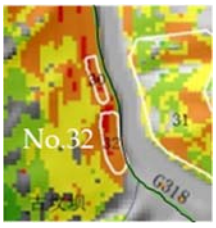
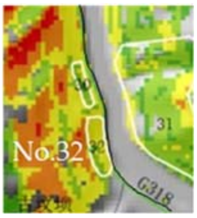
Number	MLRM	RFM	Field photos	Expert evaluation
No.28				The integrity of the rock mass is good, but there are blocks piled up on the slope. It is a high-medium class. The result from MLRM is accurate.
No.30				The slope surface is gentle, and there is no rockfall accumulation, which is a middle-class-prone area. The result from RFM is accurate.
No.31				The vegetation coverage rate is high and the slope surface is gentle. It is a low-class prone area. The result from RFM is accurate.
No.32				The vegetation coverage rate is very high. There are no exposed rock blocks. It is a medium-low class prone area. The result from RFM is accurate.

Table 4. The Proportion of historical source area in each probability grades classified by RFM (Area: km^2)

Probability grade	A- Area(percentage)	B- Historical rockfall source area (percentage)	B/A- Proportion of historical rockfall source areas in different probability grades (%)
0 -0.250	68.140 (89.220%)	0.060 (28.070%)	0.088
0.250-0.600	6.400 (8.390%)	0.040(17.390%)	0.625
0.600-0.880	1.290 (1.700%)	0.040 (18.530%)	3.101
0.880-1	0.535(0.700%)	0.080 (36.010%)	15.094
Sum	76.365(100%)	0.220(100%)	18.908

Quantitatively, the simulation efficiency of our approach can be improved by 40 times, without losing data of historical or field survey determined rockfalls. Detailed data are shown in Table 5.

Table 5. Comparison between SAT model and our approach in terms of simulation efficiency

	SAT model	Our approach	Benefit or Loss
Number of potential source cells	160823	4002	Benefit: about 40 times
Number of historical or field survey determined rockfalls	1364	1345	Loss: about 0.014 times

4.2 Temporal and size probability of rockfall sources

According to the records of the rockfall database in the study area, the historical time is 31 years. According to Equation 1, temporal probabilities in different recurrence periods (5, 20, 50 years) for each unit are calculated and the relative maps were generated (Figure 9). The map for 50 years, for example, shows that the highest temporal probability is 0.798 (Figure 9a). In order to enhance the distinction, the temporal probability maps are divided into five classes by the Natural Breaks method. The areas with the highest class with temporal probability from 0.536 to 0.798 mainly distribute along the road G318 in the north part.

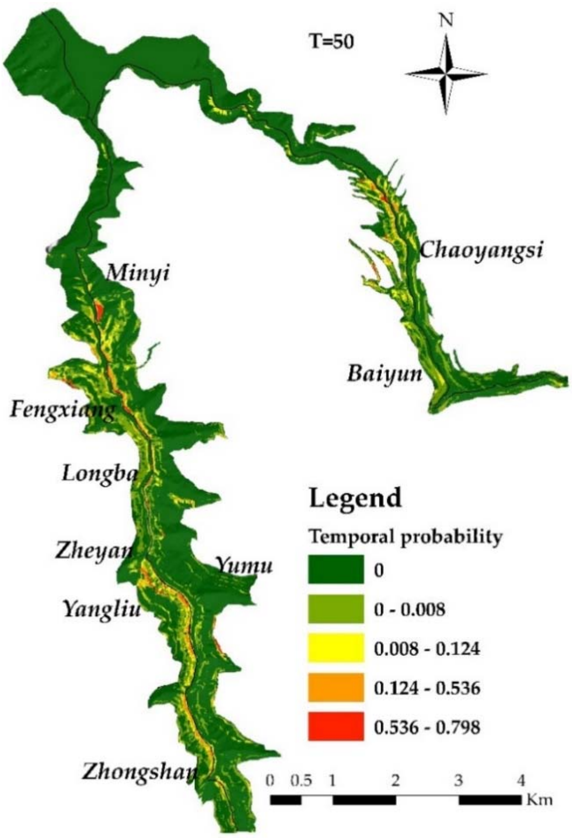


Figure 9. Fifty years return period maps of rockfall temporal probability along the national road G318

By using 31 historical rockfalls with volume records, the size probability curve is created according to Equation 2 with an R^2 value of 0.956 (Figure10). The rockfall volume ranges from 1 000 m^3 and 10 000 m^3 with size probability from 0.826 and 0.395, which means: (1) the occurrence probability of rockfalls with volume greater than 1 000 m^3 in the study area is 0.830; (2) the occurrence probability of rockfalls with volume more than 10 000 m^3 is 0.395. It indicates that small-scale rockfalls are more frequent than the larger ones, which is consistent with the real performance of rockfall hazards in the study area.

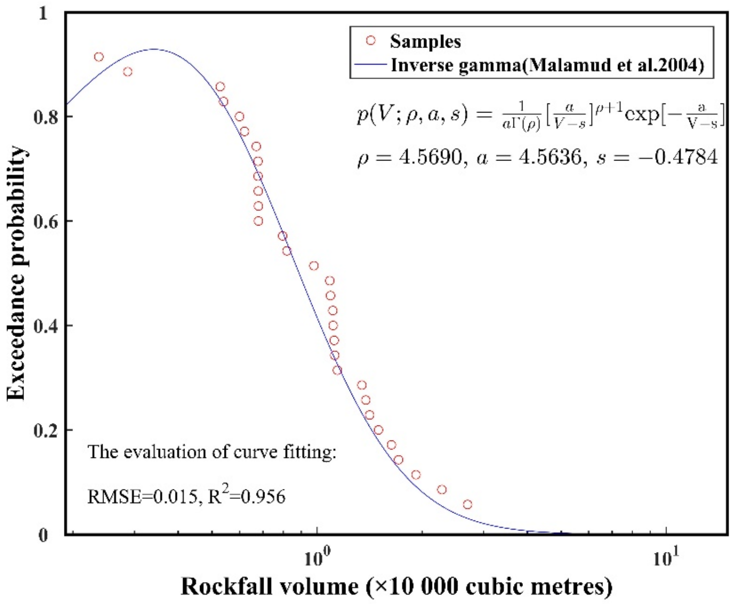


Figure10. Size probability distribution curve fitted by inverse gamma for rockfalls along the national road G318

According to Equation 4, the hazard probability maps are generated for three return periods (5, 20, and 50 years) and two-volume scenarios (10 00 m^3 and 10 000 m^3). The probability values of 10 00 m^3 volume scenarios in Figure 9b comprise five categories from very low (0 - 0.112) to very high (0.538- 0.801). The area with obvious high probability is close to the road section from Minyi to Fengxiang. It is located in the south of Longjuba anticline and the core part of Matouchang syncline, with strong geological tectonic activities. Lithology in the area is mainly covered by purplish-red mudstone, sandstone, and shell stone as interlayers. Rocks are seriously weathered in this section.

