Using short-interval landslide inventories to build short-term and overall spatial prediction models for earthquake-triggered landslides based on machine learning for the 2018 Lombok earthquake sequence

Changhu Xue1 · Kejie Chen1 · Hui Tang2 · Chaoqi Lin1 · Wenfeng Cui1

Received: 17 April 2022 / Accepted: 17 July 2022 / Published online: 8 August 2022 © The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

During an earthquake sequence, there are often multiple recurring landslides. Understanding the spatial distribution of the landslides triggered by the first earthquake can help us predict the landslide susceptibility for subsequent shakes over a short term. This study used two landslide inventories from the Lombok earthquake sequence in Indonesia in 2018 to construct a short-term secondary disaster prediction model and an overall spatial prediction model using four machine learning algorithms. The average accuracy of the positive samples predicted by the prediction model was 7.1% lower than that of the short-term model. The highest accuracy of the overall prediction model was 14.9% higher, on average, and the area under the ROC curve (AUC) score was 8.1% higher, on average, but the corresponding probability thresholds were lower. The reason for this difference is that, in the short-term prediction model, since most of the landslides in the first landslide inventory were prone to fail two or more times due to the effect of multiple earthquakes, the prediction results have a high positive rate. This feature of the short-term prediction model makes it suitable for landslide rescue guidance in a sequence of earthquakes. In contrast, the overall prediction model can better represent the spatial distribution of the earthquake-triggered landslides in the area.

Keywords Earthquake-induced landslides · Machine learning · Landslide susceptibility analysis

* Kejie Chen
  chenkj@sustech.edu.cn

1 Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen 518055, Guangdong, China

2 GFZ German Research Centre for Geoscience, Telegrafenberg, 14473 Potsdam, Germany
1 Introduction

As one of the most destructive geological disasters on Earth, earthquakes commonly induce a large number of landslides (Gorum et al., 2011; Li et al., 2014). For example, May 12, 2008, the Wenchuan earthquake in China (Mw 7.9) triggered hundreds of thousands of landslides, causing a lot of human casualties and economic losses (Xu et al., 2013). As early as the 1970s and 1980s, studies have been conducted on the mechanism, process, and impact of earthquake-induced landslides (Pain, 1972; Garwood et al., 1979; Pearce et al. 1986). Since the Chi–Chi earthquake in Taiwan in 1999, many studies of the mechanism and distribution characteristics of earthquake-triggered landslides have been conducted (Shou et al., 2003; Lin et al. 2004; Tang et al., 2009). These studies have revealed the mechanism of earthquake-induced landslides and the landslide distribution characteristics from different perspectives.

From the 1900s to the 2010s, much effort has been devoted to investigating the relationship between earthquake-triggered landslides and the possible impact factors (Keefer, 1984; Rodriguez et al., 1999; Mahdavifar et al., 2006). These studies have shown that the factors that have a great influence on earthquake-triggered landslides are the topographical factors (Yagi et al., 2009; Yuan et al., 2015; Serey et al., 2019), the geological structure (Mahdavifar et al., 2006; Sato et al., 2007; Huang et al., 2009), the ground shaking parameters (Meunier et al., 2007; Nowicki et al., 2014; Huang et al. 2021). Many scholars have used remote sensing images from before and after the earthquake to identify co-seismic landslides and generate earthquake-triggered landslide inventories (ETLIs) (Sato et al., 2007; Wartman et al., 2013; Tanyas et al., 2017; Valkaniotis et al., 2018; Ferrario, 2019). Based on the above background, some scholars have constructed spatial prediction models for earthquake-triggered landslides using machine learning method and obtained good results. For example, Niu et al. (2014) used an improved support vector machine model to identify and predict the landslides triggered by the Lushan earthquake in China in 2013 and evaluated the susceptibility to landslides in the area of high seismic intensity. Chuang et al. (2021) selected the 1999 Chi–Chi earthquake in Taiwan and used the logistic regression (LR) algorithm for machine learning modeling. They then tested the model using data from the 1998 Jueili earthquake in Taiwan and gave suggestions for probability threshold. Fan et al. (2021) used a random forest (RF) model to track the landslides that occurred after the 2008 Wenchuan earthquake over ten years and established multi-temporal landslide susceptibility maps.

However, there can often be earthquake sequences or a large earthquake followed by many aftershocks over a short period in areas with high earthquake occurrences. One ETLI results from a continuous influence of multiple shakes, and there can be many slopes that suffer more than one failure under the impact of numerous shocks. For instance, Ferrario (2019) constructed two landslide inventories for the landslides that occurred in the first and second halves of the earthquake sequence in Lombok, Indonesia, in 2018. However, the previous studies used the landslides after the earthquake sequence to build an overall spatial prediction model and did not consider predicting multiple landslides in the earthquake sequence. If an ETLI is generated soon enough right after the first shock, can we predict the possibility of secondary landslides once subsequent shocks occur over a short period using machine learning? Is there any difference between the machine learning models for short-term landslide prediction and overall spatial prediction?

In response to the above questions, we chose two ETLIs for the 2018 Lombok sequence built by Ferrario (2019) and used two different strategies to perform machine learning.
modeling. One approach is to use the ETLI for the first half of the earthquake sequence for training and the ETLI for the second half for testing. The other strategy was to merge the two ETLIs and use tenfold cross-validation to evaluate the prediction results. Terrain information, vegetation coverage, and seismic parameters were used as input factors for the model training. We selected three different types of commonly used machine learning methods—logistic regression (LR), random forest (RF), and neural network (NN)—to eliminate the difference between the results of the two strategies caused by the algorithm. The NN used a basic single-layer artificial neural network (ANN) and a multi-layer deep neural network (DNN), so there was a total of four different algorithms. Thus, eight machine learning models were established for the two strategies using these four algorithms. In this paper, we compare the differences in the prediction results from four models under the two strategies and analyze the reasons for the differences in the results of the two strategies.

2 Study area and datasets

2.1 Study area

Lombok is located in southern Indonesia, at the junction between the Sunda Block and the Australian Plate (see Fig. 1a). The area has a tropical rainforest climate with warm and rainy weather all year round. Rainfall is concentrated from November to March, and the dry season is from April to October (Ferrario, 2019). The geological structure in Lombok is complex, and it is an active crustal region. The Flores Thrust located on the north of the islands marks a transition zone between subduction of oceanic crust and continental collision (Silver et al., 1983; Jones et al., 2014).

