Machine Learning-Based Evaluation of Susceptibility to Geological Hazards in the Hengduan Mountains Region, China

Jiaqi Zhao\textsuperscript{1,2,3} · Qiang Zhang\textsuperscript{1,2,3} · Danzhou Wang\textsuperscript{1,2,3} · Wenhuan Wu\textsuperscript{1,2,3} · Ruyue Yuan\textsuperscript{1,2,3}

Received: 11 February 2022 / Published online: 14 April 2022 © The Author(s) 2022

Abstract The Hengduan Mountains Region (HMR) is one of the areas that experience the most frequent geological hazards in China. However, few reports are available that address the geological hazard susceptibility of the region. This study developed six machine learning models to assess the geological hazard susceptibility. The results show that areas with medium and high susceptibility to geological hazards as a whole account for almost 21% of the total area, while both are 18% when it comes to the single hazard of landslide and rockfall respectively. Medium and high geological hazard susceptibility is found in three parts of the HMR with different characteristics: (1) the central and southern parts, where the population of the region concentrates; (2) the northern part, where higher geological hazard susceptibility is found along the mountain ranges; and (3) the junction of Tibet, Yunnan, and Sichuan in the eastern part, which is prone to larger-scale geological hazards. Of all the potential influencing factors, topographic features and climatic variables act as the major driving factors behind geological hazards and elevation, slope, and precipitation are crucial indicators for geological hazard susceptibility assessment. This study developed the geological hazard susceptibility maps of the HMR and provided information for the multi-hazard risk assessment and management of the region.

Keywords Geological hazards · Landslides · Machine learning techniques · Rockfalls · Susceptibility evaluation

1 Introduction

Geological hazards refer to geodynamic activity or phenomena that are formed under the interactions of natural or human factors, which cause losses to human life and property and damage to the environment, such as landslides, rockfalls, and so on (Huang et al. 2020; Lin et al. 2021). Besides, warming temperatures, and particularly rapid winter warming episodes, accelerate the melting process of the permafrost (Zhang et al. 2019), which may also intensify geological hazards such as landslides. Therefore, geological hazard susceptibility assessment has become a challenging task and also an urgent demand.

Geological hazard susceptibility is recognized as the likelihood of hazard occurrence in a given location (Fell et al. 2008; Saha et al. 2020; Samia et al. 2020) and is closely related to the characteristics of the geographical elements (Dahal et al. 2008a, 2008b; Pourghasemi et al. 2012). Because of the complicated interactions between the various driving factors of susceptibility, it is also a challenging task to quantify the susceptibility of geological hazards (Fell et al. 2008). Although building an indicator system and assigning weights to calculate an evaluation
index has become the most widely used method for risk and susceptibility assessment, it is often criticized for its subjectivity (Hadmoko et al. 2010; Hadji et al. 2013; Niu et al. 2015; Huang et al. 2020). The development of machine learning technologies provides alternative solutions for this issue. By analyzing the inherent geographic characteristics of potential hazard-affected sites (for example, topography, elevation, lithology, precipitation, and so on), machine learning models can detect hazards within a certain range and predict their distribution. Among all the machine learning algorithms, the random forest (RF), support vector machine (SVM), and backpropagation (BP) neural network models have been widely used in the field of disaster modeling (Lee et al. 2007; Chen et al. 2009; Yilmaz 2009; Bednarik et al. 2012; Catani et al. 2013; Huang et al. 2020) because of their high prediction performance (Cutler et al. 2007). Thus, by comparing the results of the three algorithms with the receiver operating characteristic curve (ROC), we can select the best regional model for a susceptibility assessment.

The Hengduan Mountains Region (HMR) covers an area of more than 600,000 km² characterized overall by a relatively dry climate, which is varied spatially and seasonally by zones of temperate monsoon, subtropical monsoon, and alpine plateau climates. The complex topographic features and climate conditions in the HMR render this region extremely susceptible to large-scale geological hazards (Huang 2009; Li et al. 2011; Liu et al. 2013; Huang et al. 2014; Xu et al. 2019). Although single hazard analysis in isolation is increasingly difficult to simultaneously meet the demand of a full evaluation of risks and the risk reduction of natural hazards (Peduzzi et al. 2009; Koks et al. 2019; Pourghasemi et al. 2020; Pouyan et al. 2021), few studies are available that address this issue. This is especially true in the HMR, which is not conducive to hazard by hazard land use planning and disaster risk management in this complex, multi-hazard region. Therefore, evaluation of the susceptibility of the HMR to geological hazards is challenging but also constitutes the most significant potential contribution of this current study.