4.3 Rockfall reaching probability and verification in field work

Table 6 summarizes the algorithms and parameters used in Flow-R by repeating trials on two historical rockfall events (Figure 11).In terms of reach angle, we found three possible values (15 °, 25 °, and 27 °) according to the law of reach angle distribution of historical rockfalls. To find out the most suitable value, we compared the simulated travel distance with the measured value in the field (Figure 11). Therefore, reach angle 25 ° is adopted as the preliminary reach angle value for further verification in other 35 rockfall modelings. The further simulation results show that the average difference value is 1.97 m with an evaluation error ratio of 3.66% (Figure 12). The simulated reaching area basically matches the influence areas where rock fragments are scattered (Figure 13). It indicates that simulated travel distance fits well with the value investigated in the field by using a reaching angle of 25 °.

Table 6. Initial modeling conditions and parameters used in Flow-R to perform the transportation simulation

Algorithms and Parameters	Value
Flow direction algorithm	Holmgren modified algorithm
Exponent α	1
Persistence factor	Gamma_ 2000
Friction model	Simple Coulomb friction model
Minimum reach angle	25 °

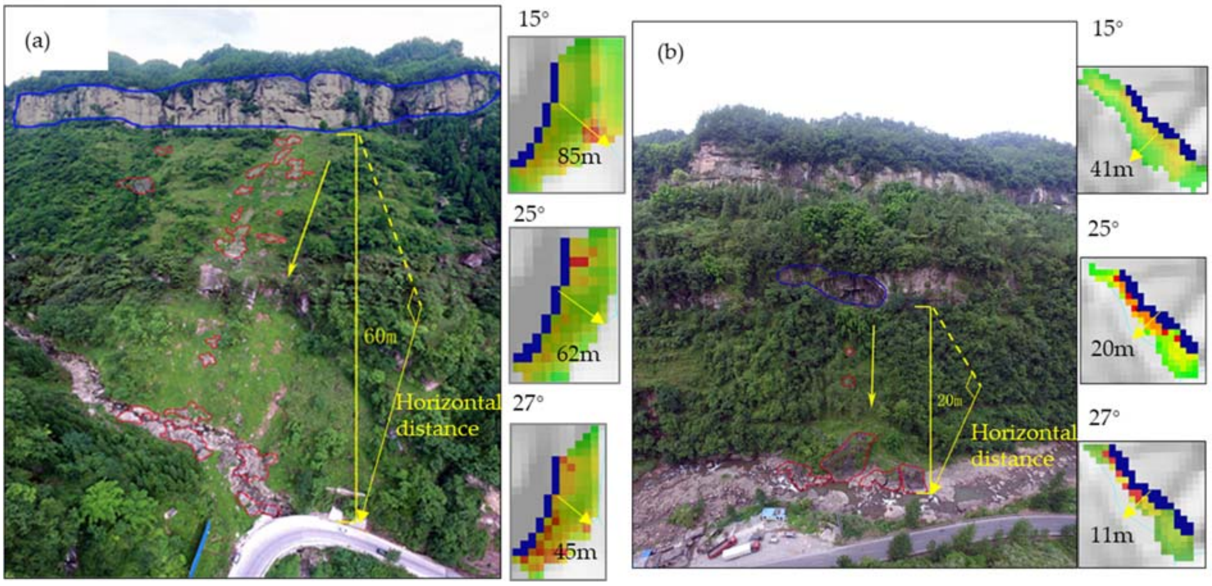


Figure11. Comparison of horizontal travel distance between the field measurement and numerical simulation for reach angle determination. (The source area is pointed out using blue closed lines. Rockfall fragments are identified using red closed lines. Horizontal travel distances marked by yellow lines.)

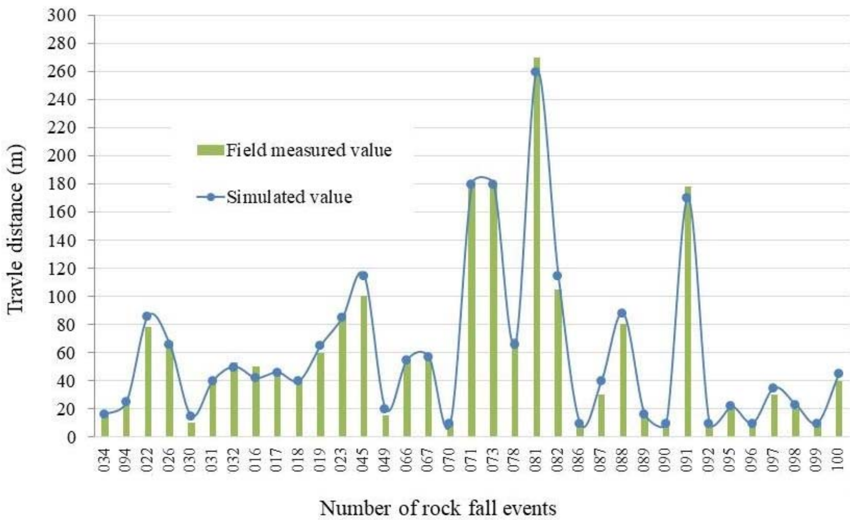


Figure 12. Horizontal travel distance comparison between the field measured value (green column) and (b) the simulated value (blue line) for 35 historical rockfalls along the national road G318.

