

RESEARCH

Open Access



The performance of using an autoencoder for prediction and susceptibility assessment of landslides: A case study on landslides triggered by the 2018 Hokkaido Eastern Iburi earthquake in Japan

Kounghoon Nam*  and Fawu Wang

Abstract

Background: Thousands of landslides were triggered by the Hokkaido Eastern Iburi earthquake on 6 September 2018 in Iburi regions of Hokkaido, Northern Japan. Most of the landslides (5627 points) occurred intensively between the epicenter and the station that recorded the highest peak ground acceleration. Hundreds of aftershocks followed the major shocks. Moreover, in Iburi region, there is a high possibility of earthquakes occurring in the future. Effective prediction and susceptibility assessment methods are required for sustainable management and disaster mitigation in the study area. The aim of this study is to evaluate the performance of an autoencoder framework based on deep neural network for prediction and susceptibility assessment of regional landslides triggered by earthquakes.

Results: By applying 12 sampling sizes and 12 landslide-influencing factors, 12 landslide susceptibility maps were produced using an autoencoder framework. The results of the model were evaluated using qualitative and quantitative assessment methods. The ratios of the sampling sizes on the non-landslide points randomly generated from the combination zone including plain and mountain (PM) and a mountainous only zone (M) affected different prediction abilities of the model's performance.

Conclusions: The 12 susceptibility maps, including the landslide susceptibility index, indicated the various spatial distributions of the landslide susceptibility values in both PM and the M. The highly accurate models explicitly distinguished the potential areas of landslide from stable areas without expanding the spatial extent of the potential landslide areas. The autoencoder is proved to be an effective and efficient method for extracting spatial patterns through unsupervised learning for the prediction and susceptibility assessment of landslide areas.

Keywords: Autoencoder, Deep neural network, Unsupervised learning, Earthquake, Landslide susceptibility

* Correspondence: geonamsoil@gmail.com

Department of Earth Science, Shimane University, 1060 Nishikawatsu-cho, Matsue, Shimane 690-8504, Japan

Introduction

A magnitude (Mj) 6.7 earthquake occurred on 6 September 2018 at a depth of approximately 35 km in the central and eastern Iburi regions of Hokkaido in Northern Japan. The reported damage included 41 fatalities, 691 injured persons, and 1016 severely collapsed houses. The Japan Meteorological Agency (JMA) designated this earthquake the 2018 Hokkaido Eastern Iburi Earthquake (Fujiwara et al. 2019). Most of the landslides occurred between the epicenter and the highest peak ground acceleration recording station (No HKD 127, Japan).

Various landslide susceptibility methods have evaluated regional landslide areas for spatial prediction and susceptibility assessment by applying different techniques, such as logistic regression (Lee 2005; Ayalew and Yamagishi 2005; Bai et al. 2010; Aditian et al. 2018), decision trees (Saito et al. 2009; Yeon et al. 2010), analytical hierarchy process (Vijith and Dodge, 2019), knowledge driven statistical models (Roy and Saha, 2019), naïve Bayes (Tien Bui et al. 2012; Tsangaratos and Ilia. 2016), support vector machines (Yao et al. 2008; Yilmaz 2010; Xu et al. 2012; Ballabio and Sterlacchini 2012; Zhou and Fang 2015), random forest (Alessandro et al. 2015; Trigila et al. 2015; Hong et al. 2016), and artificial neural networks (Pradhan et al. 2010; Arnone et al. 2016).

Recently, with the rapid development of deep neural networks, state-of-the-art learning approaches in the field of deep learning have been successfully applied in landslide susceptibility mapping, landslide deformation prediction, and landslide time series displacement, including the following techniques: the adaptive neuro-fuzzy inference system (Park et al. 2012); recurrent neural networks (Chen et al. 2015); deep belief networks (Huang and Xiang 2018); long short-term memory (Xiao et al. 2018; Yang et al. 2019); and convolutional neural networks (Wang et al. 2019) since the classification capability of a neural network to fit a decision boundary plane has become significantly more reliable (LeCun et al. 2015).

In the field of landslide hazard assessment, unsupervised learning methods have been focused mainly on landslide inventory detection and landuse classification for image analysis using interferometric synthetic-aperture radar (Mabu et al. 2019) and high-resolution satellite imaging (Liu and Wu 2016; Romero et al. 2016; Lu et al. 2019) in deep learning.

