International Journal of Digital Earth

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjde20

Analytical ‘decisiveness’ as a robust measure of the absolute importance of landslide predisposing factors

Langping Li & Hengxing Lan

To cite this article: Langping Li & Hengxing Lan (2024) Analytical ‘decisiveness’ as a robust measure of the absolute importance of landslide predisposing factors, International Journal of Digital Earth, 17:1, 2356161, DOI: 10.1080/17538947.2024.2356161

To link to this article: https://doi.org/10.1080/17538947.2024.2356161

Published online: 03 Jun 2024.

INTERNATIONAL JOURNAL OF DIGITAL EARTH
2024, VOL. 17, NO. 1, 2356161
https://doi.org/10.1080/17538947.2024.2356161

Analytical 'decisiveness' as a robust measure of the absolute importance of landslide predisposing factors

Langping LI [a,b] and Hengxing LAN [a,c,d]

[a]State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, People's Republic of China; [b]University of Chinese Academy of Sciences, Beijing, People's Republic of China; [c]School of Geological Engineering and Geomatics, Chang'an University, Xi'an, People's Republic of China; [d]Key Laboratory of Ecological Geology and Disaster Prevention of Ministry of Natural Resources, Chang'an University, Xi'an, People's Republic of China

ABSTRACT

Although some methods have been proposed to measure the absolute importance of landslide predisposing factors, they are unfortunately complicated by inherent difficulties. In this study, a robust analytical index – 'decisiveness,' is proposed for measuring factor absolute importance. The idea is to evaluate factor importance according to the capability of differentiating landslide-susceptible and landslide-insusceptible areas; that is, those factors that can zone the target area into units either highly favorable or highly unfavorable to landslides will be of high importance for landslides. For a specific factor, the favorability values of all grid cell units in the target area are first calculated using the certainty factor method. Then, the magnitudes of all those favorability values are integrated to constitute the decisiveness index. The decisiveness ranges within [0, 1], with higher values indicating greater importance. Updated versions of the open software ALSA can be used to calculate favorability values and further decisiveness. The significance and robustness of decisiveness were validated in a case study. Although decisiveness has been proposed to measure the factor importance for landslides, the novel idea and index can also be used to evaluate the importance of factors in determining the spatial distribution of other objects or phenomena.

ARTICLE HISTORY
Received 11 January 2024
Accepted 11 May 2024

KEYWORDS
Landslide; predisposing factor; absolute importance; decisiveness; certainty factor

1. Introduction

Measuring the importance of landslide predisposing factors is fundamental for weighting, selecting, and interpreting predisposing factors in landslide spatial analysis. Many indices have been proposed or used to measure the factor importance (e.g. Kavzoglu and Teke 2022; Ma, Mei, and Piccialli 2021), most of which are derived using data-driven methods. The various factor importance indices can be broadly categorized into two types: 'relative importance' and 'absolute importance' indices. For the relative scenario, the importance index is derived based on

CONTACT Hengxing LAN lanhx@igsnrr.ac.cn State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, People's Republic of China; School of Geological Engineering and Geomatics, Chang'an University, Xi'an, 710054, People's Republic of China; Key Laboratory of Ecological Geology and Disaster Prevention of Ministry of Natural Resources, Chang'an University, Xi'an, 710054, People's Republic of China

comparisons among factors; that is, the importance measurement of any individual factor uses the information of all considered factors. For the absolute scenario, the importance index of a specific factor is derived according to the correlation between landslides and this specific factor only; that is, the importance measurements of factors are independent of each other. The relative index essentially makes comparisons instead of providing absolute measurements of factor importance. The advantage of absolute importance over relative importance is straightforward. First, the absolute importance of a factor can be calculated even if data is available for only that specific factor; however, deriving relative importance requires data for at least two factors. Second, the absolute importance of a factor will change only if the data used for that specific factor changes; whereas, the relative importance of a factor might fluctuate with changes in the data of other factors, further complicating the evaluation of importance. Third, absolute measurement makes it convenient to compare the importance of the same factor in different cases, which is difficult for relative evaluations of factor importance.

The authors conducted a systematic literature survey based on the Web of Science SCIE and SSCI collections. Publications with a topic including the terms 'landslide' AND 'predisposing OR conditioning OR influencing' AND 'factor' AND 'importance' were searched out and surveyed. The details of this literature review will be presented in another contribution, while some fundamental truths are reported herein. Among the searched 334 publications, 84 have measured the importance of landslide predisposing factors, of which ∼1% have actually measured the importance of factor classes (e.g. Donati and Turrini 2002), ∼95% have measured the 'relative importance' of factors, and only ∼4% have measured the 'absolute importance.' The authors have noticed more than 30 different approaches used in measuring factor relative importance, and popular ones include the Analytic Hierarchy Process (AHP) method (e.g. Ma et al. 2013), Random Forests (RF) machine learning method (e.g. He, Wang, and Liu 2021), Information Gain Ratio (IGR) method (e.g. Zhao et al. 2021), Jackknife-based methods (e.g. Gullà, Conforti, and Borrelli 2021), and permutation-based methods (e.g. Wang et al. 2021). However, the authors are aware of only four approaches for measuring factor absolute importance in a quantitative way, i.e. the accountability & reliability method (e.g. Blahut, van Westen, and Sterlacchini 2010), AUC (area under the success rate curve) method (e.g. Barella, Sobreira, and Zêzere 2019), geographical detector method (e.g. Yang et al. 2019), and sensitivity index method (e.g. Wang et al. 2014). The scarcity of measurements of absolute importance is surprising when considering that, intuitively, the absolute importance of a factor should be measured and used instead of the relative importance whenever possible, because it will not be disturbed by uncertainties due to the involvement of other factors.

The recognized four approaches for measuring the absolute importance of a factor, nevertheless, are not impeccable. In the accountability & reliability method (e.g. Blahut, van Westen, and Sterlacchini 2010), susceptible factor classes with a conditional landslide probability greater than the prior landslide probability are first recognized. The 'accountability' is the total landslide cell counts in those susceptible factor classes divided by the total landslide cell counts in the entire study area, and multiplied by 100; and, the 'reliability' is the total landslide cell counts in those susceptible factor classes divided by the total cell counts of those susceptible factor classes, and multiplied by 100 (e.g. Castellanos Abella 2008; Greenbaum et al. 1995). First, the accountability & reliability method, when used for factors with continuous values, requires a prior factor reclassification, introducing additional uncertainties associated with the extremely variable choice of classification numbers and breaks. Second, it computes only the summation of cell counts in those susceptible factor classes and does not consider the cell distribution among them. Third, the use of two indices complicates the evaluation of importance (e.g. Greenbaum et al. 1995). The AUC method (e.g. Barella, Sobreira, and Zêzere 2019) makes landslide predictions directly based on individual predisposing factors, plots the success rate curve for each individual factor, and uses the area under the curve (AUC) as a measure of absolute importance. However, it supposes a monotonic relationship between the factor value and landslide probability, which is not necessarily true, especially for categorized factors. Similarly, any attempt to evaluate factor importance based on a direct correlation analysis between factor values and landslide presence/absence will be problematic because an unnecessarily true monotonic correlation between factor value and landslide probability is implicitly assumed. The geographical detector method (e.g. Wang et al. 2010) evaluates factor importance according to the variance of landslide presence within each factor class. However, a prior factor reclassification is also needed for factors with continuous values, although various approaches have been proposed to optimize factor discretization (e.g. Zhang, Song, and Wu 2022). The sensitivity index method first calculates the certainty factor (CF) values for all factor classes and subtracts the maximum and minimum CF values as a measure of absolute importance (e.g. Wang et al. 2014). One obvious flaw of this method is that it only considers two factor classes with the maximum and minimum CF values; however, the sensibilities to landslides of other factor classes must also contribute to the importance of the factor. Therefore, it is worthwhile to develop a new robust measurement of the absolute importance of landslide predisposing factors.

This technical note aims to propose an analytical index – 'decisiveness,' to measure the absolute importance of landslide predisposing factors in a robust manner. First, the principle and calculation of decisiveness are introduced. Then, the applicability of this novel factor absolute importance index is evaluated and discussed using a case study.

2. Decisiveness

2.1. Principle

The absolute importance index – 'decisiveness' proposed in this paper is derived based on the spatial relationship between historical landslides and individual predisposing factors. The basic idea of decisiveness is that, the importance of a factor can be evaluated according to the ability of this factor to determine landslide occurrence, i.e. the factor's 'power of determinant' (e.g. Wang et al. 2010). Suppose the target area has been rasterized into regular grid cells; the steps for constituting the decisiveness index can be illustrated as follows.

(1) Preparation: Measure favorability for all factor values and assign favorability to all grid cell units. Favorability quantifies the 'degree of favorability to landslides' for a particular value of a particular factor. Factor values can be categorized classes or continuous values. The higher the favorability, the more favorable the factor value for landslides, and vice versa. A medium favorability value suggests that the corresponding factor value is ineffective for predicting landslide occurrence. A factor with a higher ability to determine landslide occurrence is expected to have more grid cell units with high or low favorable factor values and fewer grid cell units with medium favorable factor values, so that it can make a more distinct differentiation between landslide-prone and stable areas, thus demonstrating a higher importance for landslides.

Many data-driven methods are available for measuring favorability (e.g. Li and Lan 2023). An ideal favorability index should be within the normalized range $[-1, 1]$, and have a threshold value of 0. Positive, negative, and zero favorability values indicate favoring, not favoring, and not relevant to landslide occurrence, respectively. A factor with a higher ability to determine landslide occurrence, thus having a higher importance for landslides, is expected to have more grid cell units with favorability close to 1 or − 1 and less grid cell unit with favorability around 0. Consider the following extreme situation: A factor with a 'perfect' ability to determine landslide occurrence, thus having the highest theoretical importance, is expected to have and only have two groups of grid cells: the landslide-must-occur group which has a favorability value of 1 and the landslidemustn't-occur group which has a favorability value of – 1. The simplest extreme scenario is a factor with only two values (classes), A and B, in which all A areas are affected by landslides, whereas all B areas are free of landslides.

(2) Composition: Integrate the favorability values of all grid cell units. The normalized range [−1, 0, 1] for favorability facilitates the integration of the favorability values of all grid cell units to form a factor absolute importance index. According to the above clarification, a factor with a higher ability to determine landslide occurrence (i.e. a higher importance) is expected to have more grid cell units with absolute favorability close to 1 (close to 1 or – 1) and less grid cell units with absolute favorability around 0. Then, a straightforward measure of factor absolute importance, i.e. the analytical 'decisiveness,' will be the average of the absolute favorability values of all grid cell units. Decisiveness ranges within [0, 1]. Higher decisiveness indicates a higher importance for landslides, and vice versa. Decisiveness is an analytical index derived by explicitly determining and integrating the favorability values of all grid cell units based on landslide and predisposing factor data.

2.2. Implementation

The implementation of the decisiveness concept then comes to finding a favorability measurement ranging within [−1, 0, 1], for factors with categorized and continuous values.