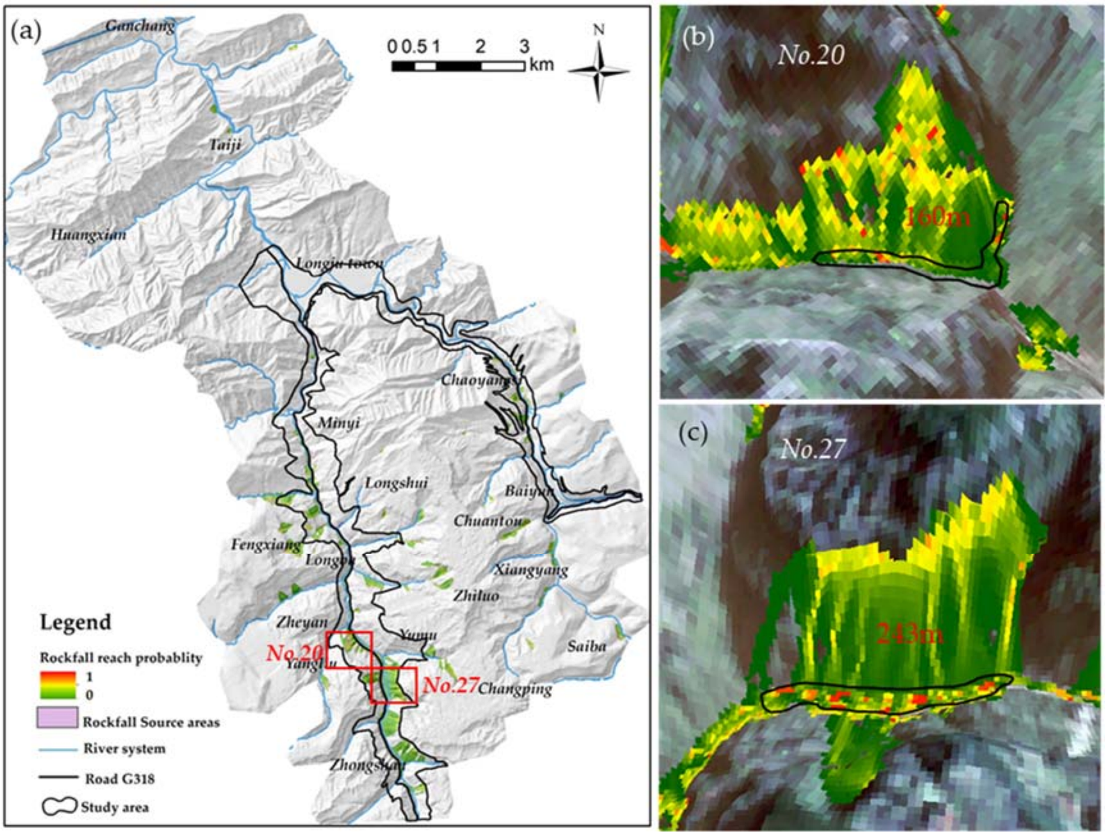


Figure 13. (a) Reaching probability map of rockfall along the national road G318. (b)(c) Enlarged view of partial rockfall. (Rockfall fragment distribution area is identified using black closed lines)

4.4 Rockfall hazard probability along the road G318

Rockfall hazard probability was calculated according to Equation 4 by overlapping the maps of spatial (Figure6), temporal (Figure 9) and size probability of rockfall sources, and the reaching probability map (Figure 13). The final maps (Figure 14) of rockfall hazard probability were set for two size scenarios and three return periods. When the volume scenario is 1 000 m^3, the maximum probability value increases from 0.123 to 0.661 with the increase of return period (Table 7). Because of the lower size probability of 10 000 m^3, the maximum hazard probabilities are generally half of the values under the size scenario 1 000 m^3.

If the above results are associated with the national road G318, we can find out the risky sections with detailed impact probability for road G318 due to rockfall fragments. Table 8 lists the length of the impacted roads under each return period and volume scenario. Among them, Minyi village, Longba village, Zheyan village, Yumu village, and Zhongshan village are located along the 318 national highway and 553 county roads in Chaoyangsi village and Xiangyang village are the most affected by the rockfall, so the protection and control should be strengthened. G318 section with high hazard is mainly located in Minyi Village and Zheyan Village.

In general, for different return periods and collapse scales, the total length of damaged sections is 8.19 km, and the damage degree of collapsed roads in the 50-year return period is higher than that in the 20-year return period and the 5-year return period. The severity of road damage caused by the collapse disaster with a scale of larger than 1000 m^3 is higher than that of the collapse disaster with a scale of larger than 1000 m^3. In the return period of 5 and 20 years, there is no very-high hazard class of road section, but mainly concentrated in the rainfall conditions of 50 years return period, and the influenced road section caused by the collapse of larger than 1000 m^3 is 0.510 km, and the influenced road section caused by the collapse of larger than 10,000 m^3 is 0.430 km. In the 5 years of the rainfall return period, there is no high-class risk of road section but mainly concentrated in the rainfall conditions of the 50 and 20 years of the return period.

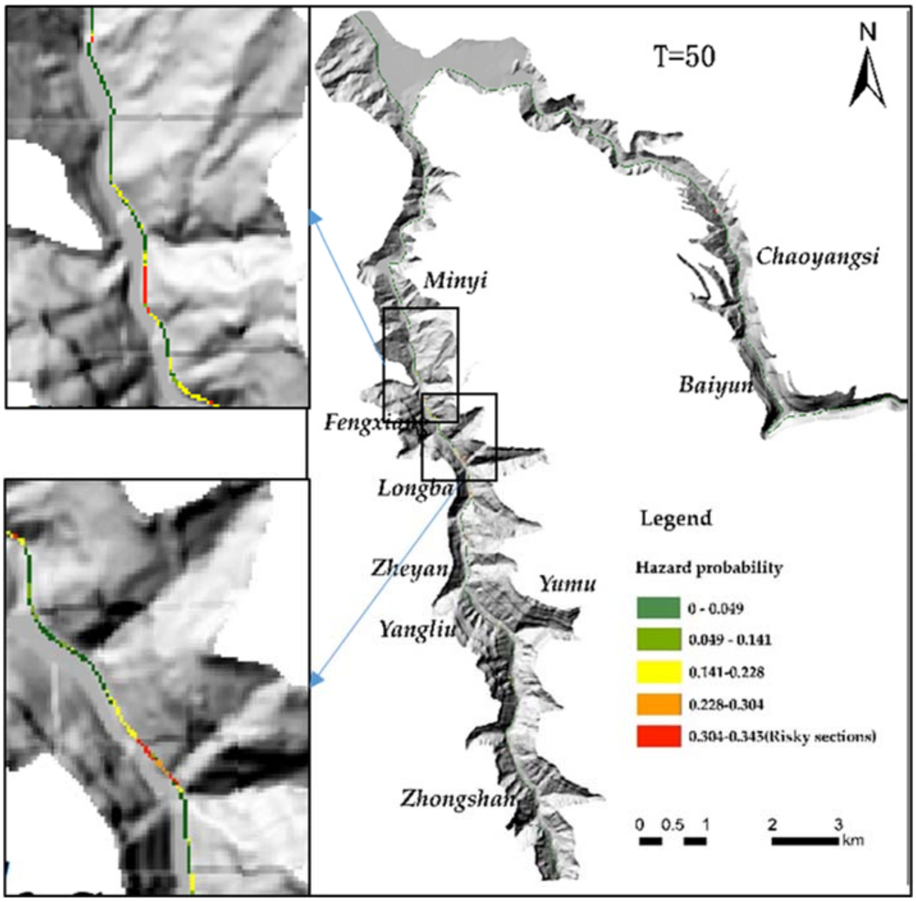


Figure14. Rockfall hazard probability map (at the volume scenario of 10 000 m^3) for the national road G318. The map is classified into five categories from high to low, which are overlapped by the road sections.

Table 7. Maximum hazard probability under three return periods (5, 20, and 50 years) and two-volume scenarios (1 000 m^3 and 10 000 m^3)

Volume scenario (m^3)	Return period (years)		
	5	20	50
1 000	0.123	0.393	0.661
10 000	0.059	0.188	0.287

Table 8. Influence length of road G318 induced by rockfalls (unit: Km) with two size scenarios and three return periods

Size scenario / Return period (years)	Very High		High		Medium		Low		Very low	
	1 000	10 000	1 000	10 000	1 000	10 000	1 000	10 000	1 000	10 000
5	0	0	0	0	0	0	0.510	0.480	7.680	7.710
20	0	0	0.430	0	0.090	0.510	1.490	1.500	6.180	6.180
50	0.510	0.430	0.010	0.080	1.480	0.010	0.010	1.490	6.180	6.180

5. Discussion

In understanding and analyzing rockfall hazard risk, it is very important to identify the source areas, predict the temporal, size, and reaching probability.