The main objective of this study is to evaluate the susceptibility of the HMR in China (Fig. 1) to geological hazards, which only include two types of hazards of landslides and rockfalls in this study. They have been selected as representative hazards because they have high occurrence frequencies, large disaster-affected areas, complex driving factors, and massive damages (Huang 2009). Six models were developed from three machine learning algorithms. Then we compared the modeling performance of these algorithms and screened out the most effective one to identify the regions with high susceptibility to geological hazards and map the susceptibility of the HMR to geological hazards. This study provides a theoretical framework for the evaluation of the susceptibility of mountainous regions to multiple geological hazards.

2 Data

The data we analyzed in this study include geological hazard data with longitude and latitude of the hazard sites (“hazard points” in the subsequent sections) and driving factors of geological hazards. The hazard points were used for training the machine learning models, while the driving factors were used to implement susceptibility assessment.

2.1 Geological Hazard Inventory

Through on-the-spot visits, we obtained the hazard inventory from the Natural Resources Bureaus of 11 cities in Sichuan Province, Tibet Autonomous Region, and Yunnan Province. This inventory records the location and scale of geological hazards, covering 9980 hazard points, wherein 7159 hazard points are for landslides and 2821 for rockfalls. Although these hazard points are not all located in the HMR but are distributed in the HMR and its immediately surrounding areas, they are all related to similar driving factors. Therefore, these records were incorporated to form our hazard inventory (Fig. 2).

To get better results by training the machine learning models, 2080 non-rockfall points and 5500 non-landslide points within the HMR were randomly selected while avoiding a 500-meter buffer zone along the rivers and roads, and around existing hazard points in the inventory. By adding them to the hazard inventory, the number of non-hazard points and hazard points is basically the same.

2.2 Geological Hazard Driving Factors

Examples of existing research that address the driving factors behind geological hazard assessment, which we consulted, include Lee and Min (2001), Süzen and Doyuran (2004), Floris et al. (2011), Erener and Düzgün (2013), Mandal and Maiti (2015), Shahabi and Hashim (2015), Strehmel et al. (2015), and Hou et al. (2016). In this study, we adopted 10 factors for geological hazard susceptibility evaluation, including elevation, slope, aspect, curvature, lithology, distance from geological fault, precipitation, Normalized Difference Vegetation Index (NDVI), distance from the river, and distance from road (Table 1), of which the first eight were used for model training and assessment of susceptibility to geological hazards. The latter two were mainly used to identify nongeological hazard points.

Our digital elevation model (DEM) is sourced from the ASTER GDEM data set jointly developed by the National
Aeronautics and Space Administration (NASA) of the United States, the Japanese Ministry of Economy, Trade and Industry (METI), and the Japan Space Agency, with a spatial resolution of 30 m. The slope, aspect, and curvature data were all processed from the DEM using the spatial analysis toolbox of ArcGIS. Lithology and the distance from the geological faults demonstrate geological structure. Hydrometeorological indicators include precipitation, NDVI, and distance from rivers, and the human activity indicator refers to the distance from roads. The precipitation data are based on a set of reanalysis data sets released by the National Qinghai-Tibet Plateau Science Data Center, the data set covers the period of 1979–2015 with a spatial resolution of 0.1°. We calculated the average precipitation using the R software package and interpolated the data to 500 m. The distribution of rivers and roads is sourced from OpenStreetMap. We removed artificial irrigation canals and country roads but kept the main roads for the analysis.

3 Methods

The scale of the geological hazards, such as the size and/or volume of the rockfalls and landslides, was taken as the basis for the susceptibility assessment and the susceptibility to the geological hazards was subdivided into 4 levels (Table 2) as suggested by the Specification of Comprehensive Survey for Landslide, Rockfall and Debris Flow (1:50,000) issued by China’s Ministry of Land and Resources (China 2014). About 70% of the hazard points were randomly screened out from each level to form the training data set, while the rest form the validation data set. Six models were developed from three types of machine learning methods, that is, random forest (RF) model, support vector machine (SVM) model with different nuclear kernels, and the gradient boosting model (GBM).
functions, and backpropagation (BP) neural network model.

### 3.1 Random Forest (RF) Model

The random forest model is an ensemble machine learning model and was developed from decision trees. Modeling based on the RF model involves four steps (Huang et al. 2020):

1. Take \( k \) samples from the training sample \( D \) to build a random tree.
2. Randomly extract \( m \) features from each sample, and select the best feature as the growth node.
3. Let each tree grow as much as possible without pruning.
4. Repeat the above steps continuously to form a random forest.

In addition, in the modeling by the RF model, the mean decrease Gini and the mean decrease accuracy index can be used to quantify the importance of each influencing factor.