Our study area locates in the area between 8°35′–8°14′S, 116°00′–116°42′E, around the Mount Rinjani volcano (see Fig. 1b). The Mount Rinjani volcano is 3726 m above sea level, and the average slope of the surrounding area is about 16°. In the center of the volcano is Segara Anak Lake, which was formed by a volcanic eruption (Lavigne et al. 2013). Figure 2 shows the basic geographic information of the Lombok area. The terrain is the one arc-second resolution raster digital elevation model (DEM) data obtained by the

![Fig. 1 Study area map. a Geographical location of the study area. b Earthquakes above Mw 6.0 that occurred during July to August, 2018. The earthquakes and fault information was obtained from the U.S. Geological Survey database (USGS, Accessed Sep 2021)](image-url)
Shuttle Radar Topography Mission (SRTM, Accessed Sep 2021), which have been further processed to show slope, aspect, and curvature. Valley lines were extracted from the DEM, and were considered as rivers or ditches (see Fig. 2f). To investigate the vegetation coverage in the study area, we used the Landsat 8 Operational Land Imager (OLI) optical images (30-m resolution) covering Lombok to calculate the normalized difference vegetation index (NDVI). Due to the active local water vapor, there are few complete images without cloud cover. In order to reduce the influence of the different seasons on the growth of vegetation, we selected three images with good observation conditions close to the season when the earthquakes occurred, for which the dates were August 26, 2017, May 09, 2018, and July 06, 2018, respectively. The three images were stitched together and processed to obtain an NDVI product with almost no cloud occlusion.

2.2 Earthquake events and landslide inventories

From July to August 2018, a series of earthquakes occurred in the Lombok area over a period of less than one month. The seismic sequence started on July 28, 2018, with a Mw 6.4 earthquake. As of August 25, there had been 30 earthquakes above Mw 5.0, nine above Mw 5.5, and four above Mw 6.0, as listed in Table 1 (USGS, Accessed Sep 2021). The epicenters were concentrated in the northeastern coastal area of the island (see Fig. 1b).

This series of earthquakes made the slopes around the volcano vulnerable, resulting in a large number of landslides. Ferrario (2019) identified the earthquake-triggered landslides by comparing the high-resolution PlanetScope optical images from July 02, 2018,
August 08, 2018, and September 13, 2018. Two landslide inventories were constructed: an inventory of the landslides triggered by earthquakes before 08/08/2018 (named “LI1”), and an inventory of the landslides occurred during the earthquake sequence (named “LI2”) (see Fig. 3). According to Ferrario (2019), the landslides in LI1 were mainly triggered by the Mw 6.9 earthquake on August 5, 2018, and the newly mapped landslides in LI2 were mainly triggered by the two earthquakes (Mw 6.3 and Mw 6.9) that took place on August 19, 2018. Additionally, landslides existed in LI1 may be reactivated affected by subsequent earthquakes, so there may be many secondary landslides in LI2. These landslides were all considered positive samples in this experiment. They resulted from the continuous influence of multiple earthquakes over a short period. The followed two major earthquakes, and LI2 followed four major earthquakes, not considering the other small (< Mw 6.0) earthquakes.

We selected three shaking variables—the modified Mercalli intensity (MMI), peak ground acceleration (PGA), and peak ground velocity (PGV)—as the influencing factors

| Date       | Time (UTC) | Location          | M<sub>W</sub> | Depth (km) |
|------------|------------|-------------------|--------------|------------|
| 28/07/2018 | 22:47:37   | 8.34S, 116.48E    | 6.4          | 14         |
| 05/08/2018 | 11:46:37   | 8.32S, 116.46E    | 6.9          | 34         |
| 19/08/2018 | 04:10:23   | 8.35S, 116.59E    | 6.3          | 16         |
| 19/08/2018 | 14:56:26   | 8.37S, 116.67E    | 6.9          | 21         |

Fig. 3  Landslide inventory maps. a LI1 map. b LI2 map. c and d are enlarged partial images of the landslides in a and b, respectively.
of earthquake-induced landslides. In order to investigate the comprehensive influence of multiple earthquakes, we take these three modeling parameters as the results of multiplying the corresponding parameters of each earthquake, followed by each landslide inventory as the training parameters. For LI1, we took:

$$\text{MMI}^* = \ln(\text{MMI}_1 \cdot \text{MMI}_2), \quad \text{PGA}^* = \ln(\text{PGA}_1 \cdot \text{PGA}_2), \quad \text{PGV}^* = \ln(\text{PGV}_1 \cdot \text{PGV}_2)$$

For LI2, we took:

$$\text{MMI}^* = \ln\left(\prod_{i=1}^{4} \text{MMI}_i\right), \quad \text{PGA}^* = \ln\left(\prod_{i=1}^{4} \text{PGA}_i\right), \quad \text{PGV}^* = \ln\left(\prod_{i=1}^{4} \text{PGV}_i\right)$$

where the subscripts represent the indices of the four earthquakes. The logarithm was taken to avoid the differences of these variables between LI1 and LI2 being too large.

Table 2 lists the details of the data used in this experiment.

The landslide polygon data were rasterized to 30-m resolution pixels. The positive samples are the pixels of the landslide areas. Since the area of non-landsides was much greater than the area of landslides, we randomly selected a batch of negative samples in the non-landslide area as negative samples, which were about twice of positive samples. The positive samples of LI1 and LI2 were 5198 and 10,734, respectively. The distribution of each factor in LI1, LI2, and the domain area is shown in Table 3.

### 3 Methods

#### 3.1 Logistic regression

Logistic regression (LR) is a widely used machine learning algorithm that can combine linear regression and classification mapping (Peng et al., 2002; Yesilnacar et al., 2005). LR involves constructing a linear combination of variables as the decision boundary, and then mapping the result to (0, 1) through a nonlinear function to classify the samples. We use the sigmoid function as the mapping function, for which the basic form is as follows:

$$P = \frac{1}{1 + e^{-z}}$$

where $P$ represents the probability of the output result being positive or negative, and $z$ is a linear combination of the factors.

$$z = \mathbf{0}^T \mathbf{X}$$

in which $\mathbf{X}$ is the vector of the input factors, and $\mathbf{0}$ is the weight vector of the factors.