2.2.1. Calculation for factors with categorized values

The certainty factor is an ideal choice for measuring the favorability to landslides for implementing decisiveness because it ranges within [−1, 0, 1] and outperforms other bivariate approaches (Li and Lan 2023). Assume that there is one landslide (L) layer and n factor (F) layers for the favorability calculation. Without loss of generality, we perform a favorability calculation based on the grid data. We denote the total count of grid cells of the entire study area by $N(A)$, total cell count of areas with landslides by $N(L)$, and ith factor by F_i. Assume that F_i is originally categorized into m classes, or has continuous values and is subdivided into m classes. Then, the total cell count of areas with the jth class of F_i ($F_{i,j}$) can be denoted by $N(F_{i,j})$, and we have:

$$\sum_{j=1}^{m} N(F_{i,j}) = N(F_i) = N(A) \quad (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \tag{1}$$

in which, $N(F_i)$ is the total cell count of the gridded ith factor layer and is equal to $N(A)$ by definition. Then, $p(L|F_i)$, which is the 'conditional probability of L given F_i', i.e. the empirical conditional probability of L given the study area A, or simply the prior probability of landslide occurrence in the study area, can be calculated by:

$$p(L|F_i) = \frac{N(L \cap F_i)}{N(F_i)} = \frac{N(L \cap A)}{N(A)} = \frac{N(L)}{N(A)} = p(L|A) = p(L) \tag{2}$$

Specifically for $F_{i,j}$ (i.e. the jth class of F_i), we can obtain $p(L|F_{i,j})$, which is the 'conditional probability of L given $F_{i,j}$', by:

$$p(L|F_{i,j}) = \frac{N(L \cap F_{i,j})}{N(F_{i,j})} \tag{3}$$

in which, $N(L \cap F_{i,j})$ is the cell count of the intersection of L and $F_{i,j}$, i.e. the count of grid cells occupied by landslides that occur within $F_{i,j}$, and can be obtained by overlying the gridded landslide layer and the gridded ith 'class-specific' factor layer.

The certainty factor (e.g. Lan et al. 2004; Li and Lan 2023) adopts a sophisticated comparison between $p(L|F_{i,j})$ and $p(L|F_i)$ to measure the favorability to landslides. It was originally proposed by Shortliffe and Buchanan (1975) and later modified by Heckerman (1986). Specifically, the 'certainty factor' (CF) for $F_{i,j}$, $CF_{i,j}$, is given by:

$$CF_{i,j} = \begin{cases} \dfrac{p(L|F_{i,j}) - p(L|F_i)}{p(L|F_{i,j}) \cdot [1 - p(L|F_i)]}, & p(L|F_{i,j}) \geq p(L|F_i) \\ \dfrac{p(L|F_{i,j}) - p(L|F_i)}{p(L|F_i) \cdot [1 - p(L|F_{i,j})]}, & p(L|F_{i,j}) < p(L|F_i) \end{cases} \quad (4)$$

The certainty factor for $F_{i,j}$ represents the change in the certainty that the proposition of landslide occurrence is true from without $F_{i,j}$ to given $F_{i,j}$. The certainty factor ranges within $[-1, 1]$ and has a threshold value of 0. A positive $CF_{i,j}$ indicates that the conditional probability of L increases from without $F_{i,j}$ to given $F_{i,j}$, and further indicates that $F_{i,j}$ favors the occurrence of landslides. In contrast, a negative $CF_{i,j}$ indicates that $F_{i,j}$ does not favor landslide occurrence. Larger (smaller) certainty factors indicate higher (lower) favorability.

Then, each grid cell can be assigned with a certainty factor value according to its factor class, and the factor absolute importance index – 'decisiveness' of the ith factor F_i, can be obtained by:

$$DC_i = \sum_{k=1}^{N(A)} |CF_{i,k}|/N(A) \ \ (i = 1, 2, \ldots, n; k = 1, 2, \ldots, N(A)) \quad (5)$$

where, the decisiveness of the ith factor F_i is denoted by DC_i, $CF_{i,k}$ is the certainty factor of the kth grid cell with respect to F_i, and $N(A)$ is the total count of grid cells in the entire study area.

It is noteworthy that this implementation makes decisiveness an integration of accountability and reliability. Consider the following extreme situation: For the ith factor F_i, suppose its certainty factor values are either 1 or – 1, i.e. the conditional probability $p(L|F_{i,j})$ is either 1 or 0. In other words, for any specific factor class $F_{i,j}$, either landslides cover all grid cells or landslides do not occur in any grid cells. This scenario has a decisiveness value of 1, and in the meantime leads to a 100 accountability and a 100 reliability (e.g. Castellanos Abella 2008; Greenbaum et al. 1995). Therefore, as a single index, decisiveness can provide information related to the two indices (accountability and reliability), making evaluations of absolute importance more concise.

2.2.2. Calculation for factors with continuous values

The classification of predisposing factors with continuous values is conventionally required in the first place by bivariate approaches for favorability calculation, including the certainty factor method. Prior factor reclassification induces two problems (e.g. Li et al. 2017; Li and Lan 2023): (1) a discontinuity problem, which means that all factor values in the same class will have the same favorability value; and (2) a subjectivity problem, which means that the choices of the number and divisions of factor classes are subjective. To address the uncertainties associated with factor classification, the authors proposed a 'classification-free modification' on bivariate approaches (Li et al. 2017; Li and Lan 2023; Zhang et al. 2020), in which the classification of factors with continuous values is avoided. The steps for the favorability calculation for factors with continuous factor values are as follows (Li and Lan 2023).

First, the continuous factor values are normalized to the range [0, 1]. This normalization allows the use of the same parameters for different factors in the following procedures.

Second, a parameter 'precision' which defines the number of digits after the decimal point is applied to get identical normalized factor values. For example, if the precision is 4, the normalized factor values of 0.21427 and 0.21421 will both change to an identical normalized factor value of 0.2142, and there will be at most 10001 identical normalized factor values. The application of precision is aimed at reducing the calculation load by reducing the number of identical normalized factor values. It must be emphasized that setting precision is not obligatory, and a value of 0 indicates that no precision will be applied, which further means that a favorability value will be calculated for each of the original factor values.

Third, a bin is created for each identical normalized factor value, which centers at this value and has a width defined by a parameter 'bin width.' This bin width ranges within [0, 1] because the factor values have been normalized. The bins of neighboring identical normalized factor values may overlap and can also have gaps if low precision and a small bin width are adopted. A minimum effective bin width exists, which is determined by precision. For example, if the precision is 4, the minimum difference between two neighboring identical normalized factor values is 0.0001. Thus, a bin width smaller than 0.0002 is ineffective because it means that there will be only one identical normalized factor value within any bin.

The principal concept of classification-free modification involves calculating the favorability of continuous bins instead of discrete classes. Because bin width is the only obligatory parameter that must be defined by users, the uncertainties introduced by manual factor classification in conventional bivariate approaches are moderated. As in the case of the categorized factors, the empirical conditional probabilities and favorability values for each bin are obtained. Then, each grid cell can be assigned a favorability (certainty factor) value according to the corresponding identical normalized factor value and bin, and the decisiveness can be calculated using Eq. (5). The assumption behind the evaluation of the favorability of a factor value based on a bin is that neighboring factor values have similarities in determining landslide presence.

2.2.3. Open software (ALSA)

The open software '*Automatic Landslide Susceptibility Analysis* (ALSA)' version 3.0 proposed by the authors has implemented a general optimization framework for eight 'conditional-probability-based' bivariate methods, including the certainty factor method (Li and Lan 2023). This paper presents versions 4.0, 4b.0, and 5.0 of the ALSA; the interface of V5.0 is shown in Figure 1. ALSA V4.0 and V4b.0 are add-ins of ESRI ArcMap and can be used to calculate favorability and further decisiveness for individual landslide predisposing factors. ALSA V4b.0 produces the same results as V4.0; the difference is that V4b.0 can perform block processing to moderate the constraint from the memory limit. Considering that ArcMap will not be supported by ESRI in the foreseeable future, ALSA V5.0, which is an add-in compatible with the next-generation ArcGIS Pro, has also been developed. ALSA V5.0, retaining the functionalities of V4.0, has advantages in calculation efficiency and memory capacity, making it the recommended choice if ArcGIS Pro is accessible. All versions of the ALSA are freely available.[1] Alternative open tools for regional landslide susceptibility analysis exist (e.g. Bragagnolo, da Silva, and Grzybowski 2020; Guo et al. 2023; Torizin, Schüßler, and Fuchs 2022), some of which can output factor importance measurements (Bragagnolo, da Silva, and Grzybowski 2020), yet only relative factor importance. To the best knowledge of the authors, ALSA V5.0 currently stands as the only open software for landslide susceptibility analysis capable of calculating and outputting absolute factor importance.

The inputs of ALSA V5.0 include landslides, predisposing factors, and processing extent data. Users should do parameters, optimization, decisivenesses, and output settings, as well as choose bivariate methods. To calculate the decisiveness index, the certainty factor method will be used by default in the background. If there is at least one factor with continuous values, inputs for precision and bin width, as well as settings for optimization and decisivenesses, are enabled. Personal experiences suggest that a precision of 4 and a bin width of 0.1 would be reasonable initial choices (e.g. Zhang et al. 2020). The bin width input will be disabled if optimization is chosen because an optimal bin width will be generated. If time permits, an optimal bin width can be derived by optimizing the landslide susceptibility assessment with all the considered predisposing factors. Decisiveness values corresponding to the input or optimized bin width will be calculated and output for continuous factors. If the decisivenesses option is checked, decisiveness values corresponding to a series of equal-interval bin widths will also be calculated and output so that the variation of