In deep learning, the autoencoder is one type of unsupervised learning method, in which a pre-training algorithm is used to address the problem of backpropagation in the absence of a teacher. The input data are used as the teacher in a neural network that is trained to be symmetrical from input to output for dimensionality reduction and feature extraction (Yu and Principe 2019). The encoder and decoder are the main frameworks for unsupervised deep learning. In some techniques used in the field

of deep learning, there is a lack of an encoder or a decoder. It is costly to compute an encoder and decoder to optimize algorithms for finding a code or sampling methods to achieve a framework. An autoencoder can capture both an encoder and a decoder in its structure by training landslide influencing factors.

The purpose of this study is to evaluate performance of an autoencoder framework for the susceptibility assessment of regional landslides triggered by earthquakes in Iburi region of Hokkaido in Northern Japan. The two primary contributions of this work were as follows. First, 12 different models were set up by various sampling methods using a large amount of data to prevent overfitting while avoiding sampling strategy issues. Second, the autoencoder framework for feature extraction and dimensionality reduction was used for landslide susceptibility mapping. An autoencoder that transforms inputs into outputs with the least possible amount of distortion was pre-trained for feature extraction through dimensionality reduction. Subsequently, landslide susceptibility maps were produced using a deep neural network by supervised learning. To assess the effectiveness of the autoencoder based on a deep neural network, qualitative and quantitative assessment methods were used to measure the imbalanced data.

Study area

The study area was located in tectonically active regions between the Pacific, North American, Eurasian, and Philippine plates, where exist the deepest trenches, such as the Northeast Honshu Arc-Japan Trench and the Kuril Arc-Trench (Kimura 1994; Tamaki et al. 2010). These trenches are mainly composed of sedimentary Quaternary deposits and Neogene rocks, and the soil layers consist of pyroclastic tephra deposits mainly derived from Tarumae caldera, including pumice, volcanic ash, and clay, which were found distributed over a wide area (Tajika et al. 2016). The total thickness of the pyroclastic tephra deposits is about 4–5 m in and around the epicentral area. The highest elevation is less than 700 m. The elevation of the terrain that was affected the most ranges from 100 to 200 m, with slope gradients of 25–30°. After a powerful typhoon (No 21, “Jebi”), the Iburi earthquake occurred. However, some reports claimed that typhoons did not pass directly through the landslide areas and that the average cumulative rain was significantly lower than in the month before the earthquake (Zhang et al. 2019).

Spatial data setting

The landslide inventory map was generated using aerial photographs of the study area, which were taken after the landslides. Then a detailed landslide inventory map incorporating 5627 points of individual landslides was

created from the landslide polygon data using a centroid technique in the ArcGIS environment. Twelve factors that influence landslides were selected from three categories: topography, hydrology, and seismic data.

Landslide inventory mapping

Aerial photographs (ortho-photographs) of the entire area affected by the earthquake were quickly taken by the Geographical Survey Institute, Japan (GSI) as well as several aerial surveying companies. The photographs were posted with analyzed satellite images on the Web as public information. The original polygon shape of 3307 landslide sites was released by the GSI several days after the main shock (Fujiwara et al. 2019). Further reconstruction was carried out to remodify the landslide polygon based on valley lines, ridgelines, hill shade, slope, and aspect, which was generated by a resolution of 10 m. Finally, a detailed landslide inventory map incorporating 5627 points of individual landslides was created by extracting the centroids of the landslide polygons (Fig. 1). This technique has been widely adopted in many landslide susceptibility methods because it is efficient in simplifying landslide data (Tsanagaratos et al. 2017). There is no guiding principle for selecting the boundaries of study areas. According to Zhang et al. (2019), the directional distribution tool

(Standard Deviational Ellipse) in ArcGIS 10.6 indicates ellipses containing certain percentages of the features through standard deviations in the landslide areas. The tool could be useful to guide the deployment of disaster relief operations and mitigation strategies. In the present study, to select the boundaries of the study area, an ellipse corresponding to standard deviations was generated by the directional distribution tool to indicate the general trend of the features. The tool may also be useful in the field of landslide susceptibility mapping for designating the boundaries of the study area, especially the area of landslides triggered by an earthquake as well as active faults in and around the epicenter.