#### 3.2 Random forest

The random forest (RF) algorithm, developed by Breiman (2001), is a robust machine learning algorithm based on ensemble learning, widely used in classification and regression problems. RF implements modeling and prediction through a set of decision trees (DTs). Each DT makes its classification based on the selected control factors. According to the bagging algorithm (Bauer et al. 1999), all DT classification results are used in the
| Data classification       | Data layers                                                                 | Source of data                                                                 | Data type | Layer resolution |
|--------------------------|------------------------------------------------------------------------------|--------------------------------------------------------------------------------|-----------|-----------------|
| Landslide inventories    | Landslides triggered by the 2018/08/05 Mw 6.9 earthquake (LI1)               | USGS Science Database ([https://www.sciencebase.gov/catalog/](https://www.sciencebase.gov/catalog/)) (Ferrario, 2019) | Polygon   |                 |
|                          | Landslides triggered by the 2018/08/19 Mw 6.9 earthquake (LI2)              |                                                                                  |           |                 |
| Topographic parameters   | DEM                                                                          | Shuttle Radar Topography Mission ([SRTM](https://srtm.csi.cgiar.org/) and [http://step.esa.int/auxdata/dem/SRTMGL1/](http://step.esa.int/auxdata/dem/SRTMGL1/)) | Grid      | ~30 m           |
|                          | Slope                                                                        |                                                                                  |           |                 |
|                          | Aspect                                                                       |                                                                                  |           |                 |
|                          | Curvature                                                                    |                                                                                  |           |                 |
| Geological map           | Distance to fault                                                            |                                                                                  | Line      | 100 m           |
| Vegetation               | NDVI                                                                         | Generated from Landsat 8 OLI data                                               | Grid      | 30 m            |
| River or ditch           | Distance to river/ditch                                                       | Generated from DEM                                                              | Grid      | 100 m           |
| Seismic parameters       | MMI                                                                          | USGS Earthquake Hazards ([https://www.usgs.gov/natural-hazards/earthquake-hazards/earthquakes](https://www.usgs.gov/natural-hazards/earthquake-hazards/earthquakes)) | Grid      | 1 arc-minute    |
|                          | PGA                                                                          |                                                                                  |           |                 |
|                          | PGV                                                                          |                                                                                  |           |                 |
|                          | Distance to epicenter                                                         |                                                                                  |           | 100 m           |
| Variables                  | Interval          | Proportion in LI1 (%) | Proportion in LI2 (%) | Proportion in domain (%) | Variables                  | Interval          | Proportion in LI1 (%) | Proportion in LI2 (%) | Proportion in domain (%) |
|----------------------------|-------------------|-----------------------|-----------------------|--------------------------|----------------------------|-------------------|-----------------------|-----------------------|--------------------------|
| Elevation (m)              | ≤500              | 9.24                  | 7.26                  | 40.18                    | Distance to faults (km)     | 60–75             | 19.44                  | 18.12                  | 19.00                    |
|                            | 500–1000          | 51.18                 | 42.11                 | 29.62                    |                            | 75–80             | 30.75                  | 25.19                  | 17.08                    |
|                            | 1000–1500         | 20.40                 | 24.38                 | 14.77                    |                            | 80–85             | 34.63                  | 34.47                  | 20.16                    |
|                            | 1500–2000         | 10.56                 | 13.21                 | 8.10                     |                            | 85–90             | 8081                   | 13.45                  | 13.88                    |
|                            | 2000–2500         | 6.66                  | 9.71                  | 5.44                     |                            | >90               | 6.37                   | 8.77                   | 29.88                    |
|                            | >2500             | 1.96                  | 3.34                  | 1.89                     | NDVI                       | ≤0.7              | 1.37                   | 2.46                   | 16.95                    |
| Slope (°)                  | ≤15               | 31.76                 | 27.83                 | 57.29                    |                            | 0.7–0.8           | 2.52                   | 2.86                   | 4.30                     |
|                            | 15–30             | 48.94                 | 40.85                 | 27.01                    |                            | 0.8–0.9           | 6.08                   | 6.93                   | 8.72                     |
|                            | 30–45             | 16.85                 | 20.35                 | 11.30                    |                            | 0.9–1             | 90.03                  | 87.75                  | 70.03                    |
|                            | 45–60             | 6.83                  | 10.20                 | 4.09                     | Distance to rivers/ditches (m) | <200             | 48.18                  | 44.56                  | 32.45                    |
|                            | >60               | 0.62                  | 0.77                  | 0.31                     |                            | 200–400           | 28.42                  | 27.03                  | 26.04                    |
| Aspect (°)                 | North             | 22.16                 | 15.70                 | 20.38                    |                            | 400–600           | 14.35                  | 15.34                  | 13.97                    |
|                            | North east        | 19.84                 | 14.91                 | 10.97                    |                            | 600–800           | 4.95                   | 6.44                   | 5.78                     |
|                            | East              | 11.93                 | 12.14                 | 9.81                     |                            | 800–1000          | 3.13                   | 4.01                   | 3.37                     |
|                            | South east        | 8.93                  | 12.39                 | 11.49                    |                            | >1000             | 0.97                   | 2.62                   | 18.39                    |
|                            | South             | 8.08                  | 12.43                 | 13.19                    | MMI* (LI1/LI2)             | <3.5 / <6.6       | 0.36                   | 3.62                   | 16.43/17.71               |
|                            | South west        | 7.92                  | 11.83                 | 11.43                    | 3.5–3.6/6.6–6.8            | 12.25             | 11.64                  | 24.24/23.35             |
|                            | West              | 8.68                  | 9.85                  | 10.27                    | 3.6–3.7/6.8–7.0           | 38.05             | 44.13                  | 30.34/29.23             |
|                            | North west        | 12.46                 | 10.75                 | 12.46                    | 3.7–3.8/7.0–7.2           | 40.78             | 29.25                  | 17.55/18.84             |
| Curvature                  | ≤−5.0             | 1.44                  | 1.97                  | 0.76                     | PGA* (LI1/LI2)            | <−0.5 /<1         | 4.19                   | 15.13                  | 15.58/18.23              |
|                            | −5 to −2          | 7.15                  | 8.59                  | 7.72                     |                            | >3.8 / >7.2       | 8.56                   | 11.36                  | 11.44/10.86              |
|                            | −2 to −1          | 12.38                 | 12.19                 | 13.05                    |                            | −0.5–0/1–1.5      | 11.52                  | 13.85                  | 17.69/21.35              |
|                            | −1 to 0           | 31.12                 | 29.28                 | 28.81                    |                            | 0.5/1.5–2         | 36.06                  | 32.54                  | 32.08/23.16              |
|                            | 0 to 1            | 29.29                 | 27.96                 | 28.90                    |                            | 0.5–1/2–2.5       | 39.65                  | 32.61                  | 26.86/22.06              |
Table 3 (continued)