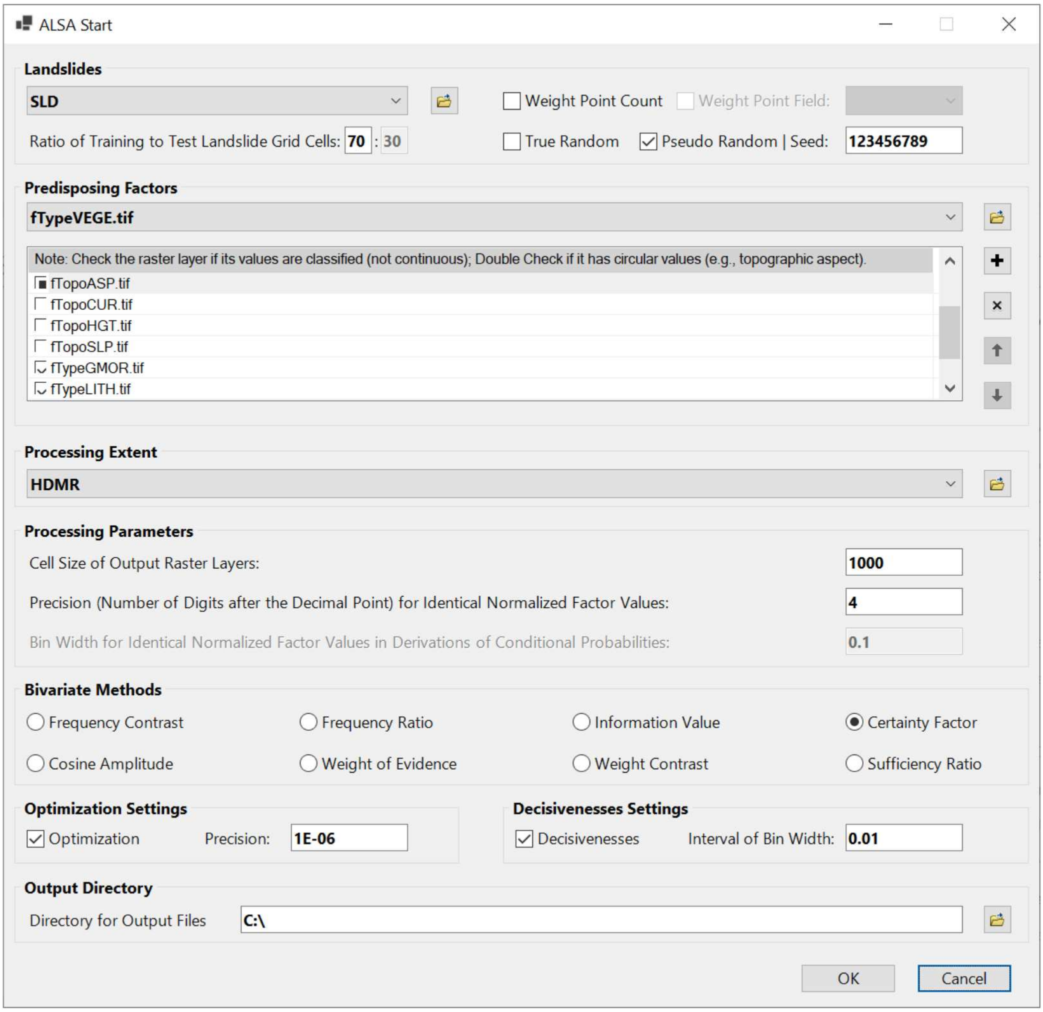

Figure 1. Interface of the open software ALSA (*Automatic Landslide Susceptibility Analysis*) version 5.0 used to calculate favorability and further decisiveness for individual landslide predisposing factors.

decisiveness with the bin width can be inspected. ALSA V5.0 outputs several files, including a log file, favorability values (including certainty factor values) for all classes and bins of identical normalized factor values of all factors, favorability raster layers of all factors, decisiveness values of all factors, a *LSI* (landslide susceptibility index) raster layer, and some other intermediate files.

2.3. Criteria for manipulating bin width and class count

It must be emphasized that decisiveness is dependent on the selection of bin width for factors with continuous values and is also influenced by the class count of categorized factors. Theoretically, a smaller bin width or a larger class count will yield a higher decisiveness, and caution should be exercised when using too small values of bin width or adopting categorized data with too many classes. This is because a smaller bin width means less overlap, larger divergence of favorability values among factor bins, thus higher decisiveness. In particular, it is meaningless to use a bin width smaller than the minimum effective value because this will probably induce an extreme situation in which the conditional probability of factor bins is either 1 or 0, leading to an irrational decisiveness of 1, especially if the precision is very small. The meaninglessness of using categorized data with an excessively large count of classes is more straightforward. The theoretical maximum class count for a study area is the total count of grid cells within it. If each grid cell corresponds to an individual class, factor classes will have a conditional probability of either 1 or 0, resulting in a meaningless decisiveness value of 1. This phenomenon can be regarded as a form of overfitting, which is also an issue that should be considered in approaches reliant on spatial data discretization (e.g. Blahut, van Westen, and Sterlacchini 2010; Wang et al. 2010); that is, over discretization of factor data leads to overfitting although higher power of determinant can be obtained. Therefore, criteria are required for selecting bin width in the decisiveness calculation of factors with continuous values and for dealing with the influence of class count on the decisiveness of categorized factors. This study proposes the following two criteria for manipulating bin width and class count in decisiveness calculation (Table 1).

(1) Criterion of susceptibility optimization. The dominance of the single parameter bin width allows for the optimization of landslide susceptibility assessment based on bivariate approaches (Li and Lan 2023; Zhang et al. 2020). An intuitive and reasonable first choice for bin width could be the optimized one, which yields the maximum prediction AUC (area under the receiver operating characteristic (ROC) curve) for landslide susceptibility analysis. Under this criterion, the optimized bin width would no longer be a subjective choice, and factor importance derived based on the optimized value would reflect the results of landslide susceptibility evaluation. However, this criterion encounters two complications when the purpose is to compare decisiveness. First, because different cases may yield different optimized bin width values, comparing the decisiveness importance of the same factor with continuous values across cases would be questionable. Second, because different categorized factors may have different class counts, comparing decisiveness among categorized factors would involve the influence of class count. Therefore, a second criterion is proposed as follows.
(2) Criterion of classification compensation. Generally, comparisons of decisiveness between categorized and continuous factors should be performed cautiously, and if a comparison is necessary, the bin width can be selected according to the classification of the categorized factors. The way in which bin width influences the decisiveness of continuous factors is in principle similar to that in which class count influences the decisiveness of categorized factors. Therefore, if a categorized factor has m classes, it is justifiable to adopt a bin width of $1/m$ for continuous factors. This is because a bin width of $1/m$ implies subdividing the normalized range [0, 1] into m non-overlapping bins. In this way, we can consider that continuous factors are first 'reclassified into m classes,' and then their decisiveness indices are compared to that of the categorized factor with m classes; i.e. if the target comparison categorized factor has m classes, the effect of factor discretization is assumed to be compensated by selecting $1/m$ as the bin width for continuous factors.

When comparing the decisiveness indices of several categorized factors with different class counts, we propose to compensate the influence of class count by two steps. In the first step, it is suggested to minimize the difference in class count among categorized factors in a physically sensible manner; specifically, by choosing the appropriate data or integrating classes based on expert

Table 1. Criteria for manipulating bin width and class count in decisiveness calculation.

Criterion	Advantage	Disadvantage
Criterion of susceptibility optimization	Decisiveness importance of factors can reflect the results of landslide susceptibility analysis.	Decisiveness comparisons might be complicated by the inconsistency of bin width and class count.
Criterion of classification compensation	Effects of bin width and class count can be compensated for decisiveness comparisons.	Decisiveness cannot reflect the results of landslide susceptibility analysis.

knowledge. For example, if higher hierarchical data with 11 classes and lower hierarchical data with 150 classes are available for vegetation type, and higher hierarchical data with 14 classes and lower hierarchical data with 116 classes are available for geomorphology, the sensible choice of data would be the 11-class vegetation type data and the 14-class geomorphology data because the difference in class count is minimized. If 14-class geomorphology data is not available, it is suggested to first produce higher hierarchical data for geomorphology based on the available 116-class data via experts' integrations of geomorphologic units, aiming for a class count as close to 11 as possible. In the second step, it is suggested to compensate the difference in class count among categorized factors in a mathematically reasonable manner. Since physically meaningful ways cannot reduce the difference in class count any more, mathematically justified procedures can be applied therein. The idea is to achieve a uniform class count across all categorized factors, targeting the minimum class count of all factors. This can be accomplished by randomly integrating the classes of categorized factors with larger class counts until all resulting categorized factors have an identical class count. For example, in the former case of selecting 11-class vegetation type data and 14-class geomorphology data, a series of 11-class geomorphology data can be generated from the 14-class data through random integrations of geomorphologic units (Table 2). Then, the mean decisiveness of those randomly generated 11-class data can be adopted as the classification compensated decisiveness of geomorphology. In practice, we can take into account either all possible random integrations of geomorphologic units (Table 2), or just a predefined number of Monte Carlo simulations for saving time. The classification compensated bin width for continuous factors would then be 1/11 (0.0909091).

Both the above two criteria have their advantages and disadvantages (Table 1). Adopting the susceptibility optimization criterion allows the derived decisiveness importance of factors to reflect the results of landslide susceptibility analysis; however, comparisons of decisiveness may be complicated due to inconsistencies in bin width and class count. On the other hand, using the classification compensation criterion enables compensation for the effects of bin width and class count in decisiveness comparisons, yet the derived decisiveness may not accurately reflect landslide susceptibility. It is suggested to choose the appropriate criterion according to the purpose of measuring factor absolute importance while acknowledging that decisiveness is an index influenced by both bin width and class count.

3. Case study

3.1. Study area and data

The case study is an extension of the one performed in our previous work that presents ALSA V3.0 (Li and Lan 2023). The case study area is the Heng-duan Mountains region (HDMR). The HDMR is located in the southeastern, marginal part of the Tibetan Plateau (Figure 2). Owing to the constraints of major active faults, six large deeply-cut rivers, the Min, Dadu, Yalong, Jinsha, Lancang, and Nu Rivers, flow approximately parallel from north to south across the HDMR (Figure 2). Lying between those parallelly distributed large rivers is a high mountain range with relief as high as 5000 M. The HDMR has suffered from severe mountain hazards owing to its high topographic relief and fragile geo-environments (e.g. Lan et al. 2003; Meng et al. 2015; Yang, Zheng, and Liu 1988). Major earthquakes in the HDMR, such as the 2013 Lushan earthquake (e.g. Yang et al. 2015) and the 2008 Wenchuan earthquake (e.g. Qi et al. 2010), triggered massive landslides. In addition, active landsliding areas have been widely detected in the HDMR (e.g. Yao et al. 2022), threatening lives, property, and engineering projects at risk. Therefore, landslide risk analysis is of great importance for HDMR (e.g. Li et al. 2017; Zhao et al. 2022).

The decisiveness indices of 12 landslide predisposing factors were calculated in this study. The eight factors with continuous values are elevation, slope, aspect, curvature, distance to fault, distance to river, distance to road, and average annual precipitation; whereas the four categorized

Table 2. Decisiveness of categorized landslide predisposing factors before and after random class integrations in the case study.

				Decisiveness after random class integrations				
Factor	Original class count	Decisiveness	Count of random class integrations	Mean	Median	Standard Deviation	Minimum	Maximum
Lithologic time	15	0.3642009	1,479,478	0.3394191	0.3483088	0.0324335	0.1106600	0.3901469
Geomorphic unit	14	0.5831247	66,066	0.5580237	0.5804508	0.0817566	0.0390517	0.5957604
Soil taxonomy	12	0.6620196	66	0.6457195	0.6620196	0.0503756	0.4076716	0.6625956
Vegetation type	11	0.2867001	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

N.A.: Not applicable.

Note: Random class integrations were performed on categorized factors with 15, 14, and 12 classes to produce a series of 11-class categorized data. All possible random integrations of classes were taken into account. Taking geomorphologic unit as an example, there can be three scenarios of class integrations. (1) Keeping 10 classes and integrating the other 4 classes. The count of possible integrations is 1001 (C_{14}^{4}). (2) Keeping 9 classes and grouping the other 5 classes into 2 integrated categories (2-calss vs. 3-class). The count of possible integrations is 20020 ($C_{14}^{5} \times C_{5}^{2}$). (3) Keeping 8 classes and grouping the other 6 classes into 3 identical integrated categories (2-calss vs. 2-calss vs. 2-class). The count of possible integrations is 45045 ($C_{14}^{6} \times C_{6}^{2} \times C_{4}^{2} / P_{3}^{3}$). Then, the total count of random class integrations for the 14-class geomorphologic unit data will be 66,066.