Landslide influencing factors

The pixel size of the factors that influenced the landslides was set to 10 m × 10 m regardless of the resolution of the original data source (Zhu et al. 2018). In landslide susceptibility modeling, a landslide may reoccur under conditions similar to past landslides (Westen et al. 2003; Lee and Talib, 2005; Dagdelenler et al. 2016). There is no common guideline for selecting the factors that influence landslides (Ayalew and Yamagishi. 2005; Yalcin 2008) In this study, 12 factors were selected to evaluate landslide susceptibility: elevation, slope angle, the normalized difference vegetation index (NDVI), distance to

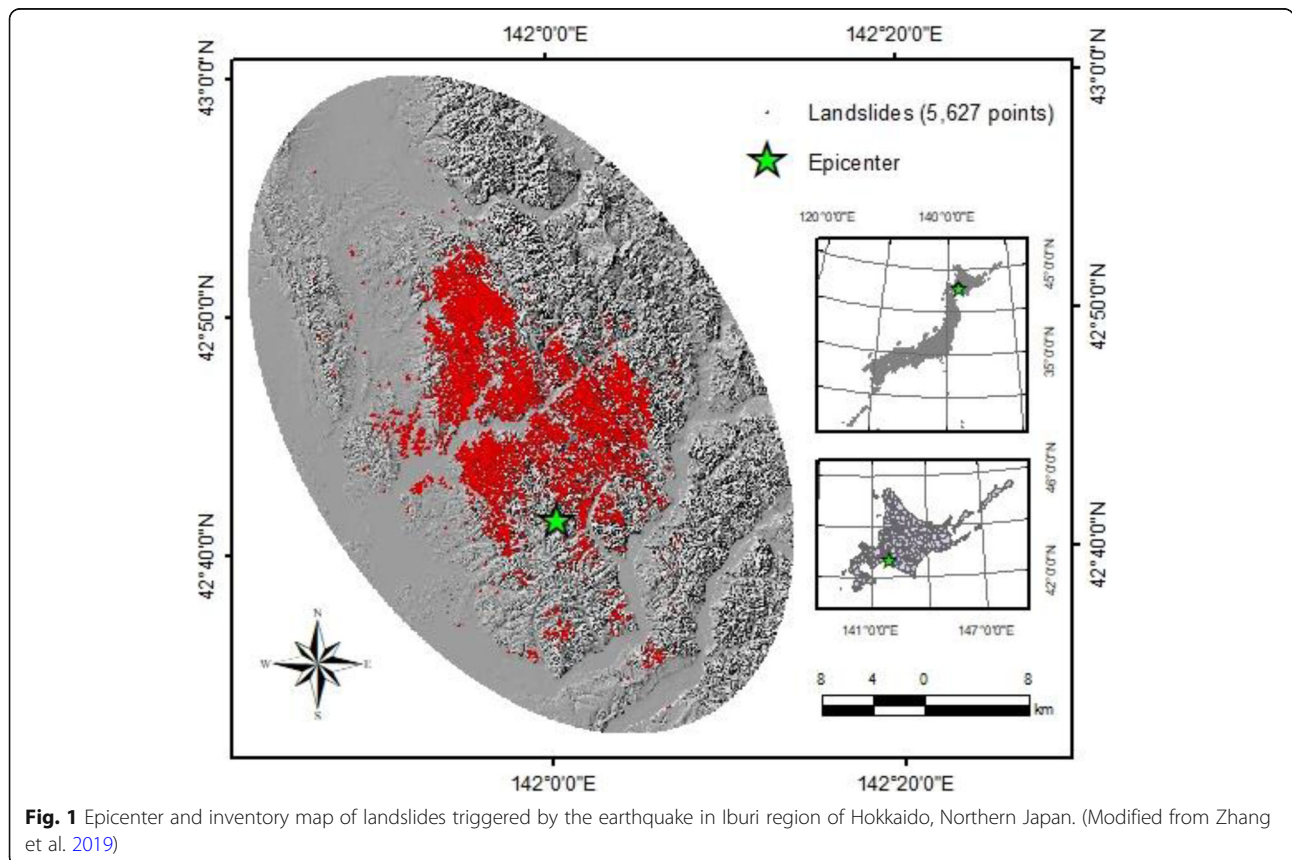


Fig. 1 Epicenter and inventory map of landslides triggered by the earthquake in Iburi region of Hokkaido, Northern Japan. (Modified from Zhang et al. 2019)

stream, stream density, plan curvature, profile curvature, lithology of geology, age of geology, distance to faults, distance to epicenter, and peak ground acceleration (PGA).