### 3.2 Support Vector Machine (SVM) Model

The SVM is a linear classifier, which maximizes the difference between categories of different features by projecting the feature points into the multi-dimensional feature space so as to find the optimal \( n \)-dimensional hyperplane to classify the data, and therefore transforms the classification problem into a solution to a convex secondary planning problem (Eqs. 1-2) (Yao et al. 2008; Xu et al. 2012; Pourghasemi et al. 2013). In essence, the SVM model is a two-classification model and can also realize the multi-classification of data.

\[
y_i(w^T x_i + b) \geq 1
\]

\[
\max \frac{1}{||w||^2}
\]

One of the important parameters in the SVM model is the kernel function, which mainly includes: linear, polynomial, radial (RBF), and sigmoid (Eqs. 3-6). We developed four kinds of SVM models according to different kernel functions for the evaluation of the susceptibility to geological hazards.

### 3.3 Back Propagation (BP) Neural Network Model

The BP neural network model is a multilayer feedforward network that incorporates an error backpropagation algorithm (Ying et al. 2017). The model includes a three-layer structure: namely an input layer, a hidden layer, and an output layer. Each structure contains multiple nodes, and the hidden layer can be set with multiple layers of nodes. Different nodes are connected by lines and assigned

---

**Table 1** Driving factors selection and data source

| Class            | Indicator          | Data source                                      |
|------------------|--------------------|--------------------------------------------------|
| Topography       | DEM                | ASTER GDEM data set (https://earthexplorer.usgs.gov/) |
| Lostorege structure | Slope          | OpenStreetMap (https://www.openstreetmap.org.)       |
|                 | Aspect             | River map GIS software                           |
| Distance from river | Distance from the geological faults | OpenStreetMap (https://www.openstreetmap.org.) |
| Hydrometeorology | Precipitation      | National Qinghai-Tibet Plateau Science Data Center (doi: https://doi.org/10.3972/westdc.002.2014.db.) |
|                 | NDVI               | OpenStreetMap (https://www.openstreetmap.org.) |

**Table 2** Criteria for the classification of susceptibility to geological hazards

| Susceptibility level | Low | Medium | High | Very high |
|----------------------|-----|--------|------|-----------|
| Scale of rockfall \(10^4\text{m}^3\) | 0   | < 1    | 1–10 | > 10      |
| Scale of landslide \(10^4\text{m}^3\) | 0   | < 10   | 10–100 | > 100   |

---
weights. Compared with ordinary neural network algorithms, the BP neural network model is characterized by the backpropagation of errors. The calculation process mainly includes forward calculation and error adjustment. The main process is as follows (Wang et al. 2005):

The output of hidden layer node:

$$y_h^k = f \left( \sum_{i=1}^{N_1} \omega_{ih} \cdot x_i^k + \theta_h \right)$$

The output of output layer node:

$$z_j^k = f \left( \sum_{h=1}^{N_2} \omega_{hj} \cdot y_h^k + \gamma_j \right)$$

Error function:

$$E = \frac{1}{2} \sum_{k,j} \left( T_j^k - z_j^k \right)^2$$

The adjustment of weight:

$$\Delta \omega = -\eta \frac{\partial E}{\partial \omega}$$

Weight adjustment:

$$\omega = \omega + \Delta \omega$$

where $\theta_h$ and $\gamma_j$ are the thresholds of the hidden layer node and the output layer node, respectively; $f$ is the transfer function; $\eta$ is the learning step; and $P$ is the number of samples, $k = 1, \ldots, P$.

### 3.4 Model Validation

The predictive accuracy of the machine learning models is evaluated by the receiver operating characteristic (ROC) curves and area under the curve (AUC) values (Gorsevski et al. 2006; Trigila et al. 2015). The accuracy of model classification consists of four types of results: true positive class (TP), false positive class (FP), false negative class (FN), and true negative class (TN) (Table 3). The true positive class and the false negative class represent the correct classification.

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

### 4 Results

In this section, we first present the ROC curves of the six models to show their accuracy and choose the most suitable model for further application; then the susceptibility to landslides, rockfalls, and geological hazards as a whole is assessed, and the susceptibility is mapped.

#### 4.1 Verification of the Model Accuracy

Figure 3 illustrates the ROC curves and AUC values of the rockfall and landslide hazards respectively. Overall, for the susceptibility to the rockfall hazard, the SVM model with sigmoid kernel function has a poor modeling performance, while the training results of other models are satisfactory with AUC above 0.8. Wherein, the RF model obtains the highest AUC of 0.88. As for the susceptibility to landslides, most models have AUC values between 0.7 and 0.8, of which the SVM model with sigmoid kernel function performs the worst, with an AUC value of 0.59, while the RF model performs the best with an AUC value of 0.83. In general, the RF model is the best choice for susceptibility assessment to the geological hazards.