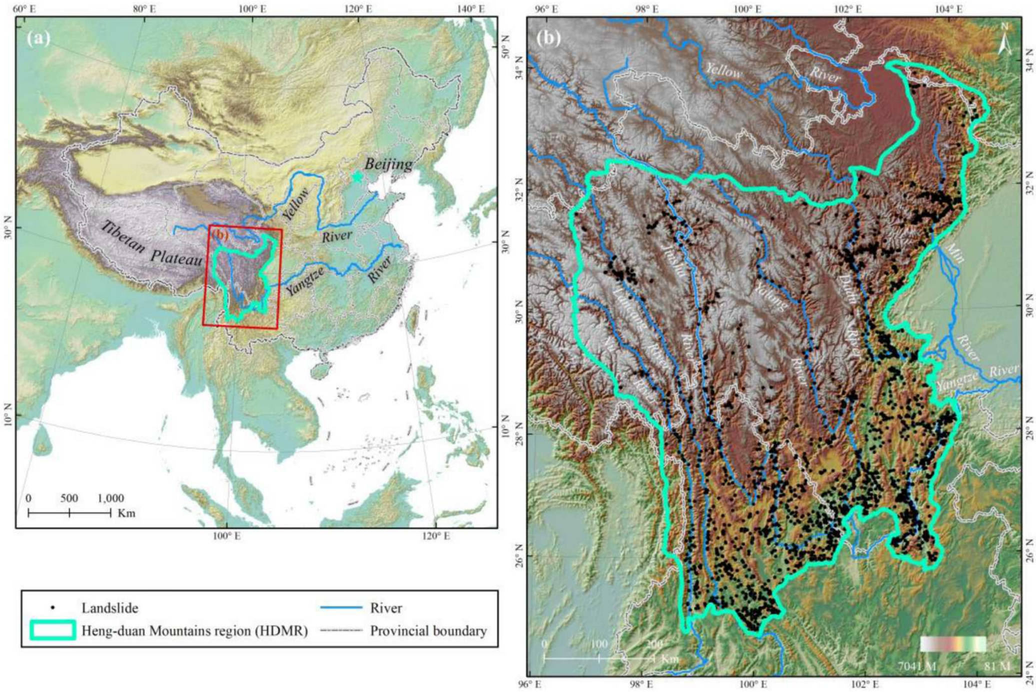

Figure 2. Heng-duan Mountains region (HDMR). The location of the HDMR relative to the Tibetan Plateau is shown in (a). Landslides used in the case study are shown in (b).

factors are lithologic time, geomorphologic unit, soil taxonomy, and vegetation type (Figure 3). The elevation, slope, aspect, and curvature data were derived from the SRTM digital elevation model (DEM) with a 30 M × 30 M spatial resolution. Fault and lithology data were obtained from a 1:500,000 geological map of China. River and road data were obtained from the 1:1,000,000 national basic geographic information dataset of China. The spatial scale of the geomorphologic, soil, and vegetation data is 1:1,000,000. The 1 Km × 1 Km grid precipitation data records the average annual precipitation from 1980 to 2015. A total of 2632 landslide records in the HDMR (Figure 2(a)) were adopted in the case study, which was obtained from a geological disaster dataset of China (RESDC 2020).

3.2. Scenarios and results

The decisiveness of the 12 landslide predisposing factors in the HDMR case study was calculated using the ALSA. The cell size was set to 1 Km. For factors with continuous values, the precision was set to 4. Both the two criteria for manipulating bin width and class count were applied. For the criterion of susceptibility optimization, the optimized bin width was 0.0581362. For the criterion of classification compensation, a bin width value of 0.0909091 (i.e. 1/11) was used. The reciprocal of 11 was selected because the categorized lithologic time, geomorphic unit, soil taxonomy, and vegetation type data used in the case study have 15, 14, 12, and 11 classes, respectively; thus 11-class was the target of random class integrations in classification compensation (Table 2). To interpret the derived decisiveness values, variations of favorability (certainty factor) with factor values and grid cell frequencies of the favorability values are shown in Figure 4 and Figure 5, respectively. In addition, decisiveness values corresponding to the 99 bin width values spacing by an interval of 0.01 are calculated for each continuous factor to inspect the variation of decisiveness with bin width (Figure 6). Similarly, for each categorized factor, decisiveness values were calculated using data with

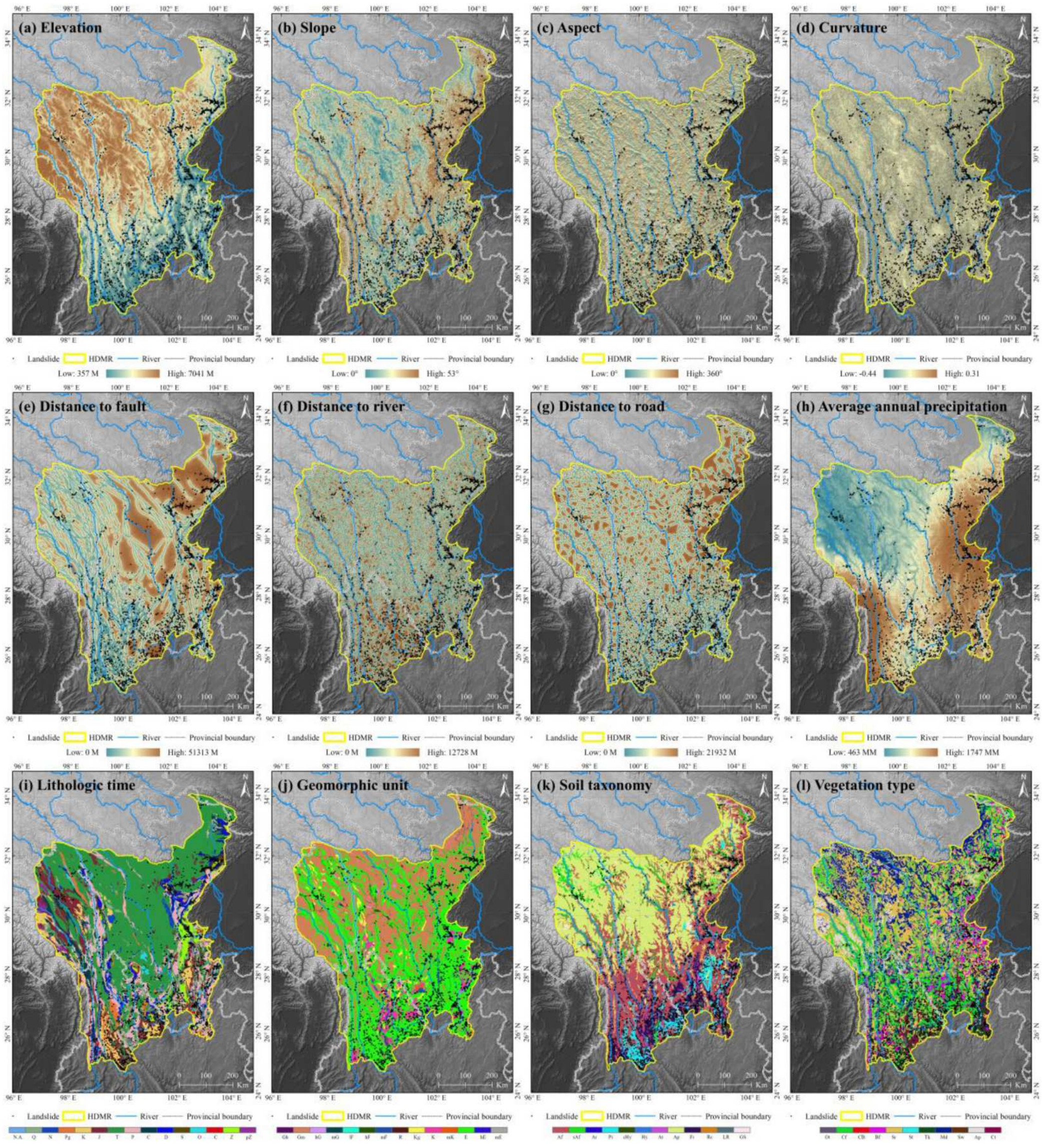

Figure 3. Predisposing factors adopted in the case study. (a) Elevation. (b) Slope. (c) Aspect. (d) Curvature. (e) Distance to fault. (f) Distance to river. (g) Distance to road. (h) Average annual precipitation. (i) Lithologic time. (j) Geomorphic unit. (k) Soil taxonomy. (l) Vegetation type. For factors with continuous values, histogram equalization stretching is used in illustrations of factor values. Lithologic time has 15 categories – N.A.: Not available; Q: Quaternary; N: Neogene; Pg: Paleogene; K: Cretaceous; J: Jurassic; T: Triassic; P: Permian; C: Carboniferous; D: Devonian; S: Silurian; O: Ordovician; Ꞓ: Cambrian; Z: Ediacaran; pZ: Pre-Ediacaran. Geomorphic unit has 14 categories – Gh: Glacial, Periglacial, Moraine Hill; Gm: Glacial, Periglacial, Moraine Mountain; hG: High-altitude Moraine, Glacial-erosion, Glacial-water, Periglacial Hill, Platform and Plain; mG: Middle-altitude Moraine, Glacial-water, Periglacial Hill, Platform and Plain; lF: Low-altitude Alluvial, Proluvial, Lacustrine, Fluvial Landform; hF: High-altitude Alluvial, Proluvial, Lacustrine, Fluvial Landform; mF: Middle-altitude Alluvial, Proluvial, Lacustrine, Fluvial Landform; R: River, Lake and Reservoir; Kg: Karst Glacial, Periglacial Mountain and Hill; K: Karst Mountain and Hill; mK: Middle-altitude Karst Platform and Plain; E: Erosion, Denudation Mountain and Hill; hE: High-altitude Erosion, Denudation Platform and Plain; mE: Middle-altitude Erosion, Denudation Platform and Plain. Soil taxonomy has 12 categories – Af: Alfisols; sAf: Semi-Alfisols; Ar: Aridisols; Pr: Primitive Soils; sHy: Semi-Hydromorphic Soils; Hy: Hydromorphic Soils; At: Anthrosols; Ap: Alpine soils; Fr: Ferralsols; Rc: Rock; LR: Lake or Reservoir; GS: Glacier or Snow. Vegetation type has 11 categories – Ot: Others; Cf: Coniferous Forest; CB: Coniferous and Broadleaf Forest; Bf: Broadleaf Forest; Sr: Shrub; St: Steppe; Tk: Tussock; Md: Meadow; Sw: Swamp; Ap: Acrophyta; Ct: Cultigen.