5.1 Difficulties in identifying source area of rockfall

The identification of the source area is the first step. The fineness of source area identification has an important impact on the following steps, such as the fragment trajectory and rockfall size analysis. However, the source area of historical rockfall hazard data is often missing or mixed with the rock debris accumulation, so it is difficult to identify the source area. Luckily, the slope angle threshold is found out to be 27°in this study, according to the relationship between the historic data and the slope. The area above this angle is preliminarily selected as source areas. After the preliminary screening of the collapse source area in the study area by using SAT method, we conducted a secondary screening of the initial source results in the study area by using various models. By using and comparing MLRM and RFM, the final source areas are determined and had a good accuracy after validation. Importantly, the efficiency of trajectory simulation followed by our approach can be improved by 40 times, without losing data of historical or field survey determined rockfalls.

Due to the special topography and geological conditions, there are a large number of multi-stage scarps in the study area (as shown in Figure 15), and more accurate source area identification is required. In the future, more detailed work will be focused on the source stage scarp identification.

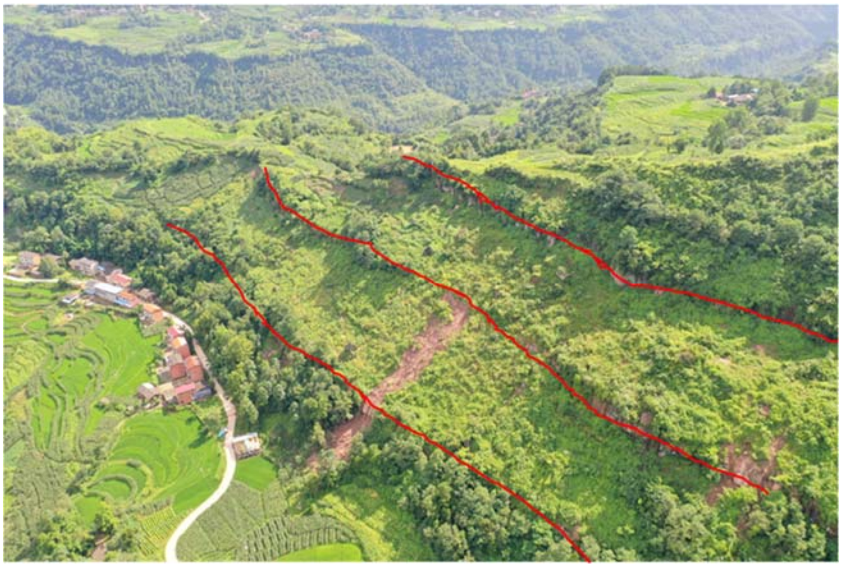


Figure15 UVA photos of multistage rockfall

5.2 Complexity and difficulty in the time probability calculation

The temporal probability and the size probability are important considerations in rockfall hazard analysis. In practice, the temporal probability calculation is a difficult problem, on the premise that there should be a large number of historical collapse time data to analyze the statistical law. However, due to the sudden occurrence of collapse, it is difficult to obtain a large number of time data, which requires a lot of monitoring work. The same is true of the probability of scale surpassing, which requires the scale data of every rock collapse in history for statistical analysis. Rockfalls in this area mostly happen along the traffic road. For road accessibility, rock fragments are quickly cleaned after hazard events so that historical influence area record is always unavailable. Both of them have always been difficult points in collapse hazard analysis.

It is difficult to calculate the time probability for multistage cliffs. There are a large number of historical rockfalls in the study area, such as the PT rockfall in Figure 16. The first occurrence time of PT rockfall is June 29, 2019. The second occurrence time is July 5, 2020. This kind of multi-stage collapse disaster causes serious economic loss and great psychological pressure on the victims' families. Therefore, it is necessary to solve the problem of how to accurately predict the time and calculate the time probability of multiple collapse disasters of multistage cliffs. But we need to do long-term monitoring and collect large amounts of historical occurrence time data for predicting these types of collapses. So, the establishment of a historical collapse time database in the study area is needed in the future.

Figure16 A typical case of multiple collapses in the study area, PT collapse (The first occurrence time of PT collapse is June 29, 2019. The second occurrence time is July 5, 2020.)

5.3 Simulation problem of hazard calculation

This paper adopted energy balance theory, GIS spatial statistical function, and flow theory to simulate the influence area of rock fragments. The parameters in the simulation are calibrated and validated by historical records collected by field investigation. The results indicated that the accuracy of the quantitative analysis is very high. However, the failure motion of collapse is various, which was ignored in the Flow-R simulation. There are multiple failure modes of collapse, such as dumping, falling, and sliding. The simulation procedure simplifies the laws governing rock-mass failures and blocks propagations.

Compared with STONE, Rockyfor3D, RAMMS, DDA, Flow-R can simulate the motion of multiple collapsing sources on the regional scale by using less time and costs. But we can not consider the failure modes by Flow-R tools. In the future, we will optimize the simulation considering rock source volume, block shape, failure modes, and mechanical parameters and achieve a three-dimensional dynamic display of the collapse process at the regional scale.

The simulation of multistage scarps should consider the energy transfer caused by the collision between the scarps or the induced collapse of the scarps. For example, in Figure 16 PT collapse is induced by the falling of a boulder in the upper layer. For the complexity of collapse, more research work is needed in the future.

6. Conclusion

A national road G318 in west Hubei China is prone to the high-frequency rockfall hazard. In this paper, rockfall hazard and its probability are quantitatively assessed. Rockfall source areas are firstly identified by the slope angle threshold method and then optimized by using the susceptibility mapping method. Slope degree 27° is determined as the threshold angle of rockfalls in the study area. The multivariate logistic regression model and random forest model are compared in terms of the model performance. Source area cells selected by the random forest model are finally chosen and applied for rockfall reaching probability assessment. Compared to the slope angle threshold method, the source areas determined by our approach are more accurate when geology data is available. Meanwhile, the advantages of trajectory simulation efficiency are obvious and without losing data of historical or field survey determined rockfalls. In addition, the size probability and temporal probability for rockfall sources are calculated considering two size scenarios (1 000 m^3 and 10 000 m^3) and three return periods (5, 20, and 50 years).

The selection of parameters is very important for the rockfall trajectory simulation. The smallest reach angle affects the farthest horizontal distance and then the reaching probability. In this paper, 25 ° is determined as the smallest reach angle. The horizontal distance is then simulated by Flow-R and then validated with the historical rockfalls with field-measured records. In the future, we will optimize the simulation considering rock source volume, block shape, failure modes, and mechanical parameters and achieve a three-dimensional dynamic display of the collapse process at the regional scale.

Rockfall hazard probability is finally obtained by integrating the spatial, temporal, size probability of source areas and the reaching probability of rock fragments. In the rainfall return period of 5 and 20 years, there is no high hazardous road section, but they are mainly concentrated in the conditions of 50 years return period. In this case, the risky road section caused by rockfalls larger than 1000 m^3 is 0.510 km. Among them, villages including Minyi village, Longba village, Zheyan village, Yumu village, and Zhongshan village are identified along the national road G318, so the protection and control are suggested in these villages. Although some limitations exist, the results show good fitness with the measurements by field investigation.