#### 4.2 Evaluation of Susceptibility to Rockfall and Landslide

Figures 4 and 5 demonstrate the spatial pattern of the susceptibility to rockfall and landslide respectively. To show the connections between altitude and slope and geological hazards, we overlay geological hazard points with our DEM (Digital Elevation Model). It is evident that the hazard points are mainly distributed along the valleys, especially the Jinsha River valley where we observe a concentration of multiple large-scale landslide points. River valleys are usually characterized by steep slopes vulnerable to erosion and erosion process by river flow also cause instability of riverbeds; all of these factors can trigger higher occurrences of geological hazards such as rockfalls and landslides. We can observe a more dispersed distribution of large-scale landslides in the southern and central areas of Sichuan Province. We also mapped the susceptibility to rockfall and landslide respectively (Figs. 4a, 5a). In general, medium and high susceptibility to landslide and rockfall is identified mainly across the central and southern HMR and is specifically distributed along the mountains.

With regard to susceptibility to landslides, areas with a very high susceptibility account for 0.18% of the study
Specifically, the areas with very high susceptibility to landslides are concentrated in the eastern parts of the Tibet Autonomous Region and are distributed along the Lancang River and the Jinsha River. The areas dominated by very high susceptibility to landslides are found near geological faults and these areas usually have a highly instable geological structure and are susceptible to geological hazards. Meanwhile, the areas across the Three Parallel Rivers (Nujiang River, Lancang River, and Jinsha River) region are dominated by distinct topographic relief, steep slopes, and low vegetation coverage, all of which are factors conducive to the occurrence of large-scale...
landslides (Huang 2009). The areas with high and medium susceptibility to landslides account for 3.27% and 14.28% respectively of the study region as a whole (Table 4). These areas are identified mainly in the central parts of the study region.

With regard to the susceptibility to rockfall, areas with very high susceptibility account for 0.02% of the study region and show dispersed spatial distribution. Areas with high susceptibility to rockfall account for 3.73% of the study region; these areas are identified mainly in the northeastern and central parts of the study region. Regions with medium susceptibility to rockfall account for 14.29% of the study region and these areas are found mainly in the central parts of the HMR. In order to further verify the reliability of the results, we compared the location of hazard points with that of the susceptibility to rockfall (Figs. 4c, 5c). We found a high degree of consistency between the location of hazard points and that of high susceptibility to rockfall, indicating the reliability of the susceptibility evaluation results.

4.3 Evaluation of Susceptibility to Geological Hazards as a Whole

By taking the average value of the susceptibility to landslides and rockfalls and grading with natural breakpoint method, we obtained the map of susceptibility to geological hazards as a whole across the study region (Fig. 6a). When compared with the mapping results of the susceptibility to geological hazards as a whole can comprehensively show the characteristics of susceptibility to multiple hazards. These regions display a distinctly different spatial pattern. In addition, nine types of

![Fig. 5](image_url)

**Table 4** Proportion of susceptibility area

| Disaster type                      | Low (%) | Medium (%) | High (%) | Very high (%) |
|-----------------------------------|---------|------------|----------|---------------|
| Rockfall                          | 81.96   | 14.29      | 3.73     | 0.02          |
| Landslide                         | 82.28   | 14.28      | 3.27     | 0.18          |
| Geological hazard as a whole      | 75.03   | 11.96      | 8.67     | 4.34          |
geological hazard susceptibility areas are delineated according to the susceptibility to landslides and rockfalls (Fig. 6b). Comparison of these two maps can benefit in-depth analysis of the characteristics of susceptibility to geological hazards. Areas with very high susceptibility to geological hazards account for 4.34% of the study region (Table 4), which is mainly distributed in the Three Parallel Rivers region, the central and the northeastern parts of the study area. Those areas with high geological hazard susceptibility account for 8.67% of the study region (Table 4), which are distributed in the central parts of the study region and the river valleys in the northeastern parts of the study region. Other areas with moderate susceptibility to the geological hazards account for 11.96% of the study region (Table 4), which are found mainly in the central parts of the study region and the Three Parallel Rivers region. The map also demonstrates that the spatial patterns of very high and high susceptibility to geological hazards are similar, indicating that large-scale landslides are more likely to occur in these areas. In addition, areas with high susceptibility to geological hazards are identical to those with moderate susceptibility to landslides and rockfalls, indicating the higher frequency of moderate-scale landslides and rockfalls in these areas. On the whole, in the central part of HMR and the Three Parallel Rivers region, massive areas with high landslide susceptibility and large areas with moderate susceptibility to rockfall are found, resulting in extremely high susceptibility to geological hazards and a high degree of risk exposure to geological hazards. These areas are also home to a large and high density population as well as a relatively high socioeconomic development level. In this sense, remarkable concerns should be attached to the human mitigation of geological disasters and effort to achieve geological disaster risk reduction.

