A rapid assessment method for earthquake-induced landslide casualties based on GIS and logistic regression model

Yuqian Dai, Xianfu Bai, Gaozhong Nie and Gang Huang

ABSTRACT
The accuracy of rapid earthquake assessment and the emergency assessment system for earthquake-induced damages could be substantially enhanced if the casualties triggered by earthquake-induced geological disasters, such as landslides, are subjected to comprehensive scientific evaluation. However, no credible solution for this purpose has been formulated yet. This study suggests a three-step rapid assessment method designed for earthquake-induced landslide casualties based on the GIS and an associated logistic regression model, as follows: (1) Partition of the region to be evaluated as a 1 km × 1 km grid in the GIS, with assignment of a certain amount of population to each of the grid cells as its population attribute. (2) Calculation of the death rate for each grid cell based upon its earthquake-induced landslide susceptibility attribute using the logistic regression model. (3) The earthquake-induced landslide casualties are first determined for each of the kilometer grid cells, and then for the entire region under evaluation. The proposed method was implemented to test the assessment of earthquake-induced landslide casualties in three earthquake-stricken regions. The study reveals the feasibility of the extensibility and applicability of the proposed rapid assessment method for earthquake-induced landslide casualties, and its suitability for similar assessments and calculations of other regions.

1. Introduction
In mountainous and canyon areas, earthquakes are often followed by severe landslides (Pain 1972). Landslides are known to be the key hazards stimulating earthquake-induced fatalities, since they generally impact the residential communities residing in...
Table 1. Casualties caused by some earthquake-induced landslides in China.

| Date            | Place                    | Magnitude ($M_s$) | Brief description                                                                 | Earthquake-induced landslide casualties |
|-----------------|--------------------------|-------------------|-----------------------------------------------------------------------------------|----------------------------------------|
| March 26, 26 BCE| Jianwei Prefecture       | 5 1/2             | Landslides occurred at both banks of Baijiang River and Juanjiang River            | 13                                     |
| August 2, 1733  | Dongchuan in Yunnan      | 7 3/4             | Hills collapsed and there was rock fallat Bigu, Ahwang, and Xiaojiang in the area  | 40                                     |
| June 7, 1789    | Huaning in Yunnan        | 7.0               | Yidu Village fell into the lake                                                   | Countless                              |
| September 6, 1833| Songming in Yunnan      | 8.0               | Massive landslides occurred in many earthquake-stricken places                    | At least dozens                        |
| August 1844     | Daguan in Yunnan         | ≥5                | Hills and houses fell at the border between Daguan and Yongshan                    | 30                                     |
| October 17, 1911| Qiaojia in Yunnan        | 5 3/4             | Cliffs collapsed in Menggu, Qiaojia, and rocks rolled down in Baiwujie, Huize      | 10                                     |
| July 31, 1917   | Daguan in Yunnan         | 6 3/4             | The Jingguo Army in Yunnan marched toward Sichuan, but encountered the earthquake at Liwupu, and was entirely buried by the unexpected collapse of hills. | 1850                                   |
| December 16, 1920| Haiyuan in Ningxia       | 8 1/2             | Landslides buried a lot of houses and cave dwellings.                             | Unclear                                |
| June 7, 1932    | Gejiu in Yunnan          | 5.0               | Collapses occurred in Fengzidong, Laochang.                                       | 9                                      |
| August 25, 1933 | Diexi in Sichuan         | 7 1/2             | Tremendous landslides were caused, and Diexi Town was buried under the two collapsed hills. | More than 500                         |
| May 13, 1941    | Lushui in Yunnan         | 5 1/2             | Hills and rocks collapsed in Yidiluo, Maozhao Town.                               | 3                                      |
| May 16, 1941    | Gengma in Yunnan         | 7.0               | Boulders rolled down.                                                             | 2                                      |
| October 10, 1948| Daguan in Yunnan         | 5 3/4             | Nearly 40 m tall rocks collapsed in Houziyan, Weining, Guizhou.                   | 3                                      |
| August 15, 1952 | Motuo in Tibet           | 8.5               | The mountains and rivers were moved, and the terrain was changed. Many peaks collapsed and blocked the Yarlung Zangbo River. Landslides caused five villages and temples to be pushed into the river. | 500                                    |
| April 22, 1973  | Yiliang in Yunnan        | 5.1               | Rocks rolled down from hills in Yankou, Yegu, and Xiaohe in the southeast of Yiliang. | 2                                      |
| August 2, 1973  | Yiliang in Yunnan        | 5.4               | Tons of rocks fell and landslides occurred.                                      | 12                                     |
| May 11, 1974    | Daguan in Yunnan         | 7.1               | Severe landslides occurred in Erpingzi, Shoubayan, Sutianba, and some other places in Yongshan. Shoubayan Village was entirely buried, and all the villagers died due to the earthquake. | 73                                     |

(continued)
The Great Alaskan Earthquake in the United States resulted in 130 deaths, including 48 fatalities due to earthquake-induced landslides (Keeper 1984). From 1969 to 1993, for over two decades, more than half the deaths observed during earthquakes exceeding Ms6.9 in Japan were ascribed to earthquake-induced landslides.

| Date                     | Place            | Magnitude \((M_s)\) | Brief description                                                                 | Earthquake-induced landslide casualties |
|--------------------------|------------------|----------------------|----------------------------------------------------------------------------------|-----------------------------------------|
| November 7, 1976         | Yanyuan in Sichuan | 6.7                  | Rocks collapsed at the upstream of Mofanggou Creek, Ninglang.                    | 2                                       |
| September 25, 1996       | Lijiang in Yunnan | 5.7                  | Earthquake-induced landslides left one person missing.                           | 1                                       |
| July 22 and August 25, 2006 | Yanjin in Yunnan | 5.1                  | Two earthquakes of Ms 5.1 occurred consecutively. The first earthquake resulted in 22 deaths including 18 persons hit by the rolling rocks caused by earthquake-induced landslides. The second earthquake caused 2 deaths because of earthquake-induced landslides. | 20                                      |
| May 12, 2008             | Wenchuan in Sichuan | 7.9                  | Massive landslides induced by the earthquake resulted in approximately 20,000 dead and missing persons. | Approximately 20,000                   |
| April 14, 2010           | Yushu in Qinghai | 7.1                  | More than 3,000 casualties. Massive landslides occurred after the earthquake, and directly impacted the residential communities. | 270                                     |
| September 7, 2012        | Yiliang in Yunnan | 5.7                  | Most of the 81 deaths were caused by rolling rocks and landslides.               | 59                                      |
| April 20, 2013           | Lushan in Sichuan | 7.0                  | Landslides occurred in some mountainous areas after the earthquake and impacted the villages. | 24                                      |
| August 3, 2014           | Ludian in Yunnan | 6.5                  | Massive landslides occurred. Some villages located on hillsides slid down with the landslides, causing a great deal of casualties. | 250                                     |
| August 8, 2017           | Jiuzhaigou in Sichuan | 7.0                  | Few deaths were attributed to the destruction of buildings, but most of the casualties were caused by the geological disasters induced by the earthquake. | Approximately 20                       |
| May 18, 2020             | Qiaojia in Yunnan | 5.0                  | The earthquake caused some hill collapses. Two persons were hit by falling rocks and then died. | 2                                       |
| May 21, 2021             | Yangbi in Yunnan | 6.4                  | The earthquake-induced partial landslides and rolling rocks. One person was hit by the rolling rocks on the highway, and then died. | 1                                       |
The 1970 Chimbote earthquake of Ms7.6 in Peru resulted in the most catastrophic earthquake-induced landslides known till date, leaving at least 20,000 individuals dead (Plafker et al. 1971). Earthquake-induced landslides are also a common phenomenon in mainland China, especially in the mountainous regions of West China. Table 1 summarizes the majority of the strong earthquakes in China that have caused devastating landslides leading to substantial casualties (Bai et al. 2013).

Historical statistics reveal that earthquake-induced landslide casualties account for about 25%–30% of the total casualties associated with earthquakes in the western mountainous regions of mainland China. For some earthquakes, the fatalities caused by earthquake-induced landslides are even greater than those caused by the destruction of buildings (Gao et al. 2011a). Additionally, earthquake-induced landslides could at times be the solitary cause for casualties for earthquakes of lower magnitudes resulting in one or few fatalities. Evidently, earthquake-induced landslides have emerged as a major cause of concern and hidden earthquake hazards for the people living in the mountainous areas of West China. The Seismic Ground Motion Parameters Zonation Map of China (GB18306-2015), indicates that the regions of western China are still at a high risk of earthquake-induced landslide casualties. To strengthen the accuracy of rapid earthquake casualty assessment method for the mountainous areas of West China, not only should the assessment techniques adopted for the casualties resulting from the destruction of buildings be further advanced and enhanced, but also the assessment methods for the casualties resulting from the earthquake-induced geological disasters(including earthquake-induced landslides) should be established and formulated at the earliest (Bai et al. 2014). Scientific analysis of fatalities from earthquake-induced geological hazards has emerged as a reliable and effective technique for enhancing the accuracy of rapid assessment methods concerning earthquake casualties, thereby strengthening the robustness of the emergency assessment system for earthquake-induced damages. At the same time, risk assessment procedures of earthquake-induced disasters and high accuracy post-earthquake rapid assessment still face many technical challenges.