7. Patents

Author Contributions: Conceptualization, Lixia Chen, Kunlong Yin; methodology, Lixia Chen, Yu Zhao, Lei Gui; investigation, Lixia Chen, Yuanyao Li, Lei Gui; writing, Lixia Chen, Yu Zhao; writing—review and editing, Dhruba Pikha Shrestha; visualization, Yu Zhao,Lixia Chen, Lei Gui; All authors have read and agreed to the published version of the manuscript.”

Funding: This research was funded by the National Natural Science Foundation of China (No.41877525).

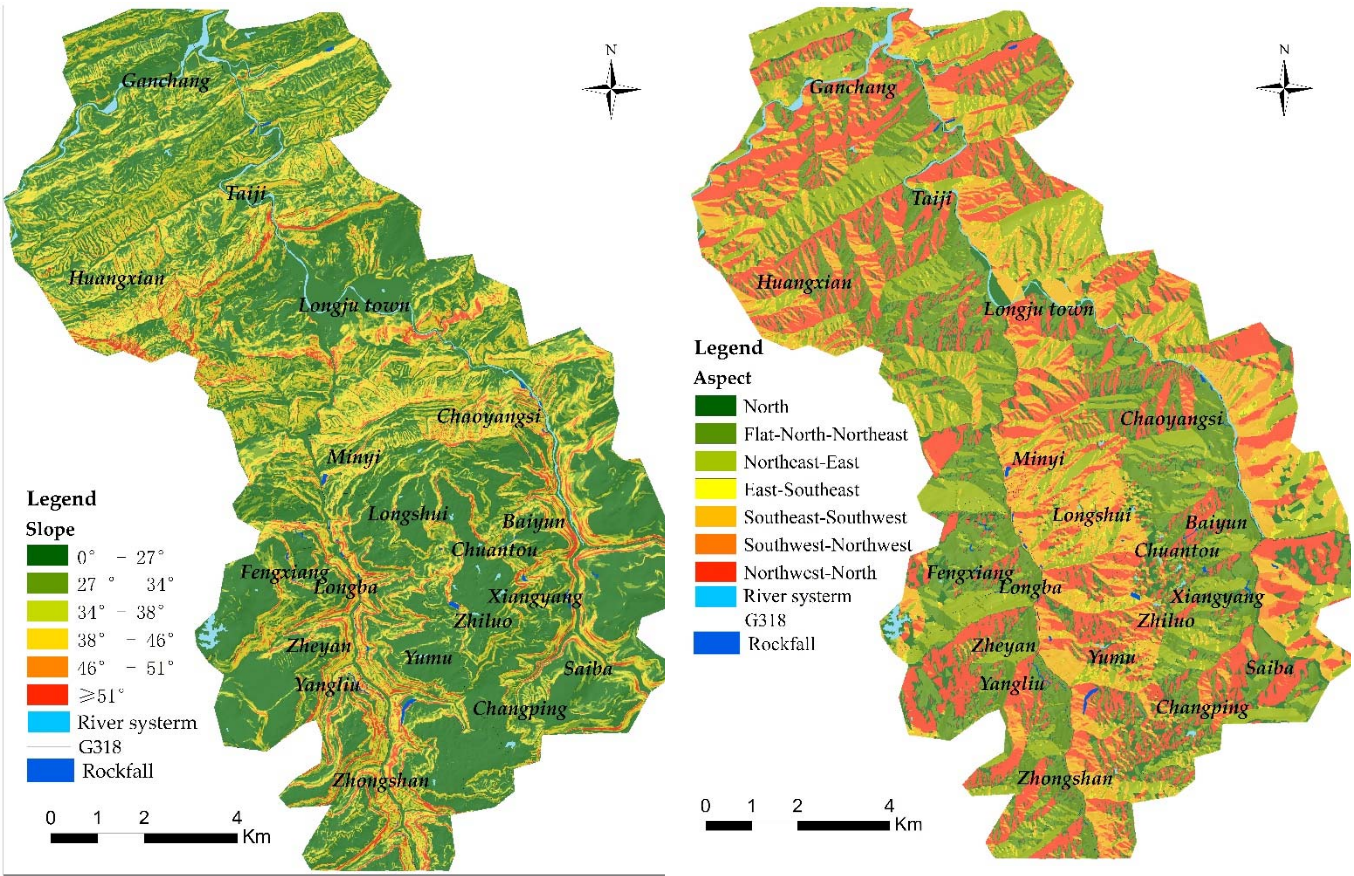
Acknowledgments: We would like to thank the editor and anonymous reviewers providing valuable contributions and constructive comments.

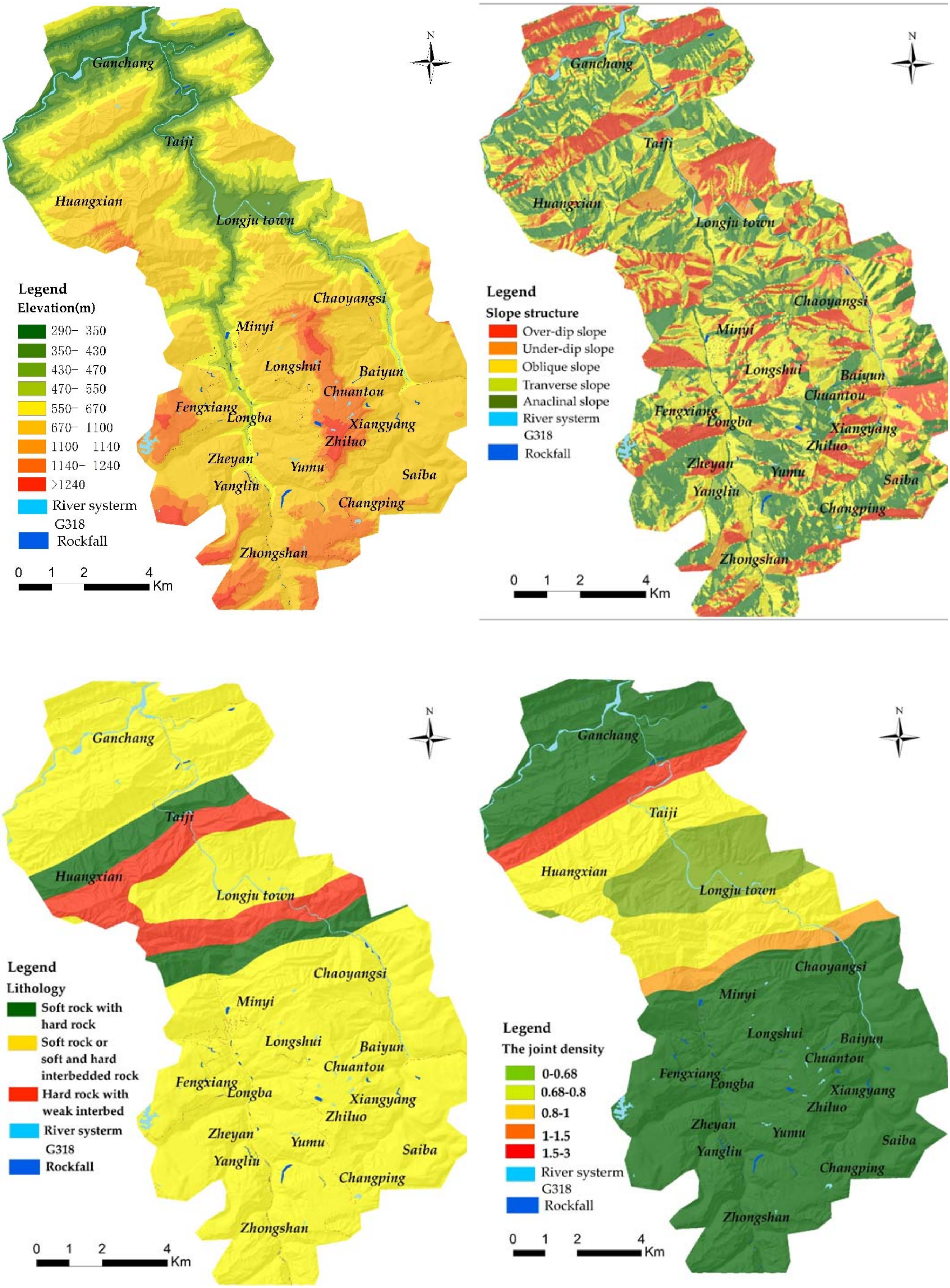
Conflicts of Interest: The authors declare no conflict of interest.