5 Discussions

In order to further investigate the characteristics of geological hazard susceptibility, we calculated and analyzed the driving factors behind susceptibility. Based on the mean decrease Gini and mean decrease accuracy indicators by the RF model, we calculated the contribution rate of each factor to geological hazard susceptibility (Fig. 7) in order to differentiate major driving factors. The mean decrease Gini is determined by calculating the influence of each variable on the heterogeneity of the observed susceptibility of geological hazards at each node of the classification tree; the mean decrease accuracy is determined by replacing the variable with a random number and comparing the degree of decline in the accuracy of the model to determine the importance of the factor to the susceptibility to geological hazards. In the landslide susceptibility assessment model, the results obtained by the two indicators are highly consistent. From the perspective of the mean decrease accuracy, driving factors behind the landslide susceptibility are elevation, curvature, precipitation, and distance from a fault. The CCR (cumulative
contribution rate of the foremost four indicators) reaches as high as 75.16%, which means that topography, geological aspects, and climatic conditions have an important impact on the occurrence of landslide hazards. According to the mean decrease Gini, elevation has the highest contribution rate to landslide hazards, reaching as high as 23.49%. The four indicators with the highest contribution rate are elevation, precipitation, curvature, and slope, with a CCR of 66.28%. Among these factors, precipitation provides sufficient hydrodynamic conditions to trigger the occurrence of landslides and rockfall. In the rockfall susceptibility model, the four indicators with the highest contribution rate are precipitation, elevation, curvature, and slope. From the mean decrease accuracy, the CCR reaches 76.04%, while from the mean decrease Gini it is 71.2%. To sum up, elevation, slope, and precipitation are crucial factors for the susceptibility assessment of geological hazards such as landslides and rockfalls, which can provide a reference for the selection of indicators in geological hazard susceptibility evaluation.

In this geological hazard susceptibility assessment, hazard points data collection is a basic and important work. The hazard points used in this study are sourced from statistical data provided by the local administrative units of each county and city, which records the occurrence of hazards within the jurisdiction of the county or city. But omissions of geological hazard events may exist. The China Institute for Geological Environmental Monitoring (CIGEM) has compiled a map of historical hazard points of landslides and rockfall, which shows a hazard distribution that is similar to our hazard susceptibility map (CIGEM 2011). Since geological hazard susceptibility is closely related to the historical record of hazard points, the results of the machine learning methods for hazard prediction are reasonably accurate.

Other than providing post-disaster relief, how to reduce exposure to landslides and rockfalls, especially in mountainous areas, has become a growing concern in recent years. With climate change, the problem would be even more challenging. To put this into perspective, implementing hazard susceptibility assessment and risk assessment before planning to effectively reduce the irrationality of construction and development is the key to tackle this challenge (Qiu 2014). For example, Sidle and Ziegler (2012) claimed that land management may pose potential problems without adequate road network planning and landslide hazard assessment. In the developing regions of Southeast Asia, there is an urgent demand to make full preparations to avoid potential disasters (Sidle and Ziegler 2012).

Chen et al. (2019) mapped the hazard-prone areas of landslide according to hazard frequency, which shows a similar distribution as does our landslide susceptibility map in the HMR area, but we further mapped the propensity for hazards of different scales. The multi-hazard map shows that the high landslide-high rockfall susceptibility areas are distributed in the northwest of Yunnan Province. This distribution can be partly explained by the huge elevation

---

**Fig. 7** Contribution rate of impact factors on landslide and rockfall susceptibility. **a** Mean decrease accuracy and **b** mean decrease Gini of landslide susceptibility model. **c** Mean decrease accuracy and **d** mean decrease Gini of rockfall susceptibility model. CCR is cumulative contribution rate of the foremost four indicators.
difference and intensive road construction in the area. Landslide rates in Southeast Asia sharply increased when development that used improper road design took place in steep-slope mountainous areas, especially in Yunnan, during the decade from 1970 to 1980. The landslides associated with roads were hundreds of times more frequent than those in the mountainous Pacific Northwest of the United States (Sidle and Ochiai 2006; Sidle et al. 2011). This is also an indirect evidence that supports our results.