The rapid assessment of earthquake-induced landslide casualties is conducted in two stages: The first stage concerns post-earthquake susceptibility assessment of earthquake-induced landslides, where the aim is to understand and figure out the kind of earthquakes that can cause different landslides, as well as the regions in which these landslides may occur. Several researchers have extensively explored and analyzed the assessment of earthquake-induced landslide susceptibility for years, and subsequently devised several methods for this purpose. The qualitative assessment has gradually transformed into semi-quantitative or wholly quantitative assessment (Chen et al. 2011; Yuan et al. 2013; Liu et al. 2016; Xu et al. 2019). In the second stage, the possibility of the post-earthquake generated landslides being potent enough to cause casualties and the likely number of casualties they may cause is assessed, i.e., hazard assessment of earthquake-induced landslides. Earthquake-related hazard statistics in China have historically concentrated on the casualties brought about by the damage to buildings. The fatalities resulting from earthquake-induced landslides were surveyed only occasionally, that too without the recording of any vital details, such as
specific location. The casualties could be evaluated by exploring their correlation with seismic parameters, and damages caused by earthquake to buildings. However, due to inadequate historical information about seismic parameters, such as the location of the landslide, location of the residential community, and earthquake-induced landslide casualties, it became difficult to conduct any inclusive research. Deficient historical records led to the undertaking of only a few research publications on the hazards of earthquake-induced landslides. Thus, for a long time, the prediction of earthquake-induced landslide casualties has by and large relied only on the field expertise with respect to the rapid assessment of earthquake-induced losses or the preliminary loss assessment in key hazardous areas, thereby its accuracy could not be guaranteed. This paper is an initiative to address and further explore this issue.

This paper is organized in five sections. In section 2, we present the research ideas and propose a rapid assessment method for earthquake-induced landslide casualties based on GIS and logistic regression model. Section 3, we use 3 cases to further evaluate the validity of the method. In section 4, we offer a discussion of methodology and results. Finally, we sum up the main conclusions of our research.

2. Research ideas and methods

2.1. Basic idea

Earthquake-induced landslide casualties are primarily and directly attributed to people being buried by landslides or being hit by rock fall. Normally, these casualties are affected by four factors, namely, scale of the earthquake hazard, earthquake-induced landslide susceptibility, population exposure, and human behavior. The scale of the earthquake hazard is often indicated by parameters such as earthquake magnitude, seismic intensity, focal depth, and peak ground acceleration (PGA). Of these parameters, seismic intensity is often taken as the basis for estimating earthquake disaster losses and damages to buildings in China, since it could reflect the post-earthquake damages more accurately, and could be immediately evaluated as soon as the earthquake occurred. Earthquake-induced landslide susceptibility reflects the possibility of landslide under the effect of ground motion during an earthquake, which can be characterized by landslide density per unit area, cell landslide probability, and landslide susceptibility level. In China, earthquake-induced landslide susceptibility level is preferred in the rapid assessment of earthquake disasters for its simple description and convenient calculation. The exposure of the population to this hazard is crucial to evaluating earthquake-induced landslide casualties. Using the same metric for earthquake hazard and earthquake-induced landslide susceptibility, there is a higher chance of casualties being triggered by larger landslides in a more populated region. In a depopulated region, no casualty is caused by any earthquake-induced landslide. Population exposure indirectly reflects the attractiveness of a place to people at different times. Under normal circumstances in China, it is represented by the number of people in a kilometer grid. Human behavior, however, has a complicated influence on the magnitude of earthquake-induced landslide casualties, as it involves numerous phenomena such as the agency of people in avoiding earthquake-induced landslides by choice, and the ability of people to mutually rescue each other. For a certain
severity of earthquake disaster, the earthquake-induced landslide susceptibility remains unchanged at a place, whereas the population exposure, and accordingly human behavior, appear to be highly uncertain. In the early stage of exploring the rapid assessment method for earthquake-induced landslide casualties, it is therefore necessary to focus on these relatively certain factors. Further research needs to be carried out to gradually identify the influence of those factors that are uncertain. For this reason, the influence of the uncertain factors is addressed by using some common data (e.g., population exposure) or is simply ignored (e.g., human behavior) in this study.

Presently, earthquake-induced losses in mainland China are assessed promptly as per the following procedure: After an earthquake occurs, the assessor receives three key data points about the earthquake (time, epicenter, and magnitude); from these the assessor acquires the intensity assessment map of the earthquake by working through the calculation based on the regional seismic intensity attenuation model. Subsequently, the impairments including house damage, economic losses, and casualties are quickly assessed for various intensity regions in accordance with the intensity assessment map. In the end, the rapid assessment results are generated together with the recommended emergency response measures in reference to the intensity assessment map. In this rapid assessment, landslide susceptibility is often assumed to represent the possibility of a landslide. Houses bearing an identical structural type in an intensity region, are more or less, damaged in a similar manner, thus the death rate appears practically consistent. Nevertheless, the earthquake landslide susceptibility varies for different places even in a single intensity region, resulting in different death rates. For this reason, death rate is not assumed to be consistent throughout an intensity region because of the earthquake-induced landslides. In other words, the assessment of earthquake-induced landslide casualties should not directly assume the whole of the intensity region as its basic grid cell. For instance, let us say the population of a region $R$ (Figure 1a) is $P$, and the total deaths caused by earthquake-induced landslides are denoted by $D$. The region $R$ is partitioned into $N$ regular kilometer grids $R_i$ (Figure 1b). The population in the $i^{th}$ kilometer grid $R_i$ is $P_i$, and the earthquake-induced landslide casualties are $D_i$. The sum of the populations in these smaller kilometer grids equal the total population of the whole region $R$, i.e., $\sum_{i=1}^{N} P_i = P$. The sum of earthquake landslide casualties occurring in these kilometer grids equal the total of the earthquake-induced landslide casualties of the whole region $R$,
Therefore, in this manner, the assessment of earthquake landslide casualties for the entire region can be transformed into the calculation of total deaths occurring in the kilometer grids. The earthquake landslide casualties in the region $R$ are the product of the population in the region and the death rate of earthquake-induced landslides, i.e., $D = P \times M$, assuming that the death rate of earthquake-induced landslides in the region $R$ is $M$. Similarly, if the death rate of earthquake-induced landslides in the $i$th kilometer grid is $M_i$, the relationship of $P_i$, $D_i$, and $M_i$ is $D_i = P_i \times M_i$, which is substituted into $\sum_{i=1}^{N} D_i = D$ to obtain $\sum_{i=1}^{N} D_i = \sum_{i=1}^{N} P_i \times M_i = D$. However, the heterogeneity of the geographical environment can be attributed to the fact that the death rate of earthquake-induced landslides in the region $R$ normally appears different from the death rate $M_i$ in each kilometer grid $R_i$, or the sum of the death rates in these kilometer grids, i.e., $M_i \neq M$ and $\sum_{i=1}^{N} M_i \neq M$. The kilometer grids possess different death rates as well, i.e., $M_i \neq M_{i+1}$. The deaths in each kilometer grid should instead be determined by virtue of $D_i = P_i \times M_i$ to evaluate the earthquake-induced landslide casualties for the entire region. For this purpose, a calculation method for calculating the death rate of earthquake-induced landslides based on the kilometer grid needs to be developed. If each kilometer grid is evenly populated, the death rate effectuated by the earthquake-induced landslides in the kilometer grid is primarily dependent on the landslide susceptibility in the grid cell. Hence, the death rate of earthquake-induced landslides depends on the distribution of landslide susceptibility attributes and the number of cells in the grid.

Through the above deduction, this paper recommends a rapid assessment method for earthquake-induced landslide casualties predicated on the geographic information system (GIS) and an associated logistic regression model. The fundamental idea of the method (see Figure 2) is as follows: The death rate of earthquake-induced landslides in each kilometer grid relates to the attribute of landslide susceptibility in that specific grid and the total number of cells. In conformity with this framework, a model is devised for calculating the death rate of earthquake-induced landslides in the kilometer grid. The earthquake-induced landslide casualties in the evaluated region are equal to the sum of earthquake-induced landslide casualties in the kilometer grids of the region.

![Figure 2. Basic idea for the assessment of earthquake-induced landslide casualties.](image-url)
In the proposed method for the assessment of earthquake-induced landslide casualties, it is imperative to develop a model for the death rate of earthquake-induced landslides for each kilometer grid individually. However, the earthquake-induced landslides influencing the death rate in a kilometer grid are represented by the earthquake-induced landslide susceptibility, since it is quite a challenge to accurately predict the frequency of occurrence of landslides at a specific location. As an influencing factor, earthquake-induced landslide susceptibility is correlated with the death rate, illustrated by the linear combination formed by different landslide susceptibility attributes and the number of cells in the kilometer grid. The unknown parameters in the linear combination are figured out by virtue of logistic regression to build the model for the death rate of earthquake-induced landslides in accordance with the kilometer grid. Based on this, the population in the kilometer grid is summarized to eventually construct the model for the assessment of earthquake-induced landslide casualties based on the kilometer grid. In this paper, the GIS is implemented to count the number of cells with different landslide susceptibility attributes in each kilometer grid and to calculate the earthquake landslide casualties and the death rate in the kilometer grid. These parameters are obtained to define the correlation between the death rate of earthquake-induced landslides and the number of cells with different landslide susceptibility attributes in the kilometer grid:

Figure 3. Rapid assessment process of earthquake-induced landslide casualties.
where $M_i$ represents the death rate of earthquake-induced landslides in the $i$\textsuperscript{th} kilometer grid; $f()$ indicates the inferred relationship between the influencing factor and $M_i$; and $X_{il}$ denotes the number of cells with different landslide susceptibility attributes in the $i$\textsuperscript{th} kilometer grid.

Following this basic idea, three steps are undertaken to develop the rapid assessment method for earthquake-induced landslide casualties in each kilometer grid (Figure 3): (1) The evaluated region under the GIS is partitioned into suitable 1 km × 1 km grids. The population in each grid is assigned as its population attribute. (2) The death rate of earthquake-induced landslides in each grid is calculated. With the spatial analysis feature of the GIS, the earthquake-induced landslide susceptibility level of the evaluated region is extracted from the prediction data set. The number of cells for each landslide susceptibility level in the kilometer grid is calculated as per the landslide susceptibility attribute of the grid. Based on the landslide susceptibility attribute, the logistic regression model is adopted to calculate the death rate of earthquake-induced landslides in the kilometer grid. (3) The distribution of earthquake-induced landslide casualties in the kilometer grid and the entire evaluated region is assessed.

### 2.2. Research methods

#### 2.2.1. Assignment of kilometer grid population attribute

The evaluated region is partitioned into appropriate 1 km × 1 km grid cells. The population in each kilometer grid cell is assigned as the population attribute of the grid cell. This paper adopts the data set of kilometer grid population in mainland China, which was jointly developed by the Institute of Geology, China Earthquake Administration, the Institute of Geographic Sciences and Natural Resources Research, and Chinese Academy of Sciences. To ensure the assessment results are as accurate as the actual condition, the China Earthquake Administration and other departments undertaking the rapid earthquake assessment task must update the basic database on an annual basis. This requires not only a specific system but also corresponding funding. In other words, the fund for updating the data is appropriated by the government fiscal every year. So, the applicability of the data set has been verified (Chen et al. 2012; Ding et al. 2014; An et al. 2015), and the data exhibits an effective update mechanism.

#### 2.2.2. Creation of earthquake-induced landslide susceptibility data

This study utilizes the earthquake-induced landslide susceptibility assessment module to draft the prediction map for the landslide susceptibility distribution of earthquake-stricken areas. The module tends to be the outcome of the project supported by the “Eleventh Five-year” national science and technology program, that is, “Integrated risk assessment technique for major earthquake-induced disasters and their disaster chains” (No.: 2008BAK50B03, undertaken by: Gaozhong Nie). In this program, Gaozhong Nie used the data of historical earthquake-induced landslides occurred in mainland China prior to 2008 and the data of Wenchuan earthquake-induced
landsides in 2008 to establish a set of methods for assessing the risk of earthquake-induced landslide in mainland China. This helped meet the information needs of the Chinese government for the spatial distribution of earthquake-induced landslide risk in the disaster area. The historical earthquake-induced landslide data prior to 2008 used in the model is came from the basic data of Chinese historical earthquake-induced landslide accumulated by Qinghua Gao and others (Gao et al. 2011b). The Wenchuan earthquake-induced landslide data they used came from the detailed investigation results of Wenchuan earthquake-induced landslides published by the China Geological Survey (Qin et al. 2008). Based on this, the spatial distribution of earthquake-induced landslide susceptibility levels with multiple intensities is allocated, to generate the 90 m × 90 m raster data. The data covers the entire territory of China.

The elements are regularly square. Each element represents a ground space division of 90 m × 90 m. From the 90 m × 90 m raster data collected in early 2008, slopes, rock attributes, and other information were taken for generating this data. Therefore, this data can be used to assess/inverse the spatial distribution of earthquake-induced landslide susceptibility levels in China since 2008. The earthquake-induced landslide risk assessment module established by Gaozhong Nie and his research group was used by the Chinese government for its earthquake emergency command. The data of geographical environment factors involved in the earthquake-induced landslide risk assessment are also updated every year, ensuring that the results of the earthquake emergency assessment can provide reasonable decision support to the Chinese government for future emergency rescue. In terms of earthquake-induced landslide probability, the landslide susceptibility is categorized into five levels from low to high, i.e., 1, 2, 3, 4, and 5. The scale of this attribute value, ranging from 1 to 5, signifies the

Figure 4. Spatial distribution of earthquake-induced landslide susceptibility levels with the seismic intensity VI-XI (90 m raster data).
The earthquake-induced landslide susceptibility level from very low to very high (Figure 4). The higher the attribute value, the higher is the possibility of landslide occurrence (Bai et al. 2015). In Figure 4, the attribute value of each raster indicates the earthquake-induced landslide susceptibility level at that spatial location. The data set has been validated as highly reasonable and pragmatic in the past analyses undertaken for rapid earthquake response and preliminary earthquake-induced disaster risk assessment (Bai et al. 2015; An et al. 2015; Gao 2015; Wei et al. 2016; Zhang 2016; Qiu et al. 2017; Cao and Li 2018; Pei et al., 2018; Yang et al. 2019; Yi et al. 2018; Cao and Li 2018; Deng et al. 2019; Yang et al., 2019). This data set has been utilized in this study to generate the data of the earthquake-induced landslide susceptibility in the evaluated region. Based on the intensities determined in the field survey, the data of earthquake-induced landslide susceptibility level for the corresponding intensity region is extracted from the data set. Subsequently, the mosaic feature of the GIS is employed to create the spatial distribution data for earthquake-induced landslide susceptibility of the entire evaluated region (Figure 5).

2.2.3. Assignment of kilometer grid attribute

After creating the data of the earthquake-induced landslide susceptibility levels in the evaluated region, the spatial relation statistics module of the GIS is implemented to generalize the information regarding earthquake-induced landslide susceptibility cells in each kilometer grid. The data entails the landslide susceptibility levels of the cells whose center point exists exclusively in the kilometer grid. The number of landslide susceptibility cells at each attribute value in each kilometer grid is acquired and assigned to the corresponding kilometer grid as its earthquake-induced landslide attribute (Table 2). In the sample table of kilometer grid attributes, FID stands for the field identifier. If the evaluated region is partitioned into N kilometer grids, these grids are numbered from 0 to N-1, creating N field identifiers. P is the identifier for population. In the table, $P_N$ indicates the number of persons in the $N^{th}$ kilometer grid. $X_1$ is the identifier of the earthquake-induced landslide susceptibility cell having the attribute value of 1, then $X_2$, and so on. $X_{1,1}$ indicates $X_{1,1}$ landslide susceptibility cells having the attribute value of 1 in the 1st kilometer grid, and $X_{1,2}$ denotes $X_{1,2}$ landslide susceptibility cells having the attribute value of 2 in the 1st kilometer grid, then $X_{1,3}$, and so on. In the actual table of attributes, the values of $X_{1,1}$ to $X_{N,5}$ are all non-negative integers.

2.2.4. Function for the death rate of earthquake-induced landslides in a kilometer grid

The relationship between the death rate of earthquake-induced landslides in each kilometer grid cell and the number of landslide susceptibility cells at different
attribute values inside the grid cell is described as a function $f$. In this study, the landslide susceptibility attribute is considered as a factor affecting the death rate of earthquake landslides. Based on its data attribute, it follows that:

$$M_i = f \left( X_{i,1}, X_{i,2}, X_{i,3}, X_{i,4}, X_{i,5} \right) \quad (2)$$

where $M_i$ is the death rate of earthquake-induced landslides in the $i^{th}$ kilometer grid; $f()$ represents the relationship between the inferred influencing factor and $M_i$; $X_{i,1}$-$X_{i,5}$ indicates the factor influencing the death rate of earthquake-induced landslides; $X_{i,1}$ stands for the number of cells having the attribute value of 1 for the earthquake-induced landslide susceptibility in the $i^{th}$ kilometer grid, $X_{i,2}$ denotes the number of cells having the attribute value of 2 for the earthquake-induced landslide susceptibility in the $i^{th}$ kilometer grid, and so on. Tentatively, the lowest death rate of earthquake-induced landslides in a kilometer grid is 0, and the highest death rate is 1. Nonetheless, the death rate varies depending upon the influencing factor, namely, the earthquake-induced landslide susceptibility attribute. The relationship between the variation and the influencing factor can be described in the logistic regression model. Thus, Equation (2) is expressed as the following logistic regression model:

$$M_i = K e^{a_0 + a_1 X_{i,1} + a_2 X_{i,2} + a_3 X_{i,3} + a_4 X_{i,4} + a_5 X_{i,5}} \quad (3)$$

where $M_i$ is the death rate of earthquake-induced landslides in the $i^{th}$ kilometer grid within the range of 0–1; $K$ is the coefficient of the relationship function; $a_0$ is a constant term with $a_1$, $a_2$, $a_3$, $a_4$, and $a_5$ standing for the parameters of the influencing factor (landslide susceptibility attribute); $X_{i,1}$-$X_{i,5}$ denote the number of cells having attribute values of 1–5 for the earthquake-induced landslide susceptibility levels in the $i^{th}$ kilometer grid. The unknown parameters in the model, i.e., $k$, $a_0$, $a_1$, $a_2$, $a_3$, $a_4$, and $a_5$, are derived by repeatedly observing the acquired sample data. While solving it, each piece of sample data is regarded as an observed piece of data from $N$ observations. Assuming that each piece of sample data satisfies Equation (3), it becomes:

$$M_i = K e^{a_0 + a_1 X_{i,1} + a_2 X_{i,2} + a_3 X_{i,3} + a_4 X_{i,4} + a_5 X_{i,5}} i = 1, 2, 3, \ldots, N \quad (4)$$

where $i$ represents the $i^{th}$ observation in $N$ observations. After the observation, data and relationship functions are used to estimate the unknown parameters $k$, $a_0$, $a_1$, $a_2$, $a_3$, $a_4$, and $a_5$. In statistics, the number of deaths $d_i$ is considered as a random variable, and subjected to the following distribution law:

| FID | $P$ | $X_{1,1}$ | $X_{1,2}$ | $X_{1,3}$ | $X_{1,4}$ | $X_{1,5}$ |
|----|----|----------|----------|----------|----------|----------|
| 0  | $P_1$ | $X_{1,1}$ | $X_{1,2}$ | $X_{1,3}$ | $X_{1,4}$ | $X_{1,5}$ |
| 1  | $P_2$ | $X_{2,1}$ | $X_{2,2}$ | $X_{2,3}$ | $X_{2,4}$ | $X_{2,5}$ |
| ... | ... | ... | ... | ... | ... | ... |
| N-1 | $P_N$ | $X_{N,1}$ | $X_{N,2}$ | $X_{N,3}$ | $X_{N,4}$ | $X_{N,5}$ |
The maximum likelihood estimation of the unknown parameters $k$ can be determined with the following equation:

$$L(k, a_0, a_1, a_2, a_3, a_4, a_5) = \prod_{i=1}^{N} M(d_i = D_i) = \prod_{i=1}^{N} \left[ C_p^{D_i} M_i^{D_i} (1-M_i)^{p_i-D_i} \right]$$

(6)

For the convenience of calculation, often logarithmic scale is adopted. Therefore, the logarithm likelihood function is:

$$l(k, a_0, a_1, a_2, a_3, a_4, a_5) = \sum_{i=1}^{N} D_i \ln M_i + \sum_{i=1}^{N} (p_i - D_i) \ln (1 - D_i) + \sum_{i=1}^{N} \ln C_p^{D_i}$$

(7)

Equation (4) is substituted into Equation (7) to obtain:

$$l(k, a_0, a_1, a_2, a_3, a_4, a_5) = \sum_{i=1}^{N} D_i \ln \left( \frac{e^{a_0 + a_1 X_{i,1} + a_2 X_{i,2} + a_3 X_{i,3} + a_4 X_{i,4} + a_5 X_{i,5}}}{1 + e^{a_0 + a_1 X_{i,1} + a_2 X_{i,2} + a_3 X_{i,3} + a_4 X_{i,4} + a_5 X_{i,5}}} \right)$$

$$+ \sum_{i=1}^{N} (p_i - D_i) \ln \left( \frac{k}{1 + e^{a_0 + a_1 X_{i,1} + a_2 X_{i,2} + a_3 X_{i,3} + a_4 X_{i,4} + a_5 X_{i,5}}} \right) + \sum_{i=1}^{N} \ln C_p^{D_i}$$

(8)

In the function, $N$, $P_i$, $D_i$, $X_{i,1}$, $X_{i,2}$, $X_{i,3}$, $X_{i,4}$, and $X_{i,5}$ are known quantities that form the observed sample data. Thus, it is the function of the unknown parameters $k$, $a_0$, $a_1$, $a_2$, $a_3$, $a_4$, and $a_5$. The maximum likelihood estimation of the unknown parameters $k$, $a_0$, $a_1$, $a_2$, $a_3$, $a_4$, and $a_5$ should lead to the maximum likelihood function of the logarithm $l(k, a_0, a_1, a_2, a_3, a_4, a_5)$. The last term of the expression for the function $l(k, a_0, a_1, a_2, a_3, a_4, a_5)$ is unrelated to $k$, $a_0$, $a_1$, $a_2$, $a_3$, $a_4$, and $a_5$, and hence could be omitted in the optimization. In other words, it is only necessary to solve for the maximum of the following function:

$$l(k, a_0, a_1, \ldots, a_5) = \sum_{i=1}^{N} D_i \ln \left( \frac{e^{a_0 + a_1 X_{i,1} + a_2 X_{i,2} + a_3 X_{i,3} + a_4 X_{i,4} + a_5 X_{i,5}}}{1 + e^{a_0 + a_1 X_{i,1} + a_2 X_{i,2} + a_3 X_{i,3} + a_4 X_{i,4} + a_5 X_{i,5}}} \right)$$

$$+ \sum_{i=1}^{N} (p_i - D_i) \ln \left( \frac{k}{1 + e^{a_0 + a_1 X_{i,1} + a_2 X_{i,2} + a_3 X_{i,3} + a_4 X_{i,4} + a_5 X_{i,5}}} \right)$$

(9)

The optimization of the function is resolved with the numerical simulation based on the sample data.

### 2.2.5. Assessment of earthquake-induced landslide casualties

The earthquake-induced landslide casualties for the entire region are equal to the sum of the casualties in all the kilometer grids of the region. Following the basic idea in section 2.1, if the region $R$ is partitioned into $N$ 1 km$ \times$ 1 km grids, the total casualties $D$ can be determined with the following equation:
where \( D \) is the earthquake-induced landslide casualties in the region \( R \); \( N \) represents the region \( R \), which is partitioned into \( N \) kilometer grids; \( D_i \) denotes the earthquake-induced landslide casualties in the \( i \)th kilometer grid of the region. In the estimation, the calculated results \( D \) are often rounded up.

3. Results

3.1. The evaluated regions

In this study, three regions were selected, including the Ludian Ms6.5 earthquake-stricken region in 2014 (hereafter referred to as the Ludian evaluated region), the Yiliang Ms5.6 and Ms5.7 earthquake-stricken region, both on 7th September 2012 (hereafter referred to as the Yiliang evaluated region), and the Wenchuan Ms8.0 earthquake-stricken region in 2008 (here in after referred to as the Wenchuan evaluated region) (Figure 6), as the regions to be assessed to test the proposed rapid assessment method for earthquake-induced landslide casualties based on the GIS and logistic regression model. The three evaluated regions are characteristic type of plateau mountainous areas in Sichuan and Yunnan provinces and even Southwest China, as well as other mountainous regions in China. The Ludian evaluated region, possessing the most detailed sample data of earthquake-induced casualties, was selected as the prime region to develop the rapid assessment method for earthquake-induced landslide casualties and evaluate its efficacy. The Yiliang and Wenchuan evaluated
regions were utilized to evaluate the extensibility and applicability of the developed method from a broader perspective.

3.2. Creation of kilometer grid population data

The feature module “Create Fishnet” of the geographical information software ARCGIS was used to create the vector data of 1 km × 1 km spatial grid for the evaluated regions. The population attribute was then assigned to the grids. In the Wenchuan evaluated region, the population data from the year 2007 was used to assign the population attribute to the kilometer grids. In the Yiliang and Ludian evaluated regions, the population data from the year 2011 was adopted to assign the population attribute to the kilometer grids. The spatial data of the kilometer grid with the population attribute was then derived for the evaluated regions (Figure 7).

3.3. Spatial distribution of earthquake-induced landslide susceptibility

The data of earthquake-induced landslide susceptibility spatial distribution for the evaluated regions was created as described in section 2.2.2 (Figure 8). The method mentioned in section 2.2.3 was employed to assign the landslide susceptibility attribute to the kilometer grid of the evaluated regions. Finally, the data of the kilometer grid was generated with population and landslide attributes. The attributes of the data are as shown in Table 2.

3.4. Modeling and evaluation of the Ludian evaluated region

3.4.1. Relational analysis of the death rate of earthquake-induced landslides and the earthquake-induced landslide susceptibility

On August 3rd, 2014, it was reported that the landslides induced by the Ludian Ms 6.5 earthquake directly resulted in the death of 134 people along with 116 people.
missing, all of which were regarded as the total casualties in the landslides induced by the earthquake, that is, 250 casualties in total. Based on the specific spatial information of these reported deaths and missing persons, the casualties and death rate of the earthquake-induced landslides were calculated for each kilometer grid in the evaluated region. The Ludian evaluated region has 10,395 kilometer grids, with 10,395 pieces of information generated for the same, but only a few of the results are listed in Table 3 due to space constraints. The information about the grids that recorded 0 deaths caused by earthquake-induced landslides is not listed, however, most of the grids contain no deaths in the entire evaluated region. Moreover, a large portion of the kilometer grids that recorded 0 deaths had 0 population. Most of the kilometer grids with deaths caused by earthquake-induced landslides belong to the intensity IX area. In the table of attributes (Table 3), $P$ represents the population identifier of a kilometer grid; $D$ is the identifier of earthquake-induced landslide casualties in the kilometer grid; $M$ is the identifier of the death rate of earthquake-induced landslides in the kilometer grid, which was obtained from its corresponding $D/P$; $X_1 - X_5$ represent the identifier of the earthquake-induced landslide susceptibility at the attribute value of 1–5 in the kilometer grid, respectively. In the table, the value in the $X_1$ column is the number of cells having the attribute value of 1 for earthquake-induced landslide susceptibility in the kilometer grid, and the value in the $X_2$ column is the number of cells having the attribute value of 2 for earthquake-induced landslide susceptibility in the kilometer grid, and so on.

Ascertaining the correlation between the death rate of earthquake-induced landslides and the number of cells for each type of earthquake-induced landslide susceptibility in a kilometer grid reveals whether they are independent of each other. A contingency table was employed to analyze and infer the consistency between the death rate of earthquake-induced landslides and the attribute of earthquake-induced landslide susceptibility for 10,395 kilometer grids in the Ludian evaluated region. In the analysis, the death rate of earthquake-induced casualties was classified into 0 and above 0, while the number of cells for different earthquake-induced landslide

![Figure 8. Reversal of spatial distribution for earthquake-induced landslide susceptibility in the evaluated regions.](image-url)
susceptibility levels in each kilometer grid was classified in the same manner. The number of kilometer grids was then counted for each combination formed by the death rate and the number of cells. The contingency analysis was carried out for $X_1$–$X_5$ and $M$ by devising a contingency table for the number of cells for landslide susceptibility attributes and the death rate, respectively (Table 4).

The analysis adopted both the original hypothesis $H_0$ that $X_1$–$X_5$ was independent of $M$, and the alternative hypothesis $H_1$ that $X_1$–$X_5$ was positively consistent with $M$. The test data was utilized to figure out which hypothesis was largely validated. In the contingency analysis of $X_5$ and $M$, the test $p$ value was $1.773 \times 10^{-21}$, and the consistency coefficient was 0.2359. It was a significant revelation that if the value of $X_5$ was 0, the value of $M$ tended to be 0; if the value of $X_5$ was above 0, the value of $M$ also tended to be above 0. This indicated that the original hypothesis, that is, $X_5$ and $M$ were independent of each other, was untenable, and they were positively consistent. In the contingency analysis of $X_4$ and $M$, the test $p$ value was $4.67 \times 10^{-18}$, and the consistency coefficient was 0.3718. It was significantly clear that if the value of $X_4$ was 0, the value of $M$ tended to be 0; if the value of $X_4$ was above 0, the value of $M$ tended to be above 0. This indicated that the original hypothesis, that is, $X_4$ and $M$ were independent of each other, was untenable, and they were positively consistent. In the contingency analysis of $X_3$ and $M$, the test $p$ value was $8.854 \times 10^{-34}$, and the consistency coefficient was 0.1389. It was significantly revealed that if the value of $X_3$ was 0, the value of $M$ tended to be 0; if the value of $X_3$ was above 0, the value of $M$ tended to be above 0. This indicated that the original hypothesis, that is, $X_3$ and $M$ were independent of each other, was untenable, and they were positively consistent. In the contingency analysis of $X_2$ and $M$, the test $p$ value was $7.139 \times 10^{-83}$, and the consistency coefficient was $-0.03289$. It was significantly revealed that if the value of $X_2$ was 0, the value of $M$ tended to be above 0; if the value of $X_2$ was above 0, the value of $M$ tended to be 0. This indicated that the original hypothesis, that is, $X_2$ and $M$ were

| $P$ | $D$ | $M$ | $X_5$ | $X_4$ | $X_3$ | $X_2$ | $X_1$ | $P$ | $D$ | $M$ | $X_5$ | $X_4$ | $X_3$ | $X_2$ | $X_1$ |
|-----|-----|-----|------|------|------|------|------|-----|-----|-----|------|------|------|------|------|
| 90  | 24  | 0.266667 | 10  | 108  | 16  | 1    | 0    | 28  | 1    | 0.035714 | 5    | 58   | 56   | 9    | 0    |
| 134 | 13  | 0.097015 | 0    | 82   | 51  | 0    | 0    | 129 | 4    | 0.031008 | 3    | 113  | 17   | 0    | 0    |
| 94  | 9   | 0.095745 | 24   | 81   | 25  | 0    | 0    | 343 | 9    | 0.026239 | 0    | 71   | 61   | 0    | 0    |
| 14  | 1   | 0.071429 | 0    | 72   | 51  | 0    | 0    | 98  | 2    | 0.020408 | 0    | 90   | 38   | 1    | 0    |
| 65  | 4   | 0.061538 | 0    | 32   | 107 | 0    | 0    | 50  | 1    | 0.02  | 8    | 116  | 6    | 1    | 0    |
| 163 | 8   | 0.04908  | 0    | 85   | 50  | 0    | 0    | 62  | 1    | 0.016129 | 0    | 16   | 96   | 12   | 0    |
| 45  | 2   | 0.044444 | 0    | 5    | 116 | 13   | 0    | 198 | 3    | 0.015152 | 0    | 1    | 62   | 63   | 2    |
| 25  | 1   | 0.04  | 0    | 37   | 88  | 0    | 0    | 192 | 2    | 0.010417 | 0    | 83   | 56   | 1    | 0    |
| 26  | 1   | 0.038462 | 0    | 10   | 125 | 1    | 0    | 492 | 5    | 0.010163 | 9    | 107  | 17   | 1    | 0    |
| 587 | 5   | 0.008518 | 0    | 55   | 71  | 8    | 0    | 199 | 2    | 0.01005  | 0    | 4    | 123  | 3    | 0    |
| 122 | 1   | 0.008197 | 0    | 0    | 88  | 45   | 0    | 199 | 2    | 0.01005  | 0    | 1    | 111  | 27   | 0    |
| 492 | 3   | 0.006098 | 7    | 82   | 39  | 4    | 0    | 110 | 1    | 0.009091 | 0    | 13   | 105  | 11   | 1    |
| 165 | 1   | 0.006061 | 0    | 0    | 22  | 95   | 22   | 224 | 1    | 0.004464 | 0    | 8    | 117  | 1    | 0    |
| 360 | 2   | 0.005556 | 0    | 21   | 73  | 42   | 0    | 241 | 1    | 0.004149 | 0    | 12   | 58   | 58   | 3    |
| 182 | 1   | 0.005495 | 0    | 2    | 111 | 20   | 0    | 492 | 2    | 0.004065 | 0    | 23   | 101  | 10   | 0    |
| 186 | 1   | 0.005376 | 0    | 24   | 104 | 7    | 0    | 246 | 1    | 0.004065 | 0    | 3    | 106  | 28   | 0    |
| 574 | 3   | 0.005226 | 0    | 0    | 63  | 69   | 0    | 504 | 2    | 0.003968 | 0    | 0    | 52   | 71   | 0    |
| 582 | 3   | 0.005155 | 3    | 102  | 22  | 0    | 0    | 523 | 2    | 0.003824 | 0    | 43   | 80   | 2    | 1    |
| 295 | 1   | 0.00339  | 0    | 25   | 97  | 11   | 0    | 544 | 1    | 0.001838 | 0    | 0    | 4    | 104  | 19   |
| 297 | 1   | 0.003367 | 0    | 45   | 80  | 8    | 0    | 586 | 1    | 0.001706 | 0    | 74   | 62   | 1    | 0    |
| 606 | 2   | 0.0033  | 0    | 66   | 64  | 1    | 0    | 621 | 1    | 0.001610 | 0    | 15   | 107  | 9    | 0    |
| 371 | 1   | 0.002695 | 0    | 77   | 47  | 6    | 0    | 738 | 1    | 0.001355 | 0    | 24   | 58   | 49   | 1    |
independent of each other, was untenable, and they were negatively consistent. In the contingency analysis of $X_1$ and $M$, the test $p$ value was $9.26 \times 10^{-106}$ and the consistency coefficient was $-0.03371$. It was significantly revealed that if the value of $X_1$ was 0, the value of $M$ tended to be above 0; if the value of $X_1$ was above 0, the value of $M$ tended to be 0. This indicated that the original hypothesis, that is, $X_1$ and $M$ were independent of each other, was untenable, and they were negatively consistent.

The analysis results in the contingency table revealed that the test $p$ value is constantly much lower than 0.001, indicating that the original hypothesis, that is, the death rate of earthquake-induced landslides and the earthquake-induced landslide susceptibility in a kilometer-grid are independent of each other, is untenable. In other words, the death rate of earthquake-induced landslides in a kilometer-grid relates to the earthquake-induced landslide susceptibility attribute and the number of cells in the grid. The consistency coefficient illustrates the significantly positive consistency of $X_5$, $X_4$, and $X_3$ with $M$. If the number of cells for earthquake-induced landslide susceptibility is increased for the attribute values of 5, 4, and 3 in a kilometer grid, the death rate of earthquake-induced landslides in the kilometer grid tends to be above 0. Yet, there exists a considerably negative consistency of $X_1$ and $X_2$ with the increase in $M$. If the number of cells for earthquake-induced landslide susceptibility is increased for the attribute values of 1 and 2 in a kilometer grid, the death rate of earthquake-induced landslides in the kilometer grid tends to be 0.

### 3.4.2. Assessment model for earthquake-induced landslide casualties

From 10,395 pieces of sample data of the Ludian evaluated region, 44 pieces are shown in Table 3, and 3,001 pieces without any deaths caused by earthquake-induced landslides were selected as the training set. The training set was utilized in the logistic regression modeling for earthquake-induced landslide casualties in the kilometer grids. The remaining 7,350 pieces were employed as the test set to examine the rationality of the constructed model. Therefore, the training set took up approximately 30% of the sample data, while the test set accounted for 70% of the sample data. The Newton method in the $R$ software was employed to solve the maximum

| Table 4. Contingency table about $X$ and $M$. |
|-----------------|---------------|-------------|-------------|
| $X$             | 0             | $>0$        | Total       |
| $X_5X_5$        | 10323         | 37          | 10360       |
| $>0$            | 24            | 11          | 35          |
| Total           | 10347         | 48          | 10395       |
| $X_4$           | 10107         | 8           | 10115       |
| $>0$            | 240           | 40          | 280         |
| Total           | 10347         | 48          | 10395       |
| $X_3$           | 1454          | 47          | 1501        |
| $>0$            | 8893          | 1           | 8894        |
| Total           | 10347         | 48          | 10395       |
| $X_2$           | 1103          | 11          | 1114        |
| $>0$            | 9244          | 37          | 9281        |
| Total           | 10347         | 48          | 10395       |
| $X_1$           | 1100          | 41          | 1141        |
| $>0$            | 9247          | 7           | 9254        |
| Total           | 10347         | 48          | 10395       |
value and the unknown parameters \( k, a_0, a_1, a_2, a_3, a_4, \) and \( a_5 \) in Equation (9). The output of the software was substituted into Equation (3) to obtain the logistic regression relationship between the death rate of earthquake-induced landslides and the attribute of earthquake-induced landslide susceptibility:

\[
M_i = 1.569434502505369 \frac{e^{-8.68346 - 0.09652X_{i,1} - 0.01666X_{i,2} + 0.01365X_{i,3} + 0.03130X_{i,4} + 0.04077X_{i,5}}}{1 + e^{-8.68346 - 0.09652X_{i,1} - 0.01666X_{i,2} + 0.01365X_{i,3} + 0.03130X_{i,4} + 0.04077X_{i,5}}}
\]

(11)

Equation (11) is the model for calculating the death rate of earthquake-induced landslides in the kilometer grids. As demonstrated in the logistic regression model for the death rate of earthquake-induced landslides in the kilometer grid, the absolute value of coefficients are larger for \( X_5 \) and \( X_1 \). This signifies that the death rate of earthquake-induced landslides in the kilometer grid is appreciably influenced by the variation of cells at the maximum and minimum attribute values of earthquake-induced landslide susceptibility. Amongst them, the coefficient of \( X_5 \) tends to be the highest. Hence, the increase in the number of cells having the attribute value of 5 for landslide susceptibility in the kilometer grid appears as the largest contributor to the increase of the death rate of earthquake-induced landslides. The coefficient of \( X_1 \) is the lowest. Thus, the increase in the number of cells having the attribute value of 1 for landslide susceptibility in the kilometer grid makes the largest contribution to the tendency of the death rate towards 0. In the model for the death rate, the coefficients of \( X_5, X_4, \) and \( X_3 \) are 0.04077, 0.03130, and 0.01365, respectively. When the number of cells with the attribute values of 5, 4, and 3 increases in the kilometer grid for landslide susceptibility, the death rate of earthquake-induced landslides goes up. The cells with larger attribute values make a larger contribution to the increase of the death rate. The coefficients of \( X_2 \) and \( X_1 \) are \(-0.01666\) and \(-0.09652\), respectively. When the number of cells with the attribute value of 2 and 1 in the kilometer grid for landslide susceptibility increases, the cells with a smaller attribute value make a larger contribution to the tendency of the death rate of earthquake-induced landslides towards 0. The cells with the lower attribute values make a larger contribution to the tendency of the death rate towards 0. To assess the effectiveness of the model, the \( F \) value is calculated for the training set and the test set, respectively. In statistical methodology, the larger the \( F \) value, the better the model fits and is more capable of elucidating the data. Theoretically, the model can pass the test only if \( F \geq 0.5 \). In the R software, the statement \( >1-pchisq(hp.logit$deviance, hp.logit$df.residual) \) is inserted to calculate the \( F \) value of the training set and the test set. The output is \([1] 1\) in both the cases. In other words, the \( F \) value is 1. To sum up, the death rate of earthquake-induced landslides is a good indicator for the earthquake-induced landslide casualties in a kilometer grid. Moreover, the constructed model can be utilized to calculate the death rate of earthquake-induced landslides in the kilometer grid.

Equations (10) and (11) can be combined to develop the assessment model for earthquake-induced landslide casualties in the kilometer grids. The expression of the model is simplified as:
Equation (12) is used to calculate the earthquake-induced landslide casualties in the training set for the Ludian evaluated region. In the training set, the actual number of earthquake-induced landslide casualties is 135, but the test result of the model gives 129, hence the accuracy of the model is 95.55% for the calculation of earthquake-induced landslide casualties. In the calculation of earthquake-induced landslide casualties with the test set for the Ludian evaluated region, the actual number of earthquake-induced landslide casualties is 115, but the test result of the model is 104, hence the accuracy of the model is 90.43% for the calculation of earthquake-induced landslide casualties. For the entire evaluated region, the landslides induced by the Ludian earthquake caused 250 deaths (including missing persons), but the test result of the model is 233 deaths. Thus, the error rate is 6.80%.

The statistical significance of the constructed model is examined by calculating the kappa coefficient for the actual number of earthquake-induced landslide casualties in each kilometer grid of the Ludian evaluated region and the assessment result of the developed model. For this calculation, the total population in each grid is divided into two parts, i.e., earthquake-induced landslide casualties and other people. The kappa value of the Ludian earthquake-stricken region is 0.912, signifying that the calculated result of the model is practically analogous to the actual number. Therefore, the calculation model for earthquake-induced landslide casualties appears to be statistically pivotal. Additionally, there also exists a higher consistency between the spatial distribution of the examined casualties (Figure 9a) and the spatial distribution of the actual earthquake-induced landslide casualties (Figure 6a).
3.5. Application of the model in the Wenchuan evaluated region and the Yiliang evaluated region

The Wenchuan and Yiliang earthquake-stricken regions were analyzed in the study, to test the extensibility of the proposed method and its applicability to a wider range of areas. The assessment model constructed in section 3.4.2 for earthquake-induced landslide casualties was employed for the Yiliang and Wenchuan evaluated regions without altering the implementation procedure or model parameters in any manner. The assessment result of the model for the casualties caused by the earthquake-induced landslides in the Wenchuan evaluated region was 18,732, more than 1,000 casualties less than the actual reported 20,000 deaths and missing persons that were affected by the landslides (Cheng et al. 2011). The error rate was around 6.5%. As revealed in the spatial distribution of earthquake-induced landslide casualties in the kilometer grid (Figure 9b), the spatial distribution of earthquake-induced landslide casualties was essentially consistent with the distribution of the landslides caused by the Wenchuan earthquake. No detailed information is available on the specific location of deaths caused in the Wenchuan evaluated region, hence the F value was not calculated for the death rate of earthquake-induced landslides for that region, and thereby the kappa test value was also not calculated for the earthquake-induced landslide casualties in each kilometer grid. Nevertheless, the casualties given by the assessment model were compared with the actual casualties caused by earthquake-induced landslides in the region, which was approximately 93.67% of the latter, which was an insignificant difference. The actual number of earthquake-induced landslide casualties was 59 for the Yiliang evaluated region (Bai et al. 2014), but the calculation result of the constructed model for earthquake-induced landslide casualties was 48, giving an error rate of 18.64%. The F value for the death rate of earthquake-induced landslides in the Yiliang evaluated region was 0.893, which was lower than the value for the Ludian evaluated region, but it still appeared somewhat consistent with the statistical requirements. Therefore, the model passed the mathematical test. Based on the assessment result of casualties in each kilometer grid (Figure 9a), the spatial distribution of casualties in the model still appeared consistent with the actual distribution of earthquake-induced landslide casualties (Figure 6a). The kappa test value for earthquake-induced landslide casualties in each kilometer grid was 0.889 in the Yiliang evaluated region. More than half of the deaths caused by the landslides in the Yiliang earthquake were laborers working in the mining area (Bai et al. 2014), who were registered in other places and not included in the corresponding kilometer grid. This was one of the major reasons as to why the error rate in the assessment of earthquake-induced landslide casualties for the Yiliang evaluated region was greater than that in the other two evaluated regions.

4. Discussion

4.1. Model results and influencing factors

A rapid assessment method for earthquake-induced landslide casualties based on the GIS and logistic regression model is put forward in this paper, that considers the
major influencing factors. The method addresses the influence of different earthquakes (intensities) on earthquake-induced landslide susceptibility, as well as the influence of different landslide susceptibility levels and their combination in a kilometer grid on the death rate of earthquake-induced landslides. After adding population exposure in the kilometer grid, a model is devised for rapidly calculating the earthquake-induced landslide casualties. According to the results obtained with the rapid assessment method for earthquake-induced landslide casualties based on the GIS and the logistic regression model, the earthquake-induced landslide casualties in the Ludian evaluated region were 233, which was 17 lower than the actual number of deaths and missing persons caused by earthquake-induced landslides, that is, 250. The $F$ test and kappa test values in a kilometer grid for the evaluated region were 1 and 0.912, respectively, indicating the remarkable statistical significance of the method in the Ludian evaluated region. Furthermore, the earthquake-induced landslide casualties in a kilometer grid given by the model were also highly consistent with the actual spatial distribution of the earthquake-induced landslide casualties in the kilometer grid. The actual earthquake-induced landslide casualties were approximately 20,000 in the Wenchuan evaluated region, and 59 in the Yiliang evaluated region. When the assessment method developed for the Ludian evaluated region was applied in the Wenchuan evaluated region and the Yiliang evaluated region, the following results were obtained without any change to the model parameters: the earthquake-induced landslide casualties calculated with the model for the Wenchuan evaluated region were 18,732, more than 1,000 casualties less than the actual reported 20,000 deaths and missing persons that were affected by the landslides in the earthquake; the error rate being around 6.5%. The earthquake-induced landslide casualties given by the model for the Yiliang evaluated region was 48. The $F$ test and kappa test values in a kilometer grid for the Yiliang evaluated region were 0.893 and 0.889, respectively. It is evident that the results obtained with the rapid assessment method for earthquake-induced landslide casualties based on the GIS and logistic regression model are reasonable.

For the three evaluated regions, the estimation of earthquake-induced landslide casualties in the model is lower than the actual number, which may be attributed to the sample data employed and the methods selected for the above modeling. In the sample data for the modeling of the Ludian earthquake-stricken region, the number of kilometer grids with deaths caused by earthquake-induced landslides is much lower than the number of kilometer grids with no deaths. At the time of modeling, the algorithm should not only consider the accuracy of the sample data related to casualties but should also consider the accuracy of the sample data that illustrates no deaths. In other words, the model has a much higher accuracy while predicting the casualties utilizing the sample data with no deaths as well, than just the sample data with casualties. If the sample data with no deaths caused by earthquake-induced landslides is reduced in the modeling, the number of deaths obtained in the assessment tend to be greater than the actual number of deaths, resulting in another error. However, the error is rooted in the inherent shortcomings of the data stemming from the statistical methodology. In the model debugging process of the Ludian earthquake-stricken region, it was found that the error could be minimized when the model used 3,001 pieces of data with no deaths and 44 pieces with deaths.
4.2. Influence on rapid assessment

It is a significantly important task to rapidly assess the casualties caused by earthquake in China. However, to date there was no operable method for this purpose. For a long time, the casualties caused by the secondary disasters including earthquake-induced landslides were subjectively estimated by experienced experts or even ignored in the evaluation, causing inaccuracy in the results of the rapid assessment of casualties in many earthquakes. This study actively makes an exploratory change to this situation. The specific assessment results of the model demonstrated no deviation in the above three evaluated regions and appear quite consistent with the actual number of earthquake-induced landslide casualties. Therefore, the model meets most of the requirements of the emergency assessment.

The test results for the Wenchuan and Yiliang evaluated regions exhibit reasonable extensibility and applicability of the proposed rapid assessment method for earthquake-induced landslide casualties based on the GIS and the associated logistic regression model. Hence, the proposed rapid assessment model for earthquake-induced landslide casualties emerges as a significantly extensible and applicable model and can be endorsed for the computation of earthquake-induced landslide casualties in other regions with identical acceptable errors as well. As a result, this study can help improve the accuracy of rapid assessment for the casualties caused in the earthquake and meets the urgent needs of rapid assessment for earthquake-induced landslide casualties in China to a large degree.

4.3. Limitations

The earthquake-induced landslide casualties in a kilometer grid are affected by a variety of factors including some uncertainties—earthquake-induced landslide susceptibility and population exposure are two key decisive factors among them. Additionally, earthquake-induced landslide casualties and their distribution vary with the individual difference of population, the behavior of different people (e.g., self and mutual rescue), the spatial distribution of population at different times, and the earthquake forecast. At this early stage of the research, it is however challenging to include all these factors in the assessment model. The resolution of influencing factors used in the research also exerts an effect on the assessment results even if these factors are mainly influential under normal circumstances. It is evident that the correlation between the death rate of earthquake-induced landslides in the kilometer grid and the 90 m earthquake-induced landslide susceptibility raster attribute is vital in devising the model for rapidly assessing the earthquake-induced casualties. As discussed, the model developed in this study is established under two conditions: (1) The assessment of earthquake-induced landslide casualties relies on 1km × 1km grids; (2) the data of earthquake-induced landslide susceptibility has a 90 m × 90 m resolution. Furthermore, the data of landslide susceptibility contains the parameters of ground motion influence based on intensity. The susceptibility is divided into 5 levels from low to high, and the corresponding attribute values of cells vary from 1 to 5. The effect of this method may vary when any change is introduced to the basic scale in the assessment of earthquake-induced landslide casualties. If the data of
earthquake-induced landslide susceptibility does not contain susceptibility levels or have the 90 m × 90 m resolution, the efficacy of the method tends to naturally vary as well.

While developing the assessment method for earthquake-induced landslide casualties, the relationship between the death rate of earthquake-induced landslides in the kilometer grid cells and the attribute of earthquake-induced landslide susceptibility is taken into account. The relationship may be determined with different models, but only the logistic regression model is employed to establish the relationship in this paper. Different models may generate different effects of assessment. It is, therefore, the vital need of the hour to determine various ways of developing these models as well as select the most effective of them all.

The rapid assessment method for earthquake-induced landslide casualties in the kilometer grid is preliminarily developed and applied in the Ludian, Wenchuan, and Yiliang evaluated regions. The method is validated to be effective and may also be applied to the assessment of earthquake-induced damages in other mountainous areas with similar attributes. At the same time, the method is essentially a statistical method. When the sample size is larger or the range of study is wider, the error in the assessment results with the proposed method may be lower. When the seismic intensity of the evaluated region is lower or the range of study is narrower, the relative error of the earthquake-induced landslide casualties obtained from the assessment with the proposed method may also be more discrete or higher. This may be attributed to the influence of some accidental factors such as the inevitable dynamic change of population in the spatial distribution.

As there is a lack of assessment methods for the casualties caused by earthquake-induced landslides, this paper puts forward an assessment method for earthquake-induced landslide casualties in an explorative manner. Nonetheless, there are still certain shortcomings with respect to the sufficiency of historical data related to the earthquake-induced landslide casualties, the prevalence of basic data, and the diversification of models and methods, leaving room for further improvements and developments in the said method in the future. Meanwhile, further study should be made to explore the applicability of the method to other regions around the world or to the rapid assessment of the casualties caused by geological hazards such as waterfall-induced landslides and mudslides without any change to the model parameters. The method proposed in this paper is in the preliminary stage of development and it should be further improved, but even in the present state it offers a simple and rapidly implementable way to assess and quantitatively calculate earthquake-induced landslide casualties.

5. Conclusion

A rapid assessment method for earthquake-induced landslide casualties based on the GIS and logistic regression model is put forward in this study to address the lack of rapid assessment methods for earthquake-induced landslide casualties. The method is then applied to the assessment of earthquake-induced landslide casualties for three past earthquakes. The statistical significance and accuracy of the results are also
evaluated. It is found that the death rate of earthquake-induced landslides in a kilometer grid relates to the earthquake-induced landslide susceptibility attribute and the number of cells in the grid at a given intensity. Moreover, the number of casualties in a kilometer grid is the product of the population in the grid and the death rate of earthquake-induced landslides, while the casualties in all kilometer grids are aggregated to get the earthquake-induced landslide casualties in an evaluated region.

The proposed rapid assessment method for earthquake-induced landslide casualties takes into account the influence of the earthquake-induced landslide susceptibility attribute (i.e., earthquake risk) and the number of cells (i.e., earthquake-induced landslide susceptibility) in a kilometer grid at different seismic intensities, as well as the population in the kilometer grid (i.e., population exposure). The results obtained using the proposed method are compared with the actual earthquake-induced landslide casualties for the three evaluated regions. It is revealed that the model constructed for the Ludian evaluated region is statistically significant. The number of earthquake-induced landslide casualties given by the model and their distribution do not differ much from the actual data of the Ludian evaluated region. After being applied in the Wenchuan evaluated region and the Yiliang evaluated region without any change, the model gives the number of earthquake-induced landslide casualties, which is also close to the actual data. Hence, the rapid assessment method for earthquake-induced landslide casualties based on the GIS and logistic regression model can be used to accurately assess earthquake-induced landslide casualties. The method offers an effective way for the traditional assessment of earthquake-induced landslide casualties and enhances the accuracy of rapid assessment for this purpose.

The proposed assessment method still has some limitations that may lead to insufficient accuracy of assessment results. Its applicability and reliability in regions outside China must be tested in more earthquake-stricken areas. The rapid assessment method for earthquake-induced landslide casualties based on the GIS and logistic regression model is preliminarily developed in this paper but needs to be further expanded in many aspects. For example, further study needs to be carried out to find out whether it is applicable to the rapid assessment of the casualties caused by water-fall-induced landslides, mudslides, and other geological disasters. However, the method presents a simple and rapid way to quantitatively assess earthquake-induced landslide casualties, which can help overcome some problems in the rapid assessment of earthquake-induced landslide casualties. It is therefore a promising focus for future research to make the method more inclusive and reliable for more accurate results, and for the development of better models in the rapid assessment of earthquake-induced landslide casualties.

**Acknowledgements**

We would like to thank the editors and anonymous reviewers for their constructive and detailed advice and suggestions. Also, Professor Shi Zhengtao of Yunnan Normal University, who provided the necessary support for this work. Ms. He Lijing of Zhaotong Municipal Earthquake Disaster Reduction Bureau, who provided the assistance for the field survey on the locations of casualties. Dr. Li Chun helped in writing this paper. We express our gratitude to all of them!
Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Key R&D program of Ministry of Science and Technology of the People’s Republic of China.

References

An JW, Bai X-F, Xu J-H, Nie G-Z, Wang X-Y. 2015. Prediction of highway blockage caused by earthquake-induced landslides for improving earthquake emergency response [J]. Nat Hazards. 79(1):511–536.

Bai X-F, Dai Y-Q, Dai J, Yu Q-K, Zhang F-H, He S-F, Deng S-R. 2013. Study of the main regional features of earthquake disasters in Zhaotong [J]. J Seismol Res. 36(4):514–524. (in Chinese).

Bai X-F, Dai Y-Q, Yu Q-K, Zhang FH, Li Y-S. 2014. Casualty study of the Yiliang Ms5.7 and Ms5.6 earthquakes on September 7, 2014[J]. Earthquake Res China. 30(4):571–582. (in Chinese).

Bai X-F, Dai Y-Q, Yu Q-K, Shao W-L. 2015. Risk Assessment modeling of earthquake-induced landslides and its preliminary application [J]. J Seismol Res. 38(2):301–312. (in Chinese).

Cao Y-B, Li Y-Q. 2018. The key technology and information service of earthquake emergency for Yunnan Province [M]. Beijing: Yunnan Science and Technology Press; p. 81–82. (in Chinese).

Chen Y, Booth, David C. 2011. The Wenchuan earthquake of 2008: anatomy of a disaster [M]. Springer, Science Press; p. 6–7.

Cheng X-L, Li Y, Hong Q-Y, Zhao Y-H. 2011. Numerical simulation of earthquake effects on rock slope[J]. ActaPetrol Sin. 27(6):1899–1908. (in Chinese)

Chen Z-T, Li Z-Q, Ding W-X, Han Z-H. 2012. Study of spatial population distribution in earthquake disaster reduction: A case study of 2007 Ning’er earthquake[J]. Technol Earthquake Disaster Prev. 7(3):273–284. (in Chinese)

Deng S-R, Zhang F-H, Yu Q-K, He S-F, Du H-G, Cao Y-B. 2019. A Method of determining the level of earthquake emergency response[J]. Technol Earthquake Disaster Prev. 14(2): 401–410. (in Chinese).

Ding W-X, Zhang Y-H, Chen Z-T, Temu Q-L, Han Z-H, Li Y. 2014. Spatialization of population data for Hubei Province and its application to rapid assessment of earthquake loss: A case of Badong Ms5.1 earthquake[J]. J Geodesy Geodynamics. 34(3):28–30. (in Chinese).

Gao N. 2015. A study on earthquake emergency radiation capacity[D]. Beijing: Institute of Geology, China Earthquake Administration; p. 72–73. (in Chinese).

Gao Q-H, Liu H-M, Li X-L. 2011b. Regional risk assessment of seismogeological disasters in China[M]. Beijing: China Meteorological Press; p. 4–6. (in Chinese).

Gao Q-H. 2011a. Theory of earthquake risk[M]. Beijing: Mtheorological press; p. 173–185. (in Chinese).

Keeper DK. 1984. Landslides caused by earthquakes [J]. Geol Soc Am Bull. 95(4):406–421.

Kobayashi Y. 1994. Effect of basal guided waves on landslides [J]. PAGEOPH. 142(2):329–346.

Liu J-M, Gao M-T, Wu S-R. 2016. Probabilistic seismic landslide hazard zonation method and its application[J]. Chinese J Rock Mech Engin. (35):3101–3109. (in Chinese).

Pain CF. 1972. Characteristics and geomorphic effects of earthquake-initiated landslides in the Adelbert Range, Papua New Guinea[J]. Eng Geol. 6(4):261–274.

Pei Q, Xia C-N, Liu X-Q, et al. 2018. Analysis of seismic response of prestressed anchor pile reinforced slope under strong earthquake[J]. Coal Technology, 37(9):57–58. (in Chinese).
Plafker G, Ericksen GE, Fernandez J. 1971. Geological aspects of the May 31, 1970, Peru earthquake. Bull Seismol Soc Am. 61(3):543–578.
Qin X-W, Yang J-Z, Zhang Z, Huang J, Yu D-Q, Chen Y-M, Zhang G, Gu F-P. 2008. Remote sensing emergency survey in WenChuan earthquake area[M]. Beijing: Mtheorological press; p. 229–285. (in Chinese)
Qiu D-D, Niu R-Q, Zhao Y-N. 2017. Susceptibility analysis of earthquake-induced landslides with different sampling strategies[J]. J Rock Mech Engin. (S1):3101–3109. (in Chinese).
Wei L-Y, Li H, Dong L-Y, Feng L. 2016. Study on experiment of dynamic shear module and damping ration of loess in Chengde of Hebei[J]. J Seismol Res. 39(3):513–518. (in Chinese).
Xu C, Xu X-W, Zhou B-G, Shen L-L. 2019. Probability of coseismic landslides: A new generation of earthquake-triggered landslide hazard model [J]. J Engin Geol. 27(5):1122–1130. (in Chinese).
Yang G-Y, Zhou W, Fang J-Y. 2019. Assessment of landslide susceptibility based on information quantity model and data normalization[J]. J Geoinformation Sci. 20(5):674–683. (in Chinese).
Yi S-J, Li Y-S, Huang C, Liu K. 2018. The genetic dynamics characteristics of the Jinsha river Shaweitaizi landslide[J]. J Disaster Prev Mitigation Engin. 38(2):297–304. (in Chinese).
Yuan R-M, Deng Q-H, Cunningham D, Xu C, Xu X-W, Chang C-P. 2013. Density distribution of landslides triggered by the 2008 Wenchuan earthquake and their relationships to peak ground acceleration[J]. Bull Seismol Soc Am. 103(4):2344–2355.
Zhang C. 2016. The study of kinetic mechanism of Shawei paleoseismic landslide located by the Jinsha River[D]. Chengdu: Chengdu University of Technology; p. 3–4. (in Chinese).
Zhang R-Q. 2015. A landslides forecast method based on improved fuzzy clustering iterative model and BP neural composite model[D]. Nanchang: University of Science and Technology; p. 11. (in Chinese).