| Variables                  | Interval | Proportion in L11 (%) | Proportion in L12 (%) | Proportion in domain (%) | Variables                  | Interval | Proportion in L11 (%) | Proportion in L12 (%) | Proportion in domain (%) |
|-----------------------------|----------|-----------------------|-----------------------|--------------------------|-----------------------------|----------|-----------------------|-----------------------|--------------------------|
|                             |          |                       |                       |                          |                             |          |                       |                       |                          |
|                             |          |                       |                       |                          |                             |          |                       |                       |                          |
|                             |          |                       |                       |                          |                             |          |                       |                       |                          |
| Distance to epicenter (km)  | 1 to 2   | 10.91                 | 11.38                 | 11.82                    | > 1/> 2.5                   | 8.58     | 5.87                  | 7.78/15.20              |                          |
|                             | 2 to 5   | 6.42                  | 6.89                  | 7.26                     | < 1.6/< 3.0                | 2.15     | 7.91                  | 15.38/21.17             |                          |
|                             | > 5.0    | 1.29                  | 1.74                  | 1.68                     | 1.6–1.8/3.0–3.3            | 9.04     | 23.54                 | 21.85/22.74             |                          |
|                             | ≤ 5      | 6.34                  | 9.23                  | 8.59                     | 1.8–2.0/3.3–3.6            | 36.62    | 35.56                 | 25.82/26.02             |                          |
|                             | 5–10     | 6.83                  | 9.49                  | 9.90                     | 2.0–2.2/3.6–3.9            | 42.15    | 21.11                 | 20.52/16.35             |                          |
|                             | 10–15    | 16.53                 | 15.44                 | 10.80                    | > 2.2/> 3.9                | 10.04    | 11.89                 | 16.42/13.72             |                          |
|                             | 15–20    | 19.01                 | 18.45                 | 12.70                    |                             |          |                       |                       |                          |
|                             | 20–25    | 35.23                 | 30.40                 | 17.22                    |                             |          |                       |                       |                          |
|                             | 25–30    | 9.82                  | 9.14                  | 10.54                    |                             |          |                       |                       |                          |
|                             | > 30     | 6.24                  | 7.89                  | 30.25                    |                             |          |                       |                       |                          |
voting, and the final voting result is the RF classification result. In RF, two main parameters need to be determined: the number of trees \((n)\) and the number of predictive variables \((m)\) used by the nodes of each tree. As parameter \(n\) increases, the classification results tend to converge. In this experiment, after preliminary analysis, we found that when \(n\) was over 60, the accuracy of the classification results was no longer significantly improved, so we took parameter \(n\) as 60. Parameter \(m\) is approximately the arithmetic square root of \(n\) and was set to 8 (Chen et al. 2020) (Fig. 4).

### 3.3 Neural network algorithms

Neural network (NN) algorithms are a well-developed machine learning method used for regression analysis and classification problems. NN algorithms update the parameters of the model through error backpropagation. They calculate the gradient of the loss function by derivation and then adjust the weight and offset of each node so that the output result decreases along the gradient direction, and a convergence result is iteratively obtained (Lee, 2004; Pradhan et al. 2010). The backpropagation algorithm has two stages in each iteration of the training process: forward forecast and backward update. The model input data are run forward to calculate the probability of the sample being a landslide, and are compared with the actual result. The error is then backpropagated, and the parameters of each layer are re-estimated. This study used single and multiple hidden layers to build the NN and DNN models, respectively.

An NN model includes an input layer, a hidden layer, and an output layer. These layers are connected by linear combinations and nonlinear activation functions. The performance of an NN model is related to the network structure, including the number of layers and nodes, the activation function, and the update method of the parameters (Haykin, 1998). The number of nodes in the input layer is the same as the number of selected factors, and they are usually set as the normalized values of the factors. The connection of nodes between adjacent layers is completed by a linear connection and a nonlinear activation function. Each linear connection contains two undetermined parameters, the weight and bias, which also need to be initialized and updated for NN modeling. In the forward steps, the output result is compared with the target value, and the difference between them is used to estimate the loss. Then, the loss is passed to the front in the backward process, and the parameters are re-estimated. In this experiment, a basic single hidden layer NN model was selected, with the hidden layer containing 20 nodes. In order to classify the results, we chose the sigmoid function as the activation function of the output layer.

![Random forest processing flow in the experiment conducted in this study](image)

*Fig. 4* Random forest processing flow in the experiment conducted in this study
A DNN increases the number of hidden layers in the neural network to more than two. The nodes in each layer are connected by the nodes in the upper layer through a linear combination and an activation function. Generally speaking, the increase in the number of hidden layers and nodes is conducive to the better performance of the model. However, to prevent the model from overfitting, we must ensure that the number of hyperparameters is no greater than the number of samples. In this experiment, we finally used three hidden layers and nodes to count the results by setting different numbers of layers and nodes. The total nodes in each layer were 16, 64, and 4, respectively. More layers and nodes do not consistently achieve better prediction results. The sigmoid and cross-entropy functions were selected as the classifier and the loss function, respectively (Fig. 5).

### 3.4 Training and testing strategies

We adopted two approaches for the training and validating from two different perspectives. The first approach was to use LI1 for training and LI2 for testing (named “Strategy-1” or “S1” for short). This is because, in a series of earthquakes that have occurred over a short period, there are typically many landslides or repeated landslides in many areas following the subsequent shocks. This approach predicts possible secondary landslides over the short-term based on the first landslide inventory in the earthquake sequence. Another approach is to merge LI1 and LI2 into an overall landslide inventory, which is then divided into training data and test data in proportion (named “Strategy-2” or “S2” for short). The overall landslide inventory of an earthquake sequence reflects the tendency for earthquake-affected landslides in the area. This approach provides a spatial reference for the long-term earthquake-triggered landslides and is similar to the one used in the existing research cases mentioned in the introduction. In this approach, we divided the merged landslide catalog into ten equal parts and performed tenfold cross-validation.

The performance of the models was measured by the overall accuracy, precision, recall, and the area under the receiver operating characteristic (ROC) curve (AOC):

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

![Deep learning processing flow in the experiment conducted in this study](image-url)
in which TP, TN, FP, FN refer to the number of true positive samples, true negative samples, false positive samples and false negative numbers, respectively.