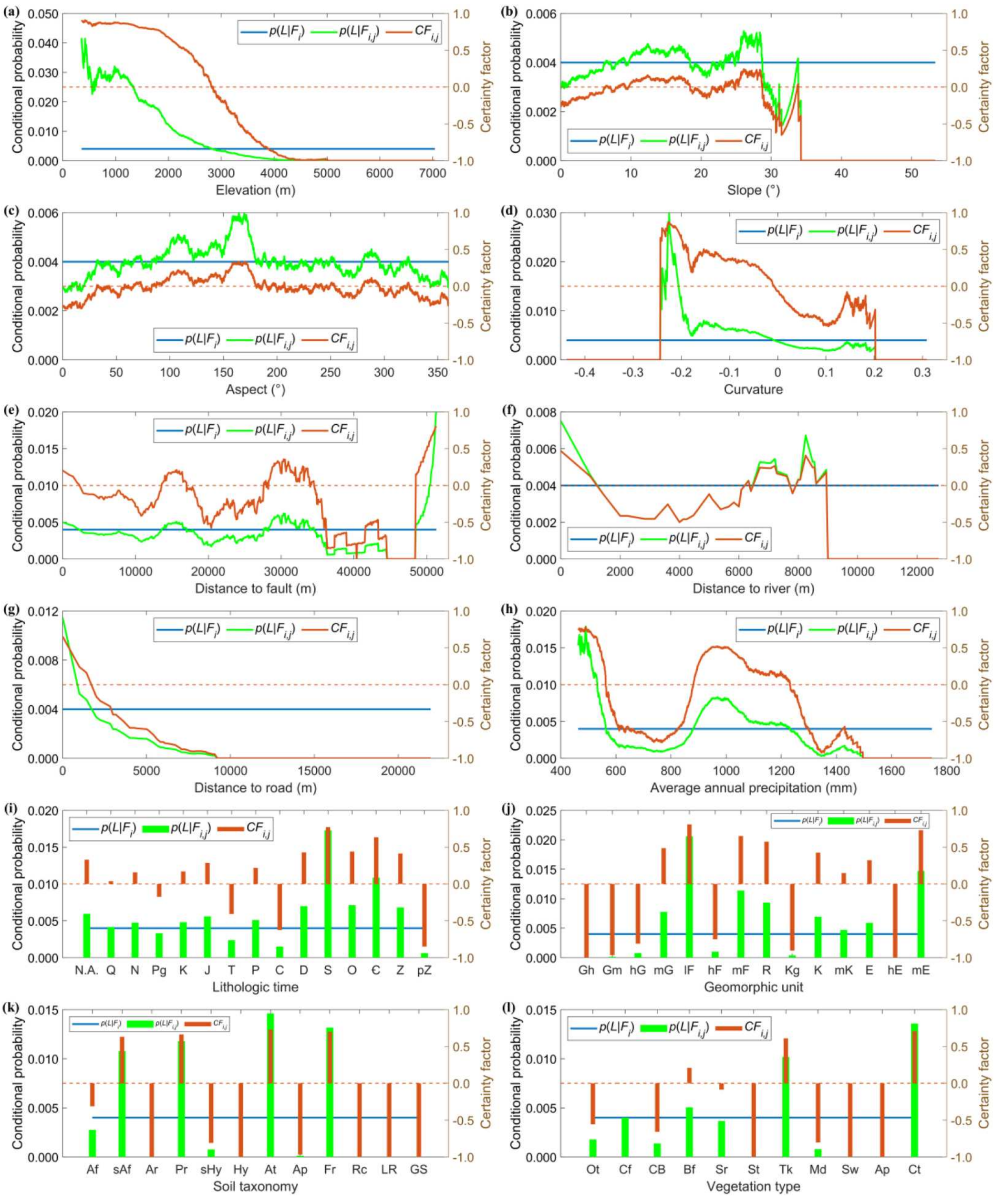

Figure 4. Variations of favorability (certainty factor) with factor values for predisposing factors adopted in the case study. (a) Elevation. (b) Slope. (c) Aspect. (d) Curvature. (e) Distance to fault. (f) Distance to river. (g) Distance to road. (h) Average annual precipitation. (i) Lithologic time. (j) Geomorphic unit. (k) Soil taxonomy. (l) Vegetation type. Categories of categorized factors are referred to Figure 3. Results for the criterion of susceptibility optimization are presented.

lower hierarchical classifications to determine the influence of the class count on decisiveness (Figure 7).

Accountability & reliability, AUC, and geographical detector methods were also used to calculate absolute importance indices for comparison. The sensitivity index method was not used because it only considers the maximum and minimum CF values in the importance evaluation (e.g. Wang et al. 2014), which is inevitably biased. The accountability & reliability and geographical detector

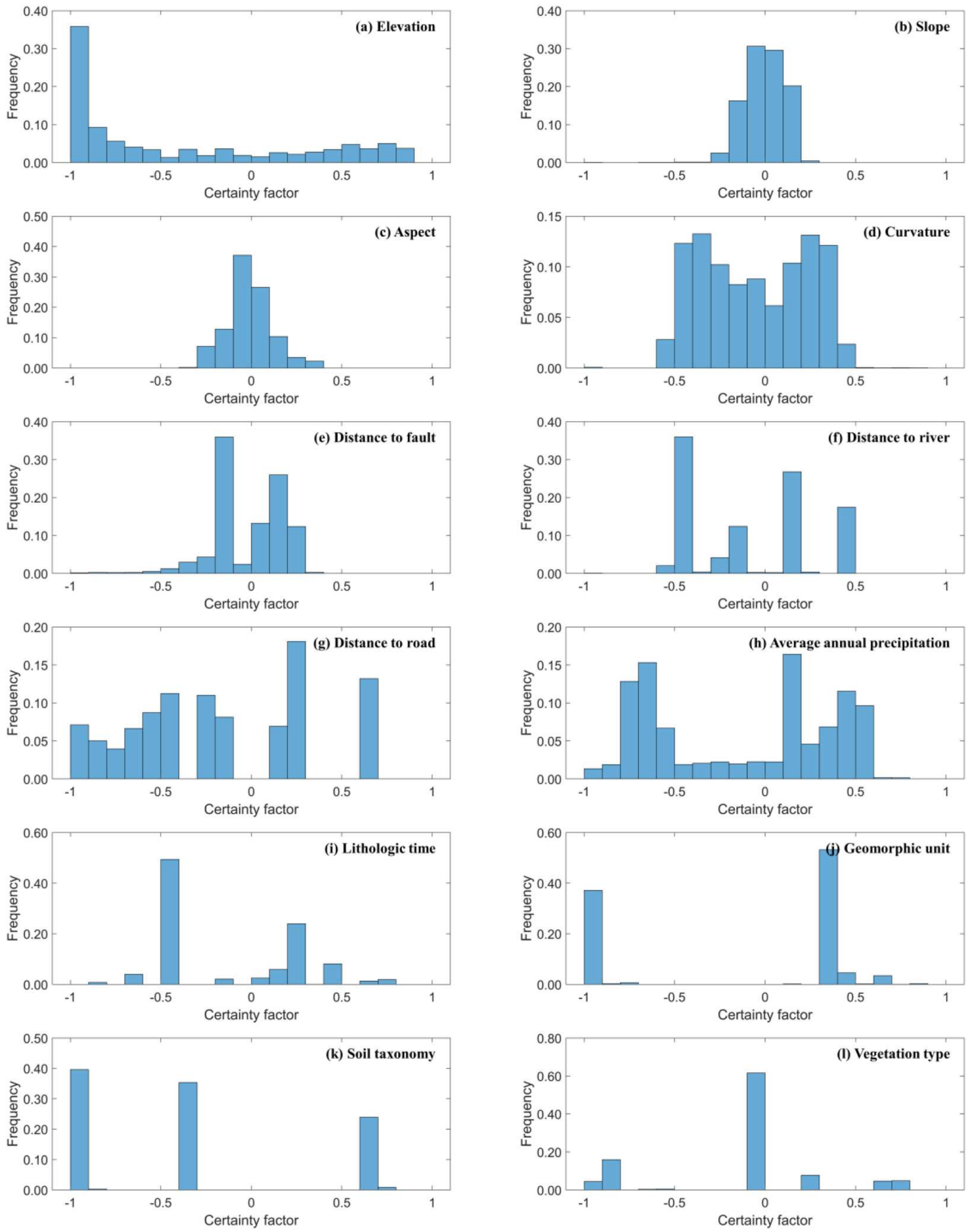

Figure 5. Grid cell frequencies of favorability (certainty factor) values for predisposing factors adopted in the case study. (a) Elevation. (b) Slope. (c) Aspect. (d) Curvature. (e) Distance to fault. (f) Distance to river. (g) Distance to road. (h) Average annual precipitation. (i) Lithologic time. (j) Geomorphic unit. (k) Soil taxonomy. (l) Vegetation type. Results for the criterion of susceptibility optimization are presented.

methods require prior factor reclassification of factors with continuous values. Without loss of generality, continuous factors were categorized into 13 classes using the Jenks natural breaks method, because 13 is the average of 15, 14, 12, and 11. The AUC values for landslide predisposing factors

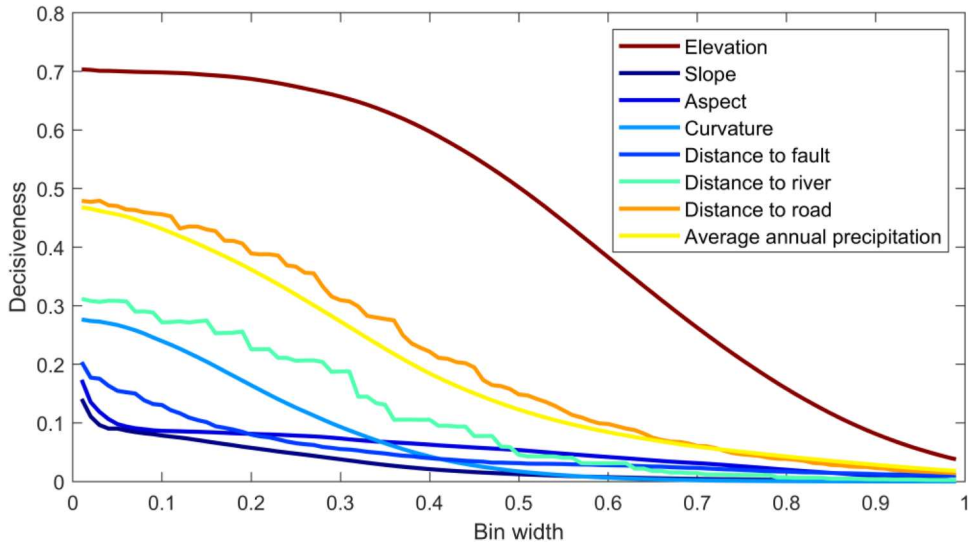

Figure 6. Variation of decisiveness with bin width for factors with continuous values.

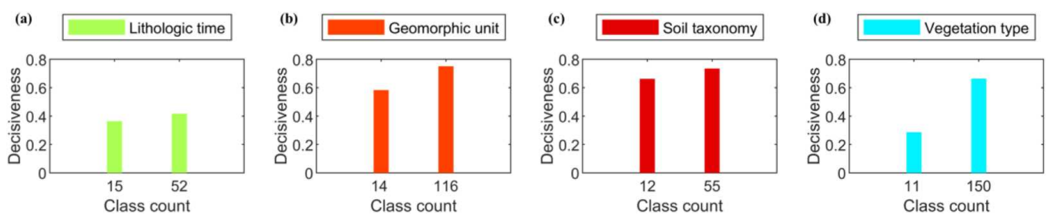

Figure 7. Variation of decisiveness with class count for factors with categorized values. Please be aware that class count is not-to-scale in the *x* axis.

were calculated based on the corresponding ROC (receiver operating characteristic) curves (Figure 8). It is noteworthy that an AUC lower than 0.5 means that a reverse prediction will give a better performance; thus, the AUC value is within the range [0.5, 1], and a value of 0.5 indicates no prediction ability at all.