Susceptibility assessment is also an essential part of risk assessment. Lin et al. (2017) assessed the empirical dependencies between the volume of landslides and their resulting casualties, and found that when the landslide volume is $0.1 - 50 \times 10^4 m^3$, the number of average casualties rise to approximately 10–20. Casualties increase even higher as larger volume landslides occur, indicating that the hazard scale is closely associated with loss of lives and economic losses. Susceptibility assessment is of great importance as a result, and further risk assessment studies need to be carried out in the future.

6 Conclusion and Findings

In this study, we adopted 10 factors in the assessment of susceptibility to the geological hazards of the Hengduan Mountains Region. Areas with different multi-hazard susceptibilities are identified, which can help to provide reference information for the risk assessment of geological hazards in the study region. Improved understanding of susceptibility to geological hazards can be conducive to the mitigation of geological disasters and sustainable development. We obtained the following findings:

(1) Based on the RF model, the SVM model, and the BP neural network model, we developed and compared six machine learning models to screen out the optimal model for the evaluation of susceptibility to geological hazards. By comparing the prediction effects of these models, we found that the RF model has the highest modeling/prediction performance when compared to the other models considered in this study. The RF model also has the advantage of strong resistance to over-fitting (with an AUC value of 0.83) compared to the other models considered in this study. This finding can help to provide reference information for the selection of models in the evaluation of susceptibility to geological hazards in other regions of the globe.

(2) We identified that areas with medium and high susceptibility to landslides and rockfalls are distributed in the central and southern parts of the Hengduan Mountains Region, and are distributed along the mountains in the northern parts of the HMR. The areas with very high susceptibility to landslides are found mainly in the eastern parts of the Tibet Autonomous Region, and along the Lancang River and the Jinsha River. The areas with very high susceptibility to rockfall have a sporadic spatial distribution. The result of prediction of the susceptibility to geological hazards is similar to the actual distribution of the geological hazard points in space. This demonstrates the reliable prediction performance of the machine learning model developed in this study.

(3) The Three Parallel River region in the eastern part of Tibet is highly susceptible to geological hazards. Meanwhile, the areas with medium and high susceptibility to geological hazards are widespread in the middle part of the Hengduan Mountains Region. These areas should be given considerable attention since the HMR is dominated by a relatively high population density and a developed economy. Mitigation of geological hazard risks and scientific planning of human activities should be done based on the evaluation results of the susceptibility to geological hazards in the HMR. Moreover, the susceptibility assessment shows that topographic and climatic factors play a critical role in the occurrence of geological hazards.

Acknowledgements This research has been financially supported by the China National Key R&D Program, Grant No. 2019YFA0606900 and by the Scientific Investigation of the Qinghai-Tibet Plateau, Grant No. 2019QZKK0900. We would like to extend our cordial gratitude to the editors, Dr. Ying Li, Dr. Juan Du and the anonymous reviewers for their pertinent and professional comments and suggestions that were greatly helpful for further quality improvement of this manuscript.

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

Bednarik, M., I. Yilmaz, and M. Marschalko. 2012. Landslide hazard and risk assessment: A case study from the Hlohovec–Sered’ landslide area in south-west Slovakia. Natural Hazards 64(1): 547–575.

Catani, F., D. Lagomarsino, S. Segoni, and V. Tofani. 2013. Landslide susceptibility estimation by random forests technique:
Sensitivity and scaling issues. Natural Hazards and Earth System Sciences 13(11): 2815–2831.

Chen, Y., W. Xie, and X. Xu. 2019. Changes of population, built-up land, and cropland exposure to natural hazards in china from 1995 to 2015. International Journal of Disaster Risk Science 10(4): 557–572.

Chen, J., S. Zhao, W. Liao, and W. Yuan. 2009. Research on natural disaster risk assessment model based on Support Vector Machine and it’s application. In Neural information processing. ICONIP 2009, ed. C.S. Leung, M. Lee, and J.H. Chan, 762–769. Heidelberg: Springer.

China, People’s Republic of. Ministry of Land and Resources. 2014. Specification of comprehensive survey for landslide, collapse and debris flow (1: 500000). Beijing: Standards Press of China. http://www.nssi.org.cn/nssi/front/87543223.html.

CIGEM (China Institute of Geological Environmental Monitoring). 2011. Rockfall and landslide disaster map of China. Beijing: SinoMaps Press.

Cutler, D.R., T.C. Edwards, K.H. Beard, A. Cutler, K.T. Hess, J.C. Gibson, and J.J. Lawler. 2007. Random forests for classification in ecology. Ecology 88(11): 2783–2792.

Dahal, R.K., S. Hasegawa, A. Nonomura, M. Yamanaka, S. Dhakal, and P. Paudyal. 2008. Predictive modelling of rainfall-induced landslide hazard in the Lesser Himalaya of Nepal based on weights-of-evidence. Geomorphology 102(3–4): 496–510.