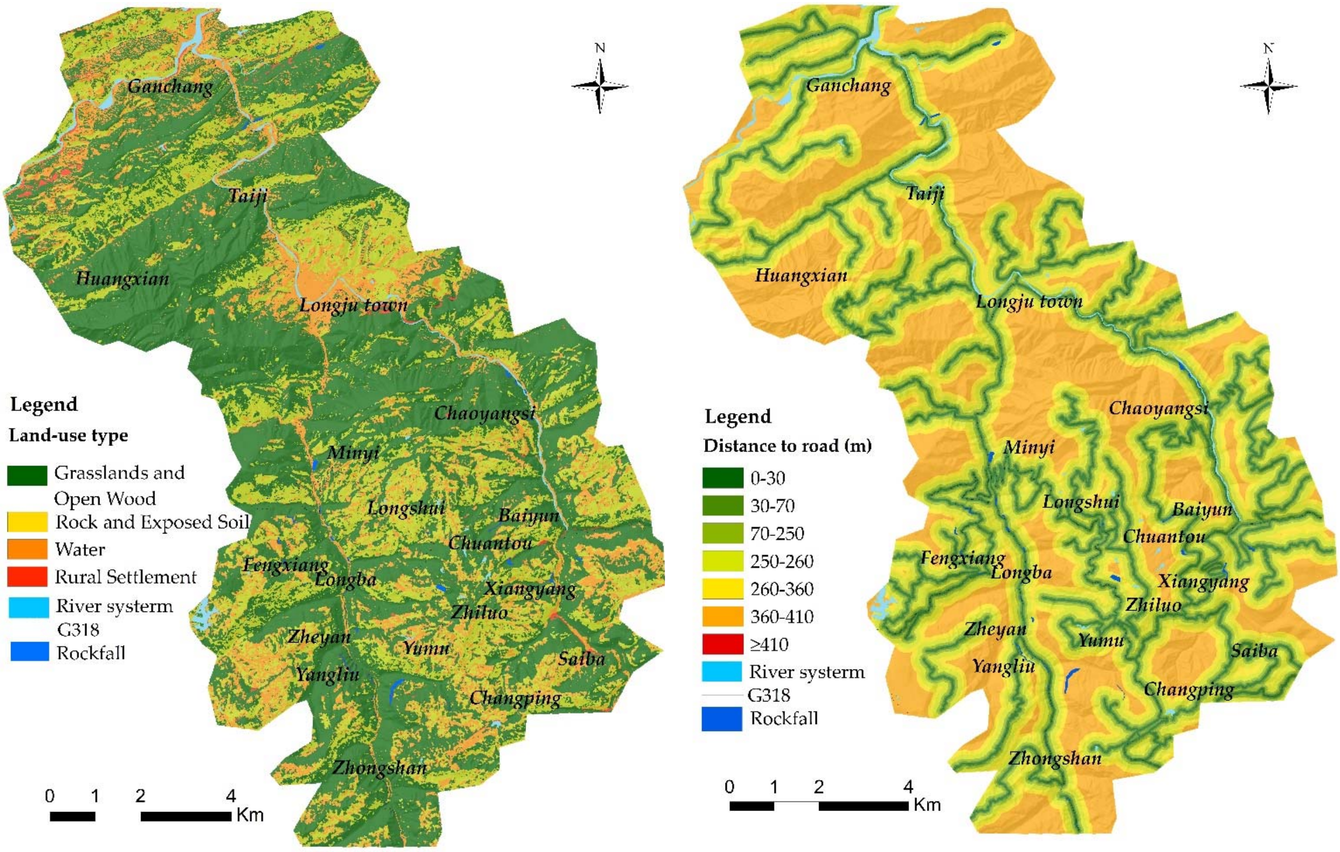
Appendix A

Factors

Eight factors including slope, aspect, elevation, slope structure, lithology, joint density, land-use type, and distance to the road are selected for source area identification (FigureA1).