4 Results

4.1 Result analysis and model evaluation

The predicted landslide probability under S1 and S2 are shown in Figs. 6 and 7, separately. The prediction target of S1 was LI2, and that of S2 was the overall landslide inventory after decile division. It can be seen from the figure that the spatial distributions of the probability predicted by the four models under these two strategies are consistent with the actual landslide distribution. The figure shows that the four types of models produce similar overall distributions of landslide probabilities. Discarding the difference in the performance of these four algorithms, the prediction results of S1 and S2 are approximately the same in spatial distribution. However, when we verified the prediction results with the test data, we found that results of S1 and S2 were quite different (Figs. 6 and 7).

Table 4 lists the average probability for landslides and non-landslides, as predicted by the four models. It shows that, compared with S1, whether it is a landslide sample or a non-landslide sample, the average probability under S2 is reduced. The average probability of landslide samples is decreased by 7.1%, and that of non-landslide samples is decreased by 14.9% (based on S1).

The ROC curves are used to investigate the performance of the models under S1 and S2 (see Fig. 8). The ROC curves under S2 are drawn by the average true positive rate (TPR) and true negative rate (TNR) values of the 10-fold cross-validation. We use the AUC to evaluate the overall performance of the model, which is close to 1 indicating good model performance. The AUC scores of the four models under S1 are 0.7602, 0.8480, 0.7954, and 0.8346, rising to 0.8337, 0.9076, 0.8519, and 0.9126 under S2, respectively. The prediction performance of each model under S2 is better than that under S1, and the AUC score is 8.1% higher, on average.

The probability threshold (p-threshold) changes from 0 to 1 in steps of 0.05. We recorded the precision and recall corresponding to the highest accuracy by adjusting the p-threshold, summarized in Table 5. The accuracy of the four models under S2 is again clearly higher, by an average of 10.5% than these from S1. The precision and recall under S2 are significantly higher than those under S1. However, considering that the precision and recall are related to the p-threshold, and the p-thresholds of each model with the highest accuracy are different, the comparison between precision and recall is not practically important here. We discuss this issue further in Sect. 5.1. It is worth noting that when the models under S1 reach the highest accuracy, the corresponding p-thresholds are all 0.7, while this drops to 0.6 under S2.

Judging from the precision, accuracy, and recall rate of the 10-fold cross-validation (Fig. 9), the graphical trends of each indicator are generally similar for the same model. In contrast, those for the different models appear to be very different. This suggests that the
four models are independent of each other. Therefore, we believe that the similar differences of these models between S1 and S2 are caused by the different training and test data selected by the two strategies. The reason for this will be discussed in Sect. 5.1.

4.2 Importance of factors

To evaluate the importance of each factor for the model, we conducted another series of training and testing. We used the accuracy from previous models as the reference accuracy, removing different variables. The other variables were then used to carry out the modeling, and we compared the final accuracy with the reference accuracy. The difference of maximum accuracy before and after excluding each variable was used as a reference, and the differences of the other factors were scaled proportionally, so as to compare the importance of each factor to the accuracy of the model. The importance of each factor in each model is calculated by Eq. (8):

![Fig. 6 Landslide probability predicted under S1: a LR, b RF, c NN, and d DNN, e distribution of actual landslides in LI2](image)
where \( k \) refers to a given model, which is LR, RF, NN, or DNN; and \( i \) is the factor index. \( w_{ki} \) represents the importance of model \( k \) after removing factor \( i \). The \( \text{Accuracy}_R^k \) is the reference accuracy of model \( k \). The importance of each factor in each model is listed in

\[
w_{ki} = \frac{\text{Accuracy}_i^k - \text{Accuracy}_R^k}{\max\{\text{Accuracy}_i^k - \text{Accuracy}_R^k\}_{i=1:11}}
\]

(8)

Fig. 7 Landslide probability predicted under S2: a LR, b RF, c NN, and d DNN, e distribution of actual landslides in the whole landslide inventory

Table 4 Average probability of landslides and non-landslides predicted under S1 and S2

| Average probability | LR S1 | S2 | RF S1 | S2 | NN S1 | S2 | DNN S1 | S2 |
|---------------------|-------|----|-------|----|-------|----|--------|----|
| Landslides          | 0.671 | 0.620 | 0.739 | 0.698 | 0.696 | 0.637 | 0.721 | 0.673 |
| Non-landslides      | 0.365 | 0.349 | 0.306 | 0.297 | 0.335 | 0.331 | 0.282 | 0.269 |
Table 6. It shows that, in most cases, the slope and elevation are the most important factors. The distance to rivers/ditches also shows a relatively high importance. This is because the occurrence of landslides shows strong spatial characteristics, that is, most landslides occur on the sides of valleys, in the middle of the mountainside. For the two strategies of S1 and S2, the importance of each factor is similar.

The aspect, curvature, and NDVI are much less important to the models, especially the curvature, which has almost no effect on the models. There is a relatively small difference in the value distribution of these factors between landslides and non-landslides (see Table 3).

By taking the average of the importance of each factor and sorting them, we obtained the factor importance ranking, as shown in Fig. 10. Slope and elevation are the two most important factors, reaching importance of more than 0.9. The importance of distance to rivers/ditches is also above 0.5. Earthquake-related factors are of moderate importance. The importance of MMI*, PGA*, and PGV* is 0.45, 0.54, and 0.50, respectively.
5 Discussion

5.1 Analysis of the differences from two strategies

The results of the two machine learning approaches (S1 and S2), show remarkable differences, as described in Sects. 4.1 and 4.2. Specifically, for each model, the difference in the results has three main manifestations:

- **Fig. 9** Tenfold cross-validation results: a accuracy, b precision, c recall, and d AUC score

| Removed factor      | LR S1 | LR S2 | RF S1 | RF S2 | NN S1 | NN S2 | DNN S1 | DNN S2 |
|---------------------|-------|-------|-------|-------|-------|-------|--------|--------|
| Elevation           | 1.00  | 0.92  | 0.88  | 0.85  | 0.96  | 0.92  | 0.90   | 0.86   |
| Slope               | 0.97  | 1.00  | 1.00  | 1.00  | 1.00  | 1.00  | 1.00   | 1.00   |
| Aspect              | 0.19  | 0.16  | 0.14  | 0.12  | 0.14  | 0.10  | 0.13   | 0.11   |
| Curvature           | −0.01 | 0.04  | 0.08  | 0.03  | 0.00  | 0.11  | 0.03   | 0.05   |
| Distance to epicenter| 0.31  | 0.38  | 0.19  | 0.25  | 0.36  | 0.43  | 0.31   | 0.23   |
| Distance to fault   | 0.33  | 0.38  | 0.22  | 0.23  | 0.34  | 0.39  | 0.30   | 0.25   |
| NDVI                | 0.20  | 0.19  | 0.12  | 0.10  | 0.18  | 0.16  | 0.20   | 0.12   |
| Distance to rivers/ditches | 0.62  | 0.57  | 0.66  | 0.65  | 0.70  | 0.68  | 0.64   | 0.58   |
| MMI                 | 0.46  | 0.42  | 0.43  | 0.42  | 0.51  | 0.44  | 0.44   | 0.40   |
| PGA                 | 0.60  | 0.47  | 0.54  | 0.49  | 0.57  | 0.56  | 0.55   | 0.46   |
| PGV                 | 0.54  | 0.48  | 0.50  | 0.46  | 0.58  | 0.51  | 0.50   | 0.44   |
(a) The mean probability of landslide samples from S2 is lower than that from S1.

(b) The highest accuracy and AUC score of the models from S2 are higher than those from S1.

(c) The \( p \)-threshold with the highest accuracy from S2 is lower than that from S1.

Since accuracy and recall depend on the value of the \( p \)-threshold, we list in Table 7 all indicators, with \( p \)-threshold = 0.6–0.75, to compare the difference between S1 and S2 with different values of \( p \)-threshold. Comparing the results from S1 and S2 for each model, the accuracy and precision from S2 are higher than those from S1, while the recall from S2 is lower than that from S1. This means that models from S2 have a poorer recognition ability for real landslides even though the overall accuracy of the S2 prediction is higher than S1.

According to the training and test data characteristics, we can infer the following reasons. Firstly, Fig. 3 shows that the landslide distributions of LI1 and LI2 have a high degree of overlap. We estimated statistics for the landslide pixels. In LI2, over 44% of the landslide pixels were landslides samples in LI1, and many of the new positive samples were reactivated by landslides in LI1. A large number of overlapping positive samples results

---

**Table 7** Accuracy, precision, and recall of each model with different \( p \)-thresholds values

| Indicator | \( p \)-threshold | LR S1 | LR S2 | RF S1 | RF S2 | NN S1 | NN S2 | DNN S1 | DNN S2 |
|-----------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Accuracy  | 0.6               | 0.7339| 0.8223| 0.7867| 0.8854| 0.7571| 0.8373| 0.7807| 0.8807|
|           | 0.7               | 0.7502| 0.8201| 0.8030| 0.8817| 0.7717| 0.8304| 0.7919| 0.8738|
| Precision | 0.6               | 0.5993| 0.6625| 0.6563| 0.7832| 0.6246| 0.6812| 0.6533| 0.7825|
|           | 0.7               | 0.6421| 0.6925| 0.7045| 0.8409| 0.6736| 0.7212| 0.6935| 0.8129|
| Recall    | 0.6               | 0.6836| 0.6796| 0.7980| 0.7882| 0.7362| 0.7320| 0.7724| 0.7652|
|           | 0.7               | 0.6188| 0.5847| 0.7360| 0.6860| 0.6524| 0.5930| 0.7081| 0.6842|
in the models trained with LI1 having a relatively high positive rate for LI2. Secondly, the number of earthquakes plays an important role in causing new landslides and reactivating landslides. The influence of more earthquakes in S2 makes the predicted landslide probability higher. This leads to those non-landslide samples with similar geographic conditions to the landslide samples also having a higher probability, increasing the false positive rate (FPR) under S1. Thirdly, the number of samples used for training in LI1 is less than that in LI2, which is also a reason for the model’s poor performance in S1. Besides, the epicenters of the two earthquakes on August 19, 2018 were to the east of that of the previous earthquakes, which caused the affected areas to move eastward. Therefore, there is an error in predicting the landslide of LI2 using the model trained by LI1, which makes the accuracy and AUC of S1 lower. However, in practice, it is difficult for the first landslide inventory to cover all the landslides occurring during the entire earthquake period. Furthermore, due to the higher FPR, the \( p \)-threshold has to take a higher value to obtain the best accuracy, which is why the \( p \)-threshold under S1 is higher than that under S2. In contrast, under S2, the models are built using the overall landslide inventory, with a uniform distribution in the training and test data. The dataset contains all landslides induced by the whole earthquake sequence. Therefore, S2 can better represent the susceptibility level of landslide affected by earthquake in the study area.

5.2 Respective advantages of two strategies in practical applications

These two different training strategies can show different practical values. On the one hand, the training and prediction data of S1 are composed of two landslide inventories during a series of earthquakes that occurred over a short period. In areas with frequent earthquakes, strong earthquakes not only trigger many landslides, but the aftershocks or subsequent shocks can also do that over the short term. According to the two landslide inventories, a very high proportion of the landslides in LI1 experienced secondary failure due to the subsequent earthquakes. Some unstable slopes may be reactivated one or more times during consecutive earthquakes, which is very critical for disaster rescue work. Accordingly, models of S1 have a better predictive ability for recurring landslides but bring a high false negative prediction. In the short-term disaster relief, we would rather believe that the local area has a high probability of landslide than omit places where there may be victims. In the numerical experiments in this study, the RF and deep learning models accurately predicted more than 85% of the secondary landslides, even when the probability threshold was simply set to 0.5. However, rapid construction of landslide inventories remains a challenge due to the weather-susceptibility of optical satellite imagery. Short-term prediction of earthquake-triggered landslides still has much to improve. UAV aerial photography has become an effective means of cataloguing landslides in recent years (Lin et al., 2017; Gupta et al. 2018; Valkaniotis et al., 2018), which can make up for this deficiency to a certain extent. Under the condition that satellite images are continuously updated, once an earthquake occurs, drones can be used to photograph surrounding surface changes with short intervals. It is expected that the joint application of multi-sensor earth observation technology will provide support for the rapid construction of landslide inventories in future.

On the other hand, S2 uses the overall landslide catalog for the training and testing, and the test results represent the landslide susceptibility level in the study area better. In this case, it is more suitable to adopt a treatment approach according to the level of landslide susceptibility. In long-term landslide prevention and management, high FPR means more effort will be spent on more landslide prevention work, which can result
in a waste of resources. In addition, for landslide management and prevention in peace-time, training with only the co-seismic landslide inventory is clearly insufficient, and that with post-seismic landslide inventory can be better. In a long-term disaster prevention, we need a more comprehensive distribution of landslide susceptibility in the area affected by earthquakes.