Dahal, R.K., S. Hasegawa, A. Nonomura, M. Yamanaka, T. Masuda, and K. Nishino. 2008. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. Environmental Geology 54(2): 314–324.

Erener, A., and H.B.S. Düzgün. 2013. A regional scale quantitative risk assessment for landslides: Case of Kumluca watershed in Bartin, Turkey. Landslides 10(1): 55–73.

Fell, R., J. Corominas, C. Bonnard, L. Cascini, E. Leroi, W.Z. Savage on behalf of the JTC-1 Joint Technical Committee on Landslides and Engineered Slopes. 2008. Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning. Engineering Geology 102: 99–111.

Floris, M., M. Iafelice, C. Squarzoni, L. Zorzi, A. De Agostini, and R. Genevois. 2011. Using online databases for landslide susceptibility assessment: An example from the Veneto Region (northeastern Italy). Natural Hazards and Earth System Sciences 11(7): 1915–1925.

Gorski, P.V., P.E. Gessler, R.B. Foltz, and W.J. Elliot. 2006. Spatial prediction of landslide hazard using logistic regression and ROC analysis. Transactions in GIS 10(3): 395–415.

Hadji, R., A.E. Boumazbeur, Y. Limani, M.A. Baghem, A.M. Chouabi, and A. Demdoum. 2013. Geologic, topographic and climatic controls in landslide hazard assessment using GIS modeling: A case study of Souk Ahras region, NE Algeria. Quaternary International 302: 224–237.

Hadmoko, D.S., F. Lavigne, J. Sartohadi, P. Hadi, and Winaryo. 2010. Landslide hazard and risk assessment and their application in risk management and landuse planning in eastern flank of Menoreh Mountains, Yogyakarta Province, Indonesia. Natural Hazards 54(3): 623–642.

Hou, J., J. Lv, X. Chen, and S. Yu. 2016. China’s regional social vulnerability to geological disasters: Evaluation and spatial characteristics analysis. Natural Hazards 84(1): 97–111.

Huang, R. 2009. Some catastrophic landslides since the twentieth century in the southwest of China. Landslides 6(1): 69–81.

Huang, J., B. Di, J. Bian, J. Zuo, and X. Hu. 2014. Analysis on the spatial distribution characteristics and human driving forces of mountain disasters in Liangshan Prefecture. Soil And Water Conservation Research 21(6): 278–283.

Huang, P., L. Peng, and H. Pan. 2020. Linking the Random Forests Model and GIS to assess geo-hazards risk: A case study in Shifang County, China. IEEE Access 8: 28033–28042.

Koks, E.E., J. Rozenberg, C. Zorn, M. Tariverdii, M. Vousdoukas, S.A. Fraser, J.W. Hall, and S. Hallegatte. 2019. A global multi-hazard risk analysis of road and railway infrastructure assets. Nature Communications 10(1): Article 2677.

Lee, S., and K. Min. 2001. Statistical analysis of landslide susceptibility at Yongin, Korea. Environmental Geology 40(9): 1095–1113.

Lee, S., J.H. Ryu, and I.S. Kim. 2007. Landslide susceptibility analysis and its verification using likelihood ratio, logistic regression, and artificial neural network models: Case study of Younign, Korea. Landslides 4(4): 327–338.

Li, Z., Y. He, C. Wang, X. Wang, H. Xin, W. Zhang, and W. Cao. 2011. Spatial and temporal trends of temperature and precipitation during 1960–2008 at the Hengduan mountains. Quaternary International 236(1): 127–142.

Lin, J., W. Chen, X. Qi, and H. Hou. 2021. Risk assessment and its influencing factors analysis of geological hazards in typical mountain environment. Journal of Cleaner Production 309(9): Article 127077.

Lin, Q., W. Ying, T. Liu, Y. Zhu, and Q. Sui. 2017. The vulnerability of people to landslides: A case study on the relationship between the casualties and volume of landslides in China. International Journal of Environmental Research and Public Health 14(2): Article 212.

Liu, C., W. Li, H. Wu, P. Lu, K. Sang, W. Sun, W. Chen, Y. Hong, and R. Li. 2013. Susceptibility evaluation and mapping of China’s landslides based on multi-source data. Natural Hazards 69(3): 1477–1495.

Mandal, S., and R. Maiti. 2015. Application of analytical hierarchy process (AHP) and frequency ratio (FR) model in assessing landslide susceptibility and risk. In Semi-quantitative approaches for landslide assessment and prediction, ed. S. Mandal, and R. Maiti, 191–226. Heidelberg: Springer.