Some adjustments can be applied to the accountability & reliability and AUC methods to moderate their inherent defects. For the accountability & reliability method, an index 'Acc&Rel', which is the average of normalized accountability and reliability, is introduced to integrate the evaluations given by the two individual indices. This simple integration avoids the dilemma of which of the two indices to use if they provide contrasting results (e.g. Greenbaum et al. 1995). For the AUC method, before the AUC calculations of the categorized factors, the factor classes were renumbered according to their favorability. Factor classes with higher ranks of favorability were assigned larger class numbers so that a monotonic relationship exists between the factor class number and empirical landslide probability. The AUC values derived before and after the adjustment are denoted by 'AUCo' and 'AUCr,' respectively. For factors with continuous values, the AUCo and AUCr are the same.

The decisiveness values of the categorized factors before and after random class integrations are listed in Table 2. The absolute importance indices of the 12 landslide predisposing factors given by the decisiveness, accountability & reliability, AUC, and geographical detector methods are listed in Table 3. In the geographical detector method, factor importance is evaluated by the

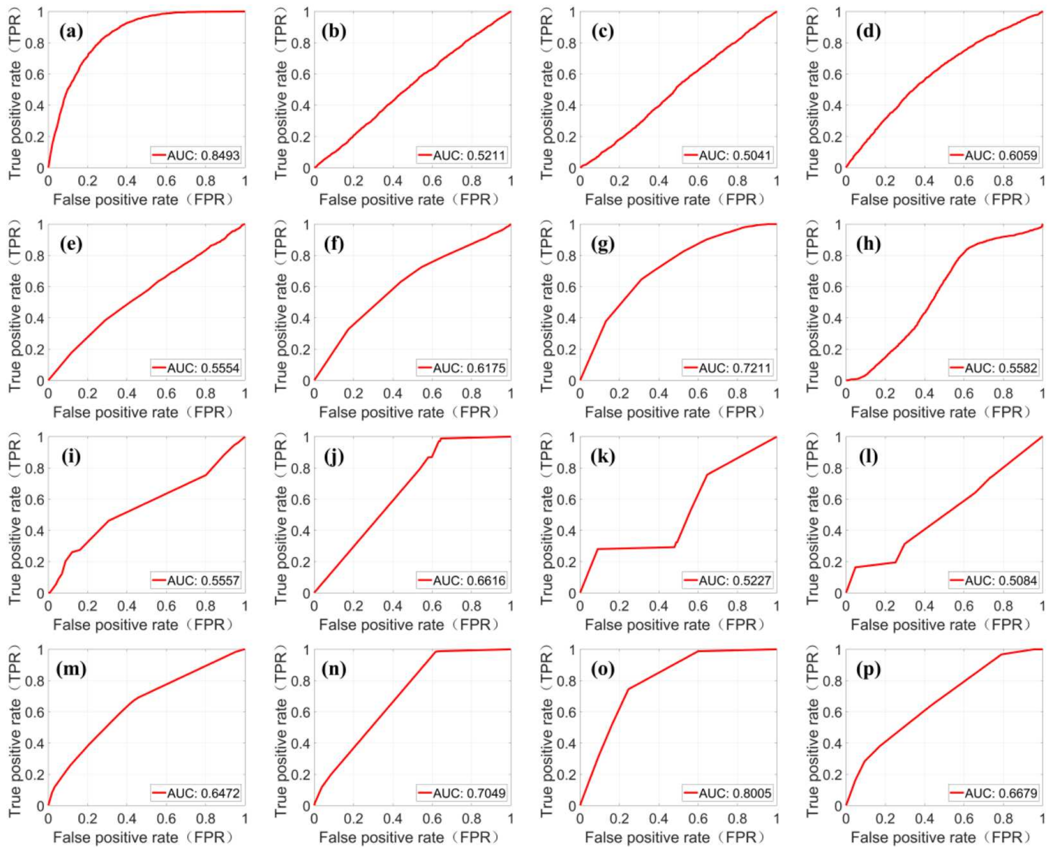

Figure 8. ROC (receiver operating characteristic) curves for predisposing factors adopted in the case study. (a) Elevation. (b) Slope. (c) Aspect. (d) Curvature. (e) Distance to fault. (f) Distance to river. (g) Distance to road. (h) Average annual precipitation. (i) Lithologic time. (j) Geomorphic unit. (k) Soil taxonomy. (l) Vegetation type. (m)-(p) ROC curves after the renumbering of classes for categorized lithologic time, geomorphic unit, soil taxonomy, and vegetation type, respectively.

index ‘q value.’ The factor absolute importance ordered by the derived indices is shown in Figure 9. It is clear that different indices provide quite different evaluations of the absolute importance of a factor.

4. Discussion

4.1. Significance of decisiveness

The proposed decisiveness reflects the ability of landslide predisposing factors to differentiate between landslide-susceptible and landslide-insusceptible areas, thus measuring the absolute importance of factors. The decisiveness rank of factors given by the susceptibility optimization and classification compensation criteria is identical (Figure 9(a, b)). The three factors with the highest decisiveness were elevation, soil taxonomy, and geomorphologic unit (Figure 9). These three factors can make the most distinct differentiations between landslide-susceptible and landslide-insusceptible areas because their favorability (certainty factor) values are concentrated closer to the decisive values of – 1 and 1, away from the indecisive value of 0 (Figures 4 and 5). In contrast, the favorability values of slope and aspect were distributed around the indecisive value of 0 (Figures 4 and 5), inducing the lowest decisiveness values (Figure 9), thus the lowest ability to differentiate landslide-prone and stable areas. In other words, factors with higher decisiveness values are

Table 3. Absolute importance indices of landslide predisposing factors given by the decisiveness, accountability & reliability, AUC, and geographical detector methods in the case study.

	Decisiveness		Accountability & Reliability			AUC		
Factor	Criterion of susceptibility optimization	Criterion of classification compensation	Accountability	Reliability	Acc&Rel	AUCo	AUCr	q value
Elevation	0.6998	0.6986	83.9814	1.1505	0.8465	0.8493	0.8493	0.0092
Slope	0.0880	0.0805	50.0871	0.4453	0.0999	0.5211	0.5211	0.0001
Aspect	0.0948	0.0872	39.4080	0.4576	0.0197	0.5041	0.5041	0.0001
Curvature	0.2636	0.2453	58.9669	0.5510	0.2432	0.6059	0.6059	0.0006
Distance to fault	0.1523	0.1319	49.5647	0.5017	0.1329	0.5554	0.5554	0.0002
Distance to river	0.3073	0.2884	63.6680	0.5717	0.2958	0.6175	0.6175	0.0006
Distance to road	0.4643	0.4578	70.6326	0.7393	0.4642	0.7211	0.7211	0.0020
Average annual precipitation	0.4531	0.4364	81.8920	0.6116	0.4730	0.5582	0.5582	0.0016
Lithologic time	0.3642	0.3394	67.3825	0.6160	0.3558	0.5557	0.6472	0.0017
Geomorphic unit	0.5831	0.5580	98.4330	0.6359	0.6260	0.6616	0.7049	0.0027
Soil taxonomy	0.6620	0.6457	74.4051	1.2018	0.8012	0.5227	0.8005	0.0057
Vegetation type	0.2867	0.2867	38.0151	0.8819	0.2886	0.5084	0.6679	0.0022

Decisiveness: Bin width values in the criterion of susceptibility optimization and criterion of classification compensation are 0.0581362 and 0.0909091, respectively. Classification compensated decisiveness values for categorized factors are referred in Table 2.

Acc&Rel: The average of normalized accountability and reliability.

AUCo and AUCr: The AUC values derived before (AUCo) and after (AUCr) the renumbering of classes for categorized factors. For a factor with continuous values, its AUCo and AUCr will be the same.

q value: The importance index calculated by the geographical detector method.

more important in determining the landslide spatial distribution, whereas factors with lower decisiveness values are less important.

Notably, the principle of decisiveness is comparable to that of the geographical detector method. For a categorized factor, a low variance of landslide presence within each factor class means a high q value, high power of determinant, thus, high importance (e.g. Wang et al. 2010). For a specific factor class, both high and low favorability values indicate a low variance of landslide presence. In other words, for a factor class, both the situation in which most of its area presents landslides (high favorability) and the situation in which most of its area does not present landslides (low favorability) indicate that it has a low variance of landslide presence. Therefore, a categorized factor has a high decisiveness value, meaning that its factor classes have either high or low favorability values. This further means that all its factor classes have a low variance of landslide presence, resulting in a high q value. Nevertheless, an obvious advantage of the decisiveness method over the geographical detector method is that variations between negatives and positives of the favorability of factor values can be inspected (Figure 4), which cannot be revealed by the variance of landslide presence. In particular, for factors with continuous values, the correlation of landslide presence with the factor value can be inspected in detail using the proposed decisiveness method; for example, in the case study, typical monotonic (Figure 4(a, g)) and non-monotonic (Figure 4(e, h)) correlation patterns can be detected.

4.2. Robustness of decisiveness

The proposed decisiveness introduces a robust measure for the absolute importance of a factor. For some factors, such as vegetation type, geomorphologic unit, and average annual precipitation, accountability and reliability give quite different evaluations of importance (Figure 9(b, c)). The integrated index ‘Acc&Rel’ gives almost the same order of factor absolute importance as decisiveness (Figure 9(a, d)), suggesting the effectiveness of the adjustment. However, the uncertainties associated with the prior classification of factors with continuous values remain unknown. Currently, we do not have a realistic way of confirming whether other uncountable choices for