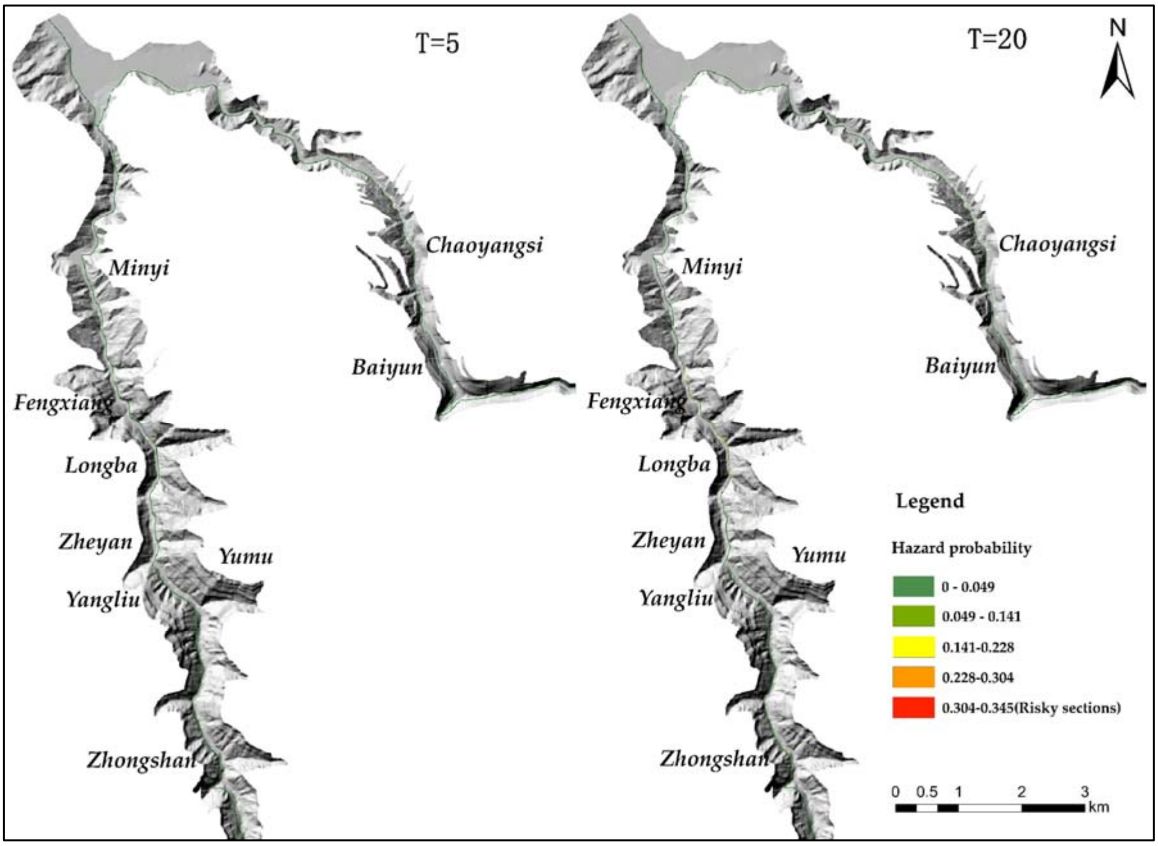



Ganchang
Taiji
Huangxian
Longju town
Chaoyangsi
Minyi
Longshui
Baiyun
Chuantou
Fengxiang
Longba
Xiangyang
Zhiluo
Zheyan
Yumu
Saiba
Yangliu
Changping
Zhongshan
N
Legend
Elevation(m)
290- 350
350- 430
430- 470
470- 550
550- 670
670- 1100
1100 1140
1140- 1240
>1240
River systerm
G318
Rockfall
0 1 2 4 Km
Legend
Slope structure
Over-dip slope
Under-dip slope
Oblique slope
Tranverse slope
Anaclinal slope
River systerm
G318
Rockfall
Legend
Lithology
Soft rock with hard rock
Soft rock or soft and hard interbedded rock
Hard rock with weak interbed
River systerm
G318
Rockfall
Legend
The joint density
0-0.68
0.68-0.8
0.8-1
1-1.5
1.5-3
River systerm
G318
Rockfall



FigureA1. Conditioning factor maps for rockfall source area identification

The complete result of the hazard probability is shown in FigureA2. FigureA2 includes the hazard probability results in the 10-year and 20-year return periods.



FigureA2. Rockfall hazard probability assessment (at the volume scenario of 10 000 m^3) in the 10-year and 20-year return periods for the areas along national road G318 in Longjuba.

References

Azzoni, A.; Labarbera, G.; Zaninetti, A. Analysis and prediction of Rockfalls using a mathematical-model. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts. 32, 709-724. https://doi.org/10.1016/0148-9062(95)00018-C, 1995.

Bak, P., Tang, C., and Wiesenfeld, K.: Self-organized criticality, Phys Rev A Gen Phys, 38, 364-374, https://doi.org/10.1103/physreva.38.364, 1988.

Blahut, J., Horton, P., Sterlacchini, S., and Jaboyedoff, M.: Debris flow hazard modelling on medium scale: Valtellina di Tirano, Italy, Natural Hazards and Earth System Sciences, 10, 2379-2390, https://doi.org/10.5194/nhess-10-2379-2010., 2010.

Brawner, C. O.; Wyllie, D. C.: Rock slope stability on railway projects: Proceedings of the American Railway Engineering Association Regional Meeting: Vancouver, British Columbia, Canada. 8 p., 1975.

Budetta, P. Assessment of Rockfall risk along roads. Natural Hazards and Earth System Sciences. 4, 71–81. SRef-ID: 1684-

9981/nhess/2004-4-71., 2004.

Bouali, E. H., Oommen, T., Vitton, S., Escobar-Wolf, R., Brooks, C. Rockfall Hazard Rating System: Benefits of Utilizing Remote Sensing. Environmental and Engineering Geoscience. 23 (3), 165–177. https://doi.org/10.2113/gseegeosci.23.3.165, 2017.

Benchelha, S., Chennaoui Aoudjehane, H., Hakdaoui, M., El Hamdouni, R., Mansouri, H., Benchelha, T., Layelmam, M., Alaoui, M., Landslide susceptibility mapping in the Commune of Oudka, Taounate Province, north Morocco; A comparative analysis of logistic regression, multivariate adaptive regression spline, and artificial neural network models. Environ. Eng. Geosci. 26(2), 185-200. https://doi.org/10.2113/EEG-2243, 2019.

Chen, W., Li, X., Wang, Y., Chen, G., and Liu, S.: Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the Three Gorges, China, Remote Sensing of Environment, 152, https://doi.org/291-301, 10.1016/j.rse.2014.07.004, 2014..

Chung, C.F., Fabbri, A.G., Van Westen, C.J. Multivariate regression analysis for landslide hazard zonation. In: Carrara, A., Guzzetti, F. (eds.), Geographical Information Systems in Assessing Natural Hazards, Springer Netherlands, Dordrecht, 107-133. https://doi.org/10.1007/978-94-015-8404-3_7, 1995.

Copons, R.; Vilaplana, J.M.; Linares, R. Rockfall travel distance analysis by using empirical models (Sol` a d'Andorra la Vella, Central Pyrenees). Nat. Hazards Earth Syst. Sci. 9, 2107-18. www.nat-hazards-earth-syst-sci.net/9/2107/2009/, 2009.

Corominas, J., Ibarbia, I., Luzuriaga, S., Navarro, J. A., Jugo, I., Jurnet, C., and Hürlimann, M.: Rockfall and Debris Flow Hazard Assessment of the Coastal Road of Gipuzkoa (Northern Spain), in: Landslide Science and Practice, 223-229, https://doi.org/10.1007/978-3-642-31319-6_31, 2013.

Crovelli, R. A. Probabilistic models for estimation of number and cost of landslides. U.S. Geological Survey Open File Report 00-249. 23. http://pubs.usgs.gov/of/2000/ofr-00-0249/ProbModels.html. , 2000.

Cruden, D.M. A simple definition of a landslide. Bulletin of the International Association of Engineering Geology. 43, 27-9. https://doi.org/10.1007/BF02590167, 1991.

Dorren L.K.A. Rockyfor3D (v5.2) revealed – Transparent description of the complete 3D rockfall model. ecorisQ paper, 33 ,www.ecorisq.org, 2016.

Frattini, P., Crosta, G., Carrara, A., and Agliardi, F.: Assessment of rockfall susceptibility by integrating statistical and physically-based approaches, Geomorphology, 94, 419-437, https://doi.org/ 10.1016/j.geomorph.2006.10.037, 2008.

Fu, S., Chen, L., Woldai, T., Yin, K., Gui, L., Li, D., Du, J., Zhou, C., Xu, Y., Lian, Z. Community-based landslide hazard probability and risk assessment: A case in west Hubei, China. Natural hazards and earth system sciences discussions, 1-31. https://doi.org/10.5194/nhess-2019-259, 2019.

Guzzetti, F., Reichenbach, P., Wieczorek, G.F. Rockfall hazard and risk assessment in the Yosemite Valley, California, USA. Nat. Hazard. Earth Sys. 3(6), 491-503. https://doi.org/10.5194/nhess-3-491-2003, 2003.