Niu, C., Q. Wang, J. Chen, W. Zhang, L. Xu, and K. Wang. 2015. Hazard assessment of debris flows in the reservoir region of Wudongde hydropower station in China. Sustainability 7(11): 15099–15118.

Pezzuti, P., H. Dao, C. Herold, and F. Mouton. 2009. Assessing global exposure and vulnerability towards natural hazards: The disaster risk index. Natural Hazards & Earth System Sciences 9(4): 1149–1159.

Pourghasemi, H.R., A.G. Jirandeh, B. Pradhan, C. Xu, and C. Gokeeoglu. 2013. Landslide susceptibility mapping using support vector machine and GIS. Journal of Earth System Science 61(15): 349–369.

Pourghasemi, H.R., N. Kariminejad, M. Amirii, M. Edalat, M. Zarafshar, T. Bласчки, and A. Cerda. 2020. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. Scientific Reports 10: Article 3203.

Pourghasemi, H.R., B. Pradhan, and C. Gokeeoglu. 2012. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. Natural Hazards 63(2): 965–996.

Pouyan, S., H.R. Pourghasemi, M. Bordbar, S. Rahmanian, and J.J. Cllauge. 2021. A multi-hazard map-based flooding, gully erosion, forest fires, and earthquakes in Iran. Scientific Reports 11(1): Article 14889.

Qiu, J. 2014. Landslide risks rise up agenda. Nature 511(7509): 272–273.

Saha, S., A. Arabameri, A. Saha, T. Blaschke, and S.S. Band. 2020. Prediction of landslide susceptibility in Rudraprayag, India using novel ensemble of conditional probability and boosted regression
tree-based on cross-validation method. *Science of the Total Environment* 764: Article 142928.

Samia, J., A. Temme, A. Bregt, J. Wallinga, and F. Ardizzone. 2020. Dynamic path dependent landslide susceptibility modelling. *Natural Hazards and Earth System Sciences* 20(1): 271–285.

Shahabi, H., and M. Hashim. 2015. Landslide susceptibility mapping using GIS-based statistical models and remote sensing data in tropical environment. *Scientific Reports* 5: Article 9899.

Sidle, R.C., and H. Ochiai. 2006. *Landslides: Processes, prediction and land use*. Washington, DC: American Geophysical Union.

Sidle, R.C., T. Furuichi, and Y. Kono. 2011. Unprecedented rates of landslide and surface erosion along a newly constructed road in Yunnan, China. *Natural Hazards* 57: 313–326.

Sidle, R.C., and A.D. Ziegler. 2012. The dilemma of mountain roads. *Nature Geoscience* 5(7): 437–438.

Strehmel, A., S. Schnbrodt-Stitt, G. Buzzo, C. Dumperth, and T. Scholten. 2015. Assessment of geo-hazards in a rapidly changing landscape: The Three Gorges Reservoir region in China. *Environmental Earth Sciences* 74(6): 4939–4960.

Süzen, M.L., and V. Doyuran. 2004. A comparison of the GIS based landslide susceptibility assessment methods: Multivariate versus bivariate. *Environmental Geology* 45(5): 665–679.

Trigila, A., C. Iadanza, C. Esposito, and G. Scarascia-Mugnozza. 2015. Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). *Geomorphology* 249: 119–136.

Wang, Y., Y. Guo, and W. Li. 2005. Prediction model based on improved BP neural network and its application. *Computer Measurement and Control* 13(1): 39–42.

Xu, C., F. Dai, X. Xu, and Y.H. Lee. 2012. GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. *Geomorphology* 145–146: 70–80.

Xu, R., X. Li, K. Hu, and Y. Nie. 2019. A dynamic hazard assessment for mountain hazards in Hengduan mountain area. *Journal of Catastrophology* 34(3): 196–201, 208.

Yao, X., L.G. Tham, and F.C. Dai. 2008. Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of Hong Kong, China. *Geomorphology* 101: 572–582.

Yilmaz, I. 2009. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat–Turkey). *Computers & Geosciences* 35(6): 1125–1138.

Ying, Z., W. Hong-Wei, W. Yu-Bing, and G. Peng-Cheng. 2017. A novel optimization algorithm for BP neural network based on RS-MEA. In *Proceedings of 2017 2nd International Conference on Image, Vision and Computing (ICIVC 2017)*, 2–4 June 2017, Chengdu, China. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Zhang, G., Z. Nan, X. Wu, H. Ji, and S. Zhao. 2019. The role of winter warming in permafrost change over the Qinghai-Tibet Plateau. *Geophysical Research Letters* 46(20): 11261–11269.