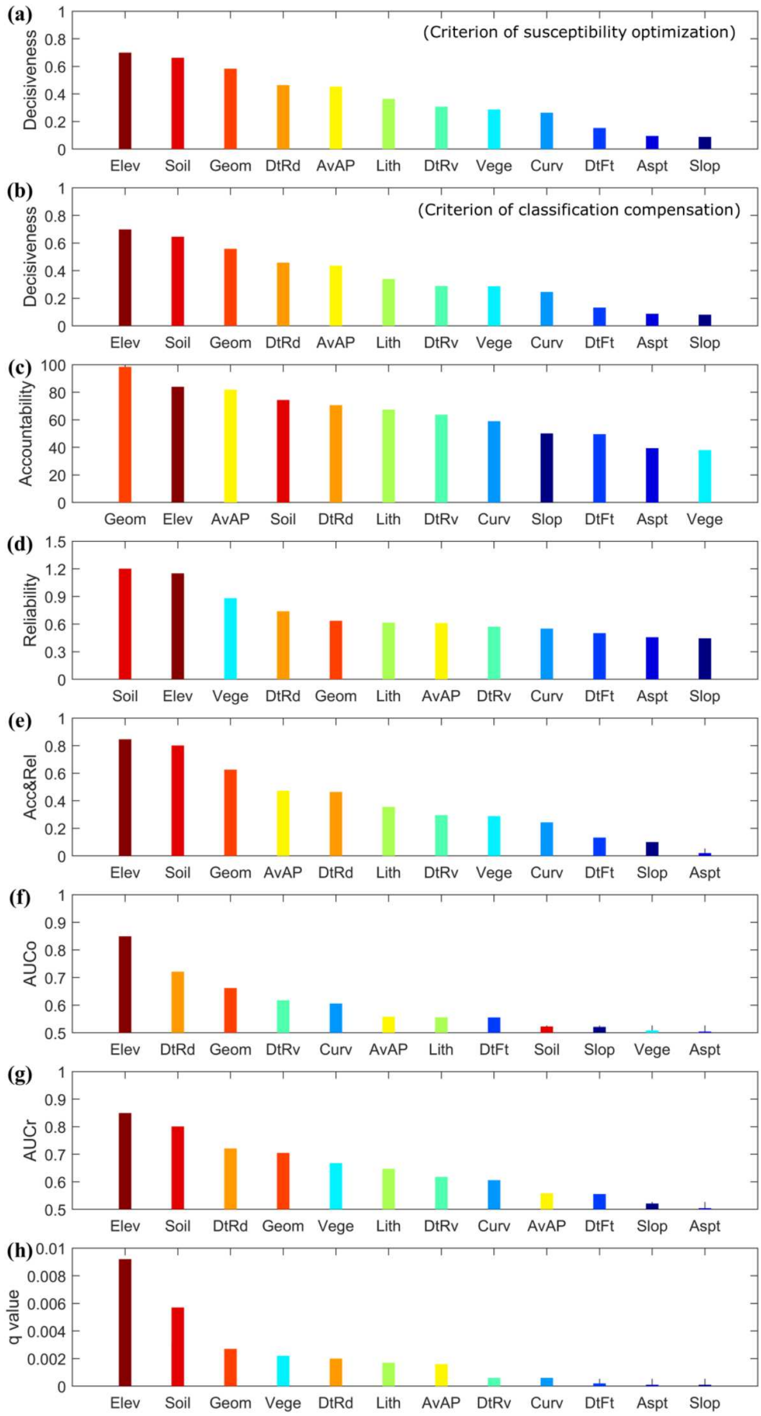

Figure 9. Factor absolute importance ordered by the derived indices in the case study. (a) Decisiveness for the criterion of susceptibility optimization. (b) Decisiveness for the criterion of classification compensation. (c) Accountability. (d) Reliability. (e) Acc&Rel, which is the average of normalized accountability and reliability. (f) AUCo and (g) AUCr are the AUC values derived before and after the renumbering of classes for categorized factors. For a factor with continuous values, its AUCo and AUCr will be the same. (h) q value. Names of landslide predisposing factors are abbreviated. Elev: Elevation; Slop: Slope; Aspt: Aspect; Curv: Curvature; DtFt: Distance to fault; DtRv: Distance to river; DtRd: Distance to road; AvAP: Average annual precipitation; Lith: Lithologic time; Geom: Geomorphic unit; Soil: Soil taxonomy; Vege: Vegetation type.

the division of classes will give quite different results. For some categorized factors, such as soil taxonomy and vegetation type, the AUC method originally gave quite low importance (Figure 9(e)). The renumbering adjustment produces a monotonic relationship between the factor class number and favorability, thus, a higher AUC value (e.g. Figure 8(k, o)). After the renumbering adjustment, the AUC values of all categorized factors increased, making the order of absolute importance more similar to that of decisiveness (Figure 9(a, f)). However, this renumbering adjustment can only be applied to the categorized factors. For continuous factors, if the factor value is far from monotonously correlated with the favorability value (e.g. Figure 4(h), Figure 8(h)), the AUC method will give fallaciously low importance, for example, average annual precipitation (Figure 8(h), Figure 9 (f)). Therefore, the accountability & reliability and AUC methods are still not robust, even though some of their inherent defects can be moderated.

The geographical detector method is still complicated by prior factor classification for continuous factors. Although optimization approaches have been proposed for factor discretization (e.g. Zhang, Song, and Wu 2022), it is still uncertain whether other possible combinations of class break-points will yield better results. When the same factor classifications are used, the q value gives a factor importance slightly different from the adjusted Acc&Rel index (Figure 9(d, g)). Similar to the AUC method, the geographical detector method gives relatively low importance for average annual precipitation, which has an obvious non-monotonic correlation with landslide presence. Another advantage of the decisiveness method is that the variation of importance with bin width (Figure 6) is much more resolved than that with class count.

As decisiveness varies with bin width and class count (Figures 6 and 7), when comparing decisiveness among continuous and categorized factors, this study suggests applying for a classification compensation; that is, first compensating the influence of class count for the categorized factors and then adopting a bin width for the continuous factors comparable to the class count of the categorized factors. The case study showed that for all categorized factors the decisiveness after random class integrations (i.e. after classification compensation) was slightly lower than the original (Table 2), consistent with theoretical expectations. Tuning down the class count and thus the decisiveness of factors with larger class counts, to enable comparability of decisiveness among categorized factors based on the same class count, is exactly the spirit of classification compensation. The change in decisiveness before and after classification compensation is minor because the class counts of factors are close (Table 2), whereas different hierarchical data exhibit a significant difference in decisiveness due to large differences in class count (Figure 7). It should be emphasized that the criterion of susceptibility optimization can be adopted if decisiveness is required to reflect landslide susceptibility results. In this case study, the susceptibility optimization and classification compensation criteria gave identical rankings of factors based on decisiveness (Figure 9).

4.3. Considerations in practice

The decisiveness importance of several landslide predisposing factors was obtained in the case study. Nevertheless, it is important to acknowledge that different rankings of factor absolute importance may be obtained in other cases. For example, the case study showed that elevation has a higher decisiveness than slope, which was consistent across all methods (Figure 9), indicating the reality of this observation. The extremely low decisiveness of slope can be clearly explained by the frequency distribution of certainty factor values (Figure 5). Most certainty factor values for slope are concentrated around 0 (Figure 5(b)), indicating that the conditional probabilities of landslides for most slope angles are neither high nor low, further suggesting a low power of determinant for landslides. In contrast, the certainty factor values for elevation are skewed around – 1 (Figure 5(a)), indicating that for many elevation values, there is a high level of confidence that the occurrence probability of landslides is very low (particularly at altitudes higher than 4500 M, Figure 4(a)), thereby contributing to a high power of determinant. Although higher importance of elevation than slope had been observed in other studies (e.g. Pourghasemi and Rahmati 2018), higher importance of slope than elevation was also reported (e.g. Bragagnolo, da Silva, and Grzybowski 2020; Youssef and Pourghasemi 2021). Contradictory importance rankings of elevation and slope can be obtained in different study areas even by the same author (e.g. Pourghasemi and Rahmati 2018; Youssef and Pourghasemi 2021). Similarly, we cannot guarantee soil taxonomy and geomorphologic units will get high decisiveness values in other cases. Therefore, it is suggested to conduct more practical applications of the decisiveness index to gain a more comprehensive picture of the importance ranking of various factors.

Another issue that should be considered in practice is the prior selection of predisposing factors for importance analysis, and we suggest three criteria. (1) Frequency of investigation. To select those factors most frequently investigated in previous studies, by referring to factor use frequencies either reported in classic review papers (e.g. Reichenbacha et al. 2018) or obtained on our own. (2) Relevance to landslide occurrence. To select those factors most relevant to landslide occurrence in the study area, which can be identified according to expert knowledge or theoretical analysis, with a focus on factors that are particularly influential in specific situations or geographic contexts. (3) Data availability and quality. To select those factors for which data of acceptable quality are available so that a data-driven quantitative measurement of factor importance can be conducted effectively and reliably.

Bias originating from the used data is a common issue for any data-driven approach, including the methodology proposed in this paper. This bias can be related to data on both landslides and predisposing factors. The incompleteness or bias of landslide inventory is a major problem that could limit the reliability of derived factor importance. For example, if field reconnaissance of landslides was conducted along or nearby roads where access is easy, then distance to road is expected to be a factor of high importance, which may not be true if the landslide dataset is completed using more extensive inventory methods such as remote sensing. In addition, using landslide headscarp areas and total affected landslide areas may yield different results. Sometimes, the importance of a factor in constraining landslide spatial distribution becomes clearer with more precise and relevant factor data (e.g. Yao et al. 2021). Therefore, we emphasize the limitation of the proposed methodology due to the bias of data in practice. If there are results that lack physically rational explanations, bias in the used data may be the 'suspect'.

5. Conclusions

In this paper, an analytical index for measuring the absolute importance of landslide predisposing factors – 'decisiveness,' is proposed. The principle of decisiveness is that factor importance can be evaluated according to the ability to differentiate between landslide-susceptible and landslide-insusceptible areas. A factor of high importance for landslides is the ability to zone the target area into units that are either highly favorable or highly unfavorable to landslides. For a specific factor, the favorability values of all grid cell units in the target area are first calculated using the certainty factor method, and the magnitudes of all favorability values are then integrated to constitute the decisiveness index. Decisiveness ranges within [0, 1]. The higher the decisiveness, the more important is the factor, and vice versa. In addition, the updated open software '*Automatic Landslide Susceptibility Analysis* (ALSA)' V4.0 and V5.0 can be used to calculate favorability values and further decisiveness for individual landslide predisposing factors with either categorized or continuous values.

A case study showed that the proposed decisiveness reflects the ability of landslide predisposing factors to differentiate between landslide-susceptible and landslide-insusceptible areas, thus measuring the absolute importance of the factors. The absolute importance of factors measured by existing relevant methods, that is, the accountability & reliability, AUC, and geographical detector methods, are still complicated by inherent difficulties, although some effective moderations can be applied. Decisiveness is not accompanied by these inherent complications thus introducing a robust measure of absolute importance.

It is noteworthy that with this novel factor absolute importance index (decisiveness), not only can the importance of different factors in a single application case be evaluated, but also the importance of the same factor in different application cases can be evaluated and compared. The importance of a factor in determining the spatial distribution of landslides could differ in different areas or different triggering events. In addition, although decisiveness is proposed to measure the absolute importance of landslide predisposing factors, the idea and this novel index can be used to analyze factor importance in other similar application scenarios that are not relevant to landslides, particularly for those evaluating the importance of factors in determining the spatial distribution of an object or phenomenon.

Note

1. Please contact the authors or visit https://github.com/lilangping/alsa/.

Acknowledgements

The authors would like to thank NCSFGI (National Catalogue Service for Geographic Information), RESDC (Resource and Environment Science and Data Center), and the Shuttle Radar Topography Mission, for providing free-access data. Ms. Wenxuan HUANG assisted with the literature review. Dr. Chengdong XU is appreciated for discussing the application of the geographical detector method. Dr. Ci SONG is particularly appreciated for discussing the permutation and combination problem. This study is a contribution to the Deep-time Digital Earth (DDE) Big Science Program.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Natural Science Foundation of China [42177150, 42041006], the Strategic Priority Research Program of Chinese Academy of Sciences (CAS) [XDA23090301], and the Key Project of Innovation LREIS [KPI007].

Data and codes availability

The data that support the findings of this study are available upon reasonable request. The open software ALSA presented in this study is available at: https://github.com/lilangping/alsa/.