Elevation and slope angle are key factors that have been widely employed in landslide susceptibility modeling (Colkesen et al. 2016; Althuwaynee et al. 2016). The elevation values in the study area were divided into equal intervals of 100 m: 0–100 m, 100–200 m, 200–300 m, 300–400 m, 400–500 m, and more than 500 m (Fig. 2a). The slope angle values were extracted from the DEM and classified into six groups at intervals of 10°: 0–10°, 10–20°, 20–30°, 30–40°, 40–50°, and 50–60° (Fig. 2b). The NDVI was used to qualitatively evaluate the conditions of vegetation coverage on slope surfaces, which were calculated from the near-infrared and the red band of Landsat 8 OLI (Chen et al. 2019). The NDVI values were arranged into five classes: (–0.141)–(0.191), 0.191–0.268, 0.268–0.325, 0.325–0.388, and 0.388–0.584 (Fig. 2c). The distance to stream and stream density were evaluated as the role of the runoff and the influence on the slope erosion process by streams in landslides. Previous studies showed that the distance to stream is an important factor that controls landslide occurrence (Devkota et al. 2013; Guo et al. 2015). The distance to stream (Meten et al. 2015) was calculated for each pixel. The streams were classified as follows: 0–100 m, 100–200 m, 200–300 m, 300–400 m, and more than 400 m (Fig. 2d). The stream density was determined from the terrain hydrographic network: 0–1 m, 1–2 m, 2–3 m, 3–4 m, and 4–5 m (Fig. 2e). The plan curvature values represented the steep degrees of slopes that influenced the characteristics of surface runoff contributing to terrain instability (Chen et al. 2019). The plan curvature values were derived from the DEM and classified according to the natural break method into five groups: (–17.246)–(–3.024), (–3.024)–(–0.806), (–0.806)–(0.368), 0.368–1.803, and 1.803–16.025 (Fig. 2f). Profile curvature is the curvature in the vertical plane parallel to the slope direction (Yilmaz et al. 2012). The profile curvature values acquired through DEM were divided into five classes using the natural break method: (–19.381)–(–2.218), (–2.218)–(–0.618), (–0.618)–(0.690), 0.690–2.726, and 2.726–17.708 (Fig. 2g). The geological map of the study area was obtained from the Geological Survey of Japan, AIST, and the lithology map was categorized as follows (Fig. 2h):

- (1) higher terrace;
- (2) lower terrace;
- (3) mafic plutonic rocks;
- (4) marine and non-marine sediments;
- (5) marine sedimentary rocks;

- (6) non-alkaline pyroclastic flow volcanic rocks;
- (7) swamp deposits;
- (8) ultramafic rocks;
- (9) water.

The geological age map was divided into 12 classes as follows (Fig. 2i):

- (1) Early Miocene to Middle Miocene;
- (2) Early to Middle Miocene;
- (3) Late Cretaceous;
- (4) Late Eocene to Early Oligocene;
- (5) Late Miocene to Pliocene;
- (6) Late Pleistocene;
- (7) Late Pleistocene to Holocene;
- (8) Middle Eocene;
- (9) Middle Pleistocene;
- (10) Middle to Late Miocene;
- (11) Present;
- (12) Unknown age.

The distance to fault was computed by a buffer operation. The distance to fault was classified by the geometrical interval function: 1–2.947 km, 2.947–5.545 km, 5.545–9.008 km, 9.008–13.627 km, 13.627–19.786 km, and more than 19.786 km (Fig. 2j). The distance to epicenter was divided according to natural breaks: 1–6 km, 6–11 km, 11–15 km, 15–19 km, 19–23 km, 23–27 km, 27–31 km, and 31–38 km (Fig. 2k). The values of PGA were divided into 10 categories using the geometrical interval function: 184–466 gal, 466–604 gal, 604–671 gal, 671–703 gal, 703–719 gal, 719–752 gal, 752–819 gal, 819–956 gal, 956–1238 gal, and 1238–1817 gal (Fig. 2l). The values of PGA were acquired from K-NET station, Japan (<http://www.kyoshin.bosai.go.jp/>).

Methodology

The methodological hierarchy in this work was based on the autoencoder framework, and 12 sampling sizes were considered for landslide susceptibility mapping. The final prediction results obtained from the autoencoder modeling were evaluated using the testing data set based on qualitative and quantitative analyses to validate the performance of the models. A flowchart of the proposed autoencoder framework is illustrated in Fig. 3. Related techniques are introduced in the following subsections.

Sampling size

The sampling process is the key step in constructing landslide (events) and non-landslide points (non-events) for the database used in landslide susceptibility mapping. Several sampling strategies, such as extracting from seed cell (or gridded) points around a polygon of the landslide area (Meusburger and Alewell 2009; Van Den