The results of the 4 type of models show that these machine learning algorithms has remarkable difference in performance. RF and DNN show better performance than other algorithms both in S1 and S2. For S1, we can see in Table 5 that while RF has a higher accuracy, its recall is also higher than that of DNN. This shows that RF has better recall, which is very important for short-term disaster relief. For S2, we can also draw similar conclusions when the accuracy reaches peak ($p$-threshold = 0.6). However, the AUC of DNN (0.9107) is higher than that of RF (0.9066) in S2. The results of these two algorithms are not very different, so in the long-term landslide prevention and control, these two algorithms can be selected according to the needs.

6 Conclusion

In this study, the landslide inventories for the 2018 Lombok earthquakes, Indonesia were used to build machine learning models with two strategies. The first strategy (S1) was to use the two short-interval landslide inventories for the training and testing separately. The second strategy (S2) was to divide the overall landslide inventory set into training and test data proportionally. We chose four commonly used machine learning algorithms (LR, RF, NN, and DNN) and built a total of eight models under these two strategies. Compared with S1, the predicted probability of landslide and non-landslide areas under S2 is lower by 7.1% and 14.9%, on average, respectively, while the highest accuracy is 10.5% higher. Affected by multiple earthquakes, the model under S1 can predict the secondary landslide area with higher accuracy, but it also brings a higher rate of misrecognition. Under the influence of multiple earthquakes over a short period, due to the high landslide recurrence rate, the first strategy can more comprehensively predict the possibility of secondary landslides, but it also brings higher misidentification. The S2 strategy can better represent the landslide susceptibility level in the study area, and the AUC score is 8.1% higher, on average. We suggest that using LI1 to establish machine learning models is more suitable for secondary landslide prediction. In contrast, modeling with the overall landslide inventory is more suitable for long-term landslide prevention during non-earthquake periods.

Acknowledgements Some of the figures in this manuscript were drawn by the use of Generic Mapping Tools software (Wessel et al. 2013). The Sentinel-1 data were acquired from the Alaska Satellite Facility (https://asf.alaska.edu/). This research was funded by the National Natural Science Foundation of China (grant 42074024) and the Young Talent Promotion Project of the China Association for Science and Technology.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.
References

Bauer E, Kohavi R (1999) An empirical comparison of voting classification algorithms: bagging, boosting, and variants. Mach Learn 36(1–2):105–139. https://doi.org/10.1023/a:1007515423169
Breiman L (2001) Random forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/A:1010933404324
Chen T, Zhu L, Niu R-Q, Trinder CJ, Peng L, Lei T (2020) Mapping landslide susceptibility at the Three Gorges Reservoir, China, using gradient boosting decision tree, random forest and information value models. J Mt Sci 17(3):670–685. https://doi.org/10.1007/s11629-019-5839-3
Chuang RY, Wu BS, Liu HC, Huang HH, Lu CH (2021) Development of a statistics-based nowcasting model for earthquake-triggered landslides in Taiwan. Eng Geol 289:9. https://doi.org/10.1016/j.enggeo.2021.106177
Fan XM, Yunus AP, Scaringi G, Catani F, Subramanian SS, Xu Q, Huang RQ (2021) Rapidly evolving controls of landslides after a strong earthquake and implications for hazard assessments. Geophys Res Lett 48(1):12. https://doi.org/10.1029/2020gl090509
Ferrario MF (2019) Landslides triggered by multiple earthquakes: insights from the 2018 Lombok (Indonesia) events. Nat Hazards 98(2):575–592. https://doi.org/10.1007/s11069-019-03718-w
Garwood NC, Janos DP, Brokaw N (1979) Earthquake-caused landslides: a major disturbance to tropical forests. Science 205(4410):997–999. https://doi.org/10.1126/science.205.4410.997
Gorum T, Fan X, van Westen CJ, Huang RQ, Xu Q, Tang C, Wang G (2011) Distribution pattern of earthquake-induced landslides triggered by the 12 May 2008 Wenchuan earthquake. Geomorphology 133(3–4):152–167. https://doi.org/10.1016/j.geomorph.2010.12.030
Gupta SK, Shukla DP (2018) Application of drone for landslide mapping, dimension estimation and its 3D reconstruction. J Indian Soc Remote Sensing 46(6):903–914. https://doi.org/10.1007/s12524-017-0727-1
Haykin S (1998) Neural networks: a comprehensive foundation, 2nd edn. Prentice Hall, Upper Saddle River
Huang RQ, Li WL (2009) Analysis of the geo-hazards triggered by the 12 May 2008 Wenchuan Earthquake, China. Bull Eng Geol Environ 68(3):363–371. https://doi.org/10.1007/s10064-009-0207-0
Huang YD, Yao LK (2021) Size distribution law of earthquake-triggered landslides in different seismic intensity zones. Nonlinear Process Geophys 28(1):167–179. https://doi.org/10.5194/npg-28-167-2021
Jones E, Hayes G, Bernardino M, Dannemann F, Furlong K, Benz H and Villaseñor A (2010) Seismicity of the Earth 1900–2012 Java and vicinity: U.S. Geological Survey Open-File Report 2010–1083-N, 1 sheet, scale 1:5,000,000. Doi: https://doi.org/10.3133/ofr20101083N
Keefee DK (1984) Landslides caused by earthquakes. Geol Soc Am Bull 95(4):406–421. https://doi.org/10.1130/0016-7606(1984)95%3c406:lcbe%3e2.0. co;2
Lavigne F, Degeai JP, Komorowski JC, Guillet S, Robert V, Lahitte P, Oppenheimer C, Stoffel M, Vidal CM, Surono PI, Wassmer P, Hajdas I, Hadmoko DS, de Belizal E (2013) Source of the great A.D. 1257 mystery eruption unveiled, Samalas volcano, Rinjani Volcanic Complex, Indonesia. Proc Natl Acad Sci U S A 110(42):16742–16747. https://doi.org/10.1073/pnas.1307520110
Lee TL (2004) Back-propagation neural network for long-term tidal predictions. Ocean Eng 31(2):225–238. https://doi.org/10.1016/s0029-8018(03)00115-x
Li G, West AJ, Densmore AL, Jin Z, Parker RN, Hilton RG (2014) Seismic mountain building: landslides associated with the 2008 Wenchuan earthquake in the context of a generalized model for earthquake volume balance. Geochem Geophys Geosyst 15(4):833–844. https://doi.org/10.1002/2013gc005067
Lin J, Wang M, Yang J, Yang Q and Iop (2017) Landslide identification and information extraction based on optical and multispectral UAV remote sensing imagery. In: international symposium on earth observation for one belt and one road. IOP conference series-earth and environmental science
Lin ML, Tung CC (2004) A GIS-based potential analysis of the landslides induced by the Chi-Chi earthquake. Eng Geol 71(1–2):63–77. https://doi.org/10.1016/s0013-7952(03)00126-1
Mahdavifar MR, Solaymani S, Jafari MK (2006) Landslides triggered by the Avaj, Iran earthquake of June 22, 2002. Eng Geol 86(2–3):166–182. https://doi.org/10.1016/j.enggeo.2006.02.016
Meunier P, Hovius N, Haines AJ (2007) Regional patterns of earthquake-triggered landslides and their relation to ground motion. Geophys Res Lett 34(20):5. https://doi.org/10.1029/2007gl031337
Niu RQ, Wu XL, Yao DK, Peng L, Ai L, Peng JH (2014) Susceptibility assessment of landslides triggered by the Lushan Earthquake, April 20, 2013, China. IEEE J Sel Top Appl Earth Observ Remote Sensing 7(9):3979–3992. https://doi.org/10.1109/jstares.2014.2308553
Nowicki MA, Wald DJ, Hamburger MW, Hearne M, Thompson EM (2014) Development of a globally applicable model for near real-time prediction of seismically induced landslides. Eng Geol 173:54–65. https://doi.org/10.1016/j.enggeo.2014.02.002
Pain CF (1972) Characteristics and geomorphic effects of earthquake initiated landslides in Adelbert Range, Papua New-Guinea. Eng Geol 6(4):261–274. https://doi.org/10.1016/0013-7952(72)90011-7
Pearce AJ, Watson AJ (1986) Effects of earthquake-induced landslides on sediment budget and transport over a 50-yr period. Geology 14(1):52–55