ORCID

Langping LI http://orcid.org/0000-0001-6398-3161
Hengxing LAN http://orcid.org/0000-0002-7146-932X

References

Barella, C. F., F. G. Sobreira, and J. L. Zêzere. 2019. "A Comparative Analysis of Statistical Landslide Susceptibility Mapping in the Southeast Region of Minas Gerais State, Brazil." *Bulletin of Engineering Geology and the Environment* 78: 3205–3221. https://doi.org/10.1007/s10064-018-1341-3.

Blahut, J., C. J. van Westen, and S. Sterlacchini. 2010. "Analysis of Landslide Inventories for Accurate Prediction of Debris-Flow Source Areas." *Geomorphology* 119: 36–51. https://doi.org/10.1016/j.geomorph.2010.02.017.

Bragagnolo, L., R. V. da Silva, and J. M. V. Grzybowski. 2020. "Landslide Susceptibility Mapping with *r.Landslide*: A Free Open-Source GIS-Integrated Tool Based on Artificial Neural Networks." *Environmental Modelling & Software* 123: 104565. https://doi.org/10.1016/j.envsoft.2019.104565.

Castellanos Abella, E. A. 2008. "Multi-Scale Landslide Risk Assessment in Cuba." PhD diss., University of Utrecht. https://dspace.library.uu.nl/handle/1874/30890.

Donati, L., and M. C. Turrini. 2002. "An Objective Method to Rank the Importance of the Factors Predisposing to Landslides with the GIS Methodology: Application to an Area of the Apennines (Valnerina; Perugia, Italy)." *Engineering Geology* 63: 277–289. https://doi.org/10.1016/S0013-7952(01)00087-4.

Greenbaum, D., M. Tutton, M. R. Bowker, T. J. Browne, J. Buleka, K. B. Greally, G. Kuna, et al. 1995. *Rapid Methods of Landslide Hazard Mapping: Papua New Guinea Case Study*. Nottingham, UK: British Geological Survey. https://nora.nerc.ac.uk/id/eprint/9967/.

Gullà, G., M. Conforti, and L. Borrelli. 2021. "A Refinement Analysis of the Shallow Landslides Susceptibility at Regional Scale Supported by GIS-Aided Geo-Database." *Geomatics, Natural Hazards and Risk* 12: 2500–2543. https://doi.org/10.1080/19475705.2021.1967204.

Guo, Z., F. Guo, Y. Zhang, J. He, G. Li, Y. Yang, and X. Zhang. 2023. "A Python System for Regional Landslide Susceptibility Assessment by Integrating Machine Learning Models and Its Application." *Heliyon* 9: e21542. https://doi.org/10.1016/j.heliyon.2023.e21542.

He, Q., M. Wang, and K. Liu. 2021. "Rapidly Assessing Earthquake-Induced Landslide Susceptibility on a Global Scale Using Random Forest." *Geomorphology* 391: 107889. https://doi.org/10.1016/j.geomorph.2021.107889.

Heckerman, D. 1986. "Probabilistic Interpretations for Mycin's Certainty Factors." *Machine Intelligence and Pattern Recognition* 4: 167–196. https://doi.org/10.1016/B978-0-444-70058-2.50017-6.

Kavzoglu, T., and A. Teke. 2022. "Ensemble Conditioning Factor Selection with Markov Chain Framework for Shallow Landslide Susceptibility Mapping in Lake Sapanca Basin and Its Vicinity, Turkey." *Baltic Journal of Modern Computing* 10: 224–240. https://doi.org/10.22364/bjmc.2022.10.2.09.

Lan, H., F. Wu, C. Zhou, and L. Wang. 2003. "Spatial Hazard Analysis and Prediction on Rainfall-Induced Landslide Using GIS." *Chinese Science Bulletin* 48: 703–708. https://doi.org/10.1007/BF03325659.

Lan, H., C. Zhou, L. Wang, H. Zhang, and R. Li. 2004. "Landslide Hazard Spatial Analysis and Prediction Using GIS in the Xiaojiang Watershed, Yunnan, China." *Engineering Geology* 76: 109–128. https://doi.org/10.1016/j.enggeo.2004.06.009.

Li, L. P., and H. X. Lan. 2023. "Bivariate Landslide Susceptibility Analysis: Clarification, Optimization, Open Software, and Preliminary Comparison." *Remote Sensing* 15: 1418. https://doi.org/10.3390/rs15051418.

Li, L. P., H. X. Lan, C. B. Guo, Y. S. Zhang, Q. W. Li, and Y. M. Wu. 2017. "A Modified Frequency Ratio Method for Landslide Susceptibility Assessment." *Landslides* 14: 727–741. https://doi.org/10.1007/s10346-016-0771-x.

Ma, Z., G. Mei, and F. Piccialli. 2021. "Machine Learning for Landslides Prevention: A Survey." *Neural Computing and Applications* 33: 10881–10907. https://doi.org/10.1007/s00521-020-05529-8.

Ma, F., J. Wang, R. Yuan, H. Zhao, and J. Guo. 2013. "Application of Analytical Hierarchy Process and Least-Squares Method for Landslide Susceptibility Assessment Along the Zhong-Wu Natural Gas Pipeline, China." *Landslides* 10: 481–492. https://doi.org/10.1007/s10346-013-0402-8.

Meng, Y., H. Lan, L. Li, Y. Wu, and Q. Li. 2015. "Characteristics of Surface Deformation Detected by X-Band SAR Interferometry Over Sichuan-Tibet Grid Connection Project Area, China." *Remote Sensing* 7: 12265–12281. https://doi.org/10.3390/rs70912265.

Pourghasemi, H. R., and O. Rahmati. 2018. "Prediction of the Landslide Susceptibility: Which Algorithm, Which Precision?" *Catena* 162: 177–192. https://doi.org/10.1016/j.catena.2017.11.022.

Qi, S., Q. Xu, H. Lan, B. Zhang, and J. Liu. 2010. "Spatial Distribution Analysis of Landslides Triggered by 2008.5.12 Wenchuan Earthquake, China." *Engineering Geology* 116: 95–108. https://doi.org/10.1016/j.enggeo.2010.07.011.

Reichenbacha, P., M. Rossi, B. D. Malamud, M. Mihir, and F. Guzzetti. 2018. "A Review of Statistically-Based Landslide Susceptibility Models." *Earth-Science Reviews* 180: 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001.

RESDC (Resource and Environment Science and Data Center). 2020. "Geological Disaster Dataset of China." Accessed October 29, 2020. https://www.resdc.cn/data.aspx?DATAID=290.

Shortliffe, E. H., and B. G. Buchanan. 1975. "A Model of Inexact Reasoning in Medicine." *Mathematical Biosciences* 23: 351–379. https://doi.org/10.1016/0025-5564(75)90047-4.

Torizin, J., N. Schüßler, and M. Fuchs. 2022. "Landslide Susceptibility Assessment Tools v1.0.0b – Project Manager Suite: A New Modular Toolkit for Landslide Susceptibility Assessment." *Geoscientific Model Development* 15: 2791–2812. https://doi.org/10.5194/gmd-15-2791-2022.

Wang, Z. H., Z. W. Hu, W. J. Zhao, H. L. Gong, and J. X. Deng. 2014. "Susceptibility Analysis of Precipitation-Induced Landslide Disaster-Pregnant Environmental Factors Based on the Certainty Factor Probability Model—Taking the Hilly Area in Sichuan as Example (in Chinese with English Abstract)." *Journal of Catastrophology* 29 (2): 109–115.

Wang, J. F., X. H. Li, G. Christakos, Y. L. Liao, T. Zhang, X. Gu, and X. Y. Zheng. 2010. "Geographical Detectors-Based Health Risk Assessment and Its Application in the Neural Tube Defects Study of the Heshun Region, China." *International Journal of Geographical Information Science* 24: 107–127. https://doi.org/10.1080/13658810802443457.

Wang, H., L. Zhang, H. Luo, J. He, and R. W. M. Cheung. 2021. "AI-Powered Landslide Susceptibility Assessment in Hong Kong." *Engineering Geology* 288: 106103. https://doi.org/10.1016/j.enggeo.2021.106103.
Yang, Z., H. Lan, X. Gao, L. Li, Y. Meng, and Y. Wu. 2015. "Urgent Landslide Susceptibility Assessment in the 2013 Lushan Earthquake-Impacted Area, Sichuan Province, China." *Natural Hazards* 75: 2467–2487. https://doi.org/10.1007/s11069-014-1441-8.
Yang, Y., J. T. Yang, C. D. Xu, C. Xu, and C. Song. 2019. "Local-scale Landslide Susceptibility Mapping Using the B-GeoSVC Model." *Landslides* 16: 1301–1312. https://doi.org/10.1007/s10346-019-01174-y.
Yang, Q., D. Zheng, and Y. Liu. 1988. "Physico-geographical Feature and Economic Development of the Dry Valleys in the Hengduan Mountains, Southwest China (in Chinese with English Abstract)." *Journal of Arid Land Resources and Environment* 2: 17–24. https://doi.org/10.13448/j.cnki.jalre.1988.02.003.
Yao, J., H. Lan, L. Li, Y. Cao, Y. Wu, Y. Zhang, and C. Zhou. 2022. "Characteristics of a Rapid Landsliding Area Along Jinsha River Revealed by Multi-Temporal Remote Sensing and Its Risks to Sichuan-Tibet Railway." *Landslides* 19: 703–718. https://doi.org/10.1007/s10346-021-01790-7.
Yao, X., S. Qi, C. Liu, S. Guo, X. Huang, C. Xu, B. Zheng, Z. Zhan, and Y. Zou. 2021. "An Empirical Attenuation Model of the Peak Ground Acceleration (PGA) in the Near Field of a Strong Earthquake." *Natural Hazards* 105: 691–715. https://doi.org/10.1007/s11069-020-04332-x.
Youssef, A. M., and H. R. Pourghasemi. 2021. "Landslide Susceptibility Mapping Using Machine Learning Algorithms and Comparison of Their Performance at Abha Basin, Asir Region, Saudi Arabia." *Geoscience Frontiers* 12: 639–655. https://doi.org/10.1016/j.gsf.2020.05.010.
Zhang, Y., H. Lan, L. Li, Y. Wu, J. Chen, and N. Tian. 2020. "Optimizing the Frequency Ratio Method for Landslide Susceptibility Assessment: A Case Study of the Caiyuan Basin in the Southeast Mountainous Area of China." *Journal of Mountain Science* 17: 340–357. https://doi.org/10.1007/s11629-019-5702-6.
Zhang, Z. H., Y. Z. Song, and P. Wu. 2022. "Robust Geographical Detector." *International Journal of Applied Earth Observation and Geoinformation* 109: 102782. https://doi.org/10.1016/j.jag.2022.102782.
Zhao, Z., J. Chen, K. Xu, H. Xie, X. Gan, and H. Xu. 2021. "A Spatial Case-Based Reasoning Method for Regional Landslide Risk Assessment." *International Journal of Applied Earth Observation and Geoinformation* 102: 102381. https://doi.org/10.1016/j.jag.2021.102381.
Zhao, J., Q. Zhang, D. Wang, W. Wu, and R. Yuan. 2022. "Machine Learning-Based Evaluation of Susceptibility to Geological Hazards in the Hengduan Mountains Region, China." *International Journal of Disaster Risk Science* 13: 305–316. https://doi.org/10.1007/s13753-022-00401-w.