Guzzetti, F., Crosta, G., Detti, R., Agliardi, F. STONE: A computer program for the three-dimensional simulation of rock-falls. Comput. Geosci.-UK. 28(9), 1079-1093. https://doi.org/10.1016/S0098-3004(02)00025-0, 2002.

Heckmann, T., Hilger, L., Vehling, L., and Becht, M.: Integrating field measurements, a geomorphological map and stochastic modelling to estimate the spatially distributed rockfall sediment budget of the Upper Kaunertal, Austrian Central Alps, Geomorphology, 260, 16-31, https://doi.org/10.1016/ j.geomorph.2015.07.003, 2016.

Horton, P., Jaboyedoff, M., Rudaz, B., and Zimmermann, M.: Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale, Natural Hazards and Earth System Sciences, 13, 869-885, https://doi.org/10.5194/nhess-13-869-2013, 2013.

Jaboyedoff, M. CONEFALL 1.0- user’s guide. Quanterra: http://www.quanterra.org/ Manual_ conefall.pdf. date of access: 02.12.2014. 2003.

Jaboyedoff, M.; Labiouse, V., Preliminary assessment of Rockfall hazard based on GIS data. 10th International Congress on Rock Mechanics ISRM 2003 – Technology roadmap for rock mechanics. 575–578. Paper Number: ISRM-10CONGRESS-2003-097, 2003.

Kanari, M., Katz, O., Weinberger, R., Porat, N., and Marco, S.: Evaluating earthquake-induced rockfall hazard near the Dead Sea Transform, Natural Hazards and Earth System Sciences, 19, 889-906, https://doi.org/10.5194/nhess-19-889-2019, 2019.

Li, Z. H., Huang, H. W., Xue, Y. D., and Yin, J.: Risk assessment of rockfall hazards on highways, Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 3, 147-154, https://doi.org/10.1080/17499510902809763, 2009.

Liaw, A. Package“random Forest”. http://stat-www. Berkele. edu/users /breiman/Random Forests., 2012.

Leine, R. I.; Schweizer, A.; Christen, M.; Glover, J.; Bartelt, P. and Gerber, W. Simulation of rockfall trajectories with consideration of rock shape. Multibody System Dynamics 32, 2: 241 - 271. https://doi.org/10.1007/s11044-013-9393-4, 2014:

Lee S. Application and Cross-Validation of Spatial Logistic Multiple Regression for Landslide Susceptibility Analysis. Geosciences Journal, 9(1):63-71. https://link.springer.com/content/pdf/10.1007/BF02910555.pdf, 2005.

Liu, H., Wang, X., Liao, X., Sun, J., Zhang, S. Rockfall investigation and hazard assessment from nang county to jiacha county in tibet. Applied Sciences. 10(1), 247. https://doi.org/10.3390/app10010247, 2020.

Losasso, L., Jaboyedoff, M., and Sdao, F.: Potential rock fall source areas identification and rock fall propagation in the province of Potenza territory using an empirically distributed approach, Landslides, 14, 1593-1602, https://doi.org/10.1007/s10346-017-0807-x, 2017.

Loye, A., Jaboyedoff, M., Pedrazzini, A. Identification of potential rockfall source areas at a regional scale using a DEM-based geomorphometric analysis. Nat. Hazard. Earth Sys. 9(5), 1643-1653. https://doi.org/10.5194/nhess-9-1643-2009, 2009.

Malamud, B. D., Turcotte, D. L., Guzzetti, F., and Reichenbach, P.: Landslide inventories and their statistical properties, Earth Surface Processes and Landforms, 29, 687-711, https://doi.org/10.1002/esp.1064, 2004.

Marchelli, M. and De Biagi, V.: Optimization methods for the evaluation of the parameters of a rockfall fractal fragmentation model, Landslides, 16, 1385-1396, https://doi.org/10.1007/s10346-019-01182-y, 2019.

Matasci, B., Jaboyedoff, M., Loye, A., Pedrazzini, A., Derron, M. H., and Pedrozzi, G.: Impacts of fracturing patterns on the rockfall susceptibility and erosion rate of stratified limestone, Geomorphology, 241, 83-97, https://doi.org/10.1016/j.geomorph.2015.03.037, 2015.

Michoud, C., Derron, M. H., Horton, P., Jaboyedoff, M., Baillifard, F. J., Loye, A., Nicolet, P., Pedrazzini, A., and Queyrel, A.: Rockfall hazard and risk assessments along roads at a regional scale: example in Swiss Alps, Natural Hazards and Earth System

Sciences, 12, 615-629. https://doi.org/10.5194/nhess-12-615-2012, 2012.

Mitchell, A., Mcdougall, S., Nolde, N., Brideau, M.A., Whittall, J., Aaron, J.B. Rock avalanche runout prediction using stochastic analysis of a regional dataset. Landslides. 17(4), 777-792. https://doi.org/ 10.1007/s10346-019-01331-3, 2020.

Pelletier, J.D., Malamud, B.D., Blodgett, T., Turcotte, D.L. Scale-invariance of soil moisture variability and its implications for the frequency-size distribution of landslides. Eng. Geol. 48(3), 255-268. https://doi.org/10.1016/S0013-7952(97)00041-0, 1997.

Pierson LA, Vickle RV. Rock-fall hazard rating system participants' manual. U.S. Department of Transportation, Publication No．FHWA SA-93-057, 11, 1-99. https://vulcanhammernet.files.wordpress.com/ 2017/01/fhwa-sa-93-057.pdf, 1993.

Provost, F., Hibert, C., and Malet, J. P.: Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier, Geophysical Research Letters, 44, 113-120, https://doi.org/10.1002/2016gl070709, 2017.

Rissanen, J. J.: Modeling by the shortest data description, Automatica-J.IF AC, 14, 465–471, https://doi.org/10.1016/0005-1098(78)90005-5, 1978.

Vitanyi, P. M. B., Li, M.: Minimum description length induction, Bayesianism, and Kolmogorov complexity, IEEE Transactions on Information Theory, 46, no. 2, 446-464, https://doi.org/10.1109/18.825807 , 2000.

Wieczorek, G.F.M.M. Rockfall Hazards in the Yosemite Valley. U.S. Geological Survey. https://doi.org/10.3133/ofr98467, 1998.

Zheng, L., Chen, G., Li, Y., Zhang, Y., and Kasama, K.: The slope modeling method with GIS support for rockfall analysis using 3D DDA, Geomechanics and Geoengineering, 9, 142-152, 10.1080/17486025.2013.871070, 2014.

Žabota, B., Repe, B., and Kobal, M.: Influence of digital elevation model resolution on rockfall modelling, Geomorphology, 328, 183-195, https://doi.org/10.1016/j.geomorph.2018.12.029, 2019.

Zezere, J. L.; Pereira, S.; Melo, R.; Oliveira, S. C.; Garcia, R. A. C. Mapping landslide susceptibility using data-driven methods. Sci. Total Environ. 589, 250–267. https://doi.org/10.1016/j.scitotenv.2017.02.188. , 2017.