Peng CYJ, Lee KL, Ingersoll GM (2002) An introduction to logistic regression analysis and reporting. J Educ Res 96(1):3–14. https://doi.org/10.1080/0022020670209598786

Pradhan B, Lee S (2010) Regional landslide susceptibility analysis using back-propagation neural network model at Cameron Highland, Malaysia. Landslides 7(1):13–30. https://doi.org/10.1007/s10346-009-0183-2

Rodriguez CE, Bommer JJ, Chandler RJ (1999) Earthquake-induced landslides: 1980–1997. Soil Dyn Earthq Eng 18(5):325–346. https://doi.org/10.1016/s0267-7261(99)00012-3

Sato HP, Hasegawa H, Fujiwara S, Tobita M, Koarai M, Une H, Iwahashi J (2007) Interpretation of landslide distribution triggered by the 2005 Northern Pakistan earthquake using SPOT 5 imagery. Landslides 4(2):113–122. https://doi.org/10.1007/s10346-006-0069-5

Serey A, Pinero-Feliciangeli L, Sepulveda SA, Poblete F, Petley DN, Murphy W (2019) Landslides induced by the 2010 Chile megathrust earthquake: a comprehensive inventory and correlations with geological and seismic factors. Landslides 16(6):1153–1165. https://doi.org/10.1007/s10346-019-01150-6

Shou KJ, Wang CF (2003) Analysis of the Chiufengershan landslide triggered by the 1999 Chi-Chi earthquake in Taiwan. Eng Geol 68(3–4):237–250. https://doi.org/10.1016/s0013-7952(02)00230-2

Silver EA, Reed D, McCaffrey R, Joyodiwiry Y (1983) Back arc thrusting in the Eastern Sunda Arc, Indonesia: a consequence of arc-continent collision. J Geophys Res Solid Earth 88(B9):7429–7448. https://doi.org/10.1029/JB088iB09p07429

SRTM, Accessed Sep 2021, SRTM Data, https://srtm.cgiar.org/srtmdata/, 2021

Tang CL, Hu JC, Lin ML, Angelier J, Lu CY, Chan YC, Chu HT (2009) The Tsaoling landslide triggered by the Chi-Chi earthquake, Taiwan: insights from a discrete element simulation. Eng Geol 106(1–2):19. https://doi.org/10.1016/j.enggeo.2009.02.011

Tanyas H, van Westen CJ, Allstadt KE, Jesse MEAN, Gorum T, Jibson RW, Godt JW, Sato HP, Schmitt RG, Marc O, Hovius N (2017) Presentation and analysis of a worldwide database of earthquake-induced landslide inventories. J Geophys Res Earth Surf 122(10):1991–2015. https://doi.org/10.1002/2017JF004236

USGS, Accessed Sep 2021, Earthquake Hazard Program, https://earthquake.usgs.gov/earthquakes/, Accessed Oct 2021.

Valkaniotis S, Papathanassiou G, Ganas A (2018) Mapping an earthquake-induced landslide based on UAV imagery; case study of the 2015 Okeanos landslide, Lefkada, Greece. Eng Geol 245:141–152. https://doi.org/10.1016/j.enggeo.2018.08.010

Wartman J, Dunham L, Tiwari B, Pradel D (2013) Landslides in eastern Honshu induced by the 2011 off the Pacific coast of Tohoku Earthquake. Bull Seismol Soc Am 103(2B):1503–1521. https://doi.org/10.1785/0120120128

Xu C, Xu X, Yao X, Dai F (2013) Three (nearly) complete inventories of landslides triggered by the May 12, 2008 Wenchuan Mw 7.9 earthquake of China and their spatial distribution statistical analysis. Landslides 11(3):441–461. https://doi.org/10.1007/s10346-013-0404-6

Yagi H, Sato G, Higaki D, Yamamoto M, Yamashiki T (2009) Distribution and characteristics of landslides induced by the Iwate-Miyagi Nairiku Earthquake in 2008 in Tohoku District, Northeast Japan. Landslides 6(4):335–344. https://doi.org/10.1007/s10346-009-0182-3

Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Eng Geol 79(3–4):251–266. https://doi.org/10.1016/j.enggeo.2005.02.002

Yuan RM, Tang CL, Deng QH (2015) Effect of the acceleration component normal to the sliding surface on earthquake-induced landslide triggering. Landslides 12(2):335–344. https://doi.org/10.1007/s10346-014-0486-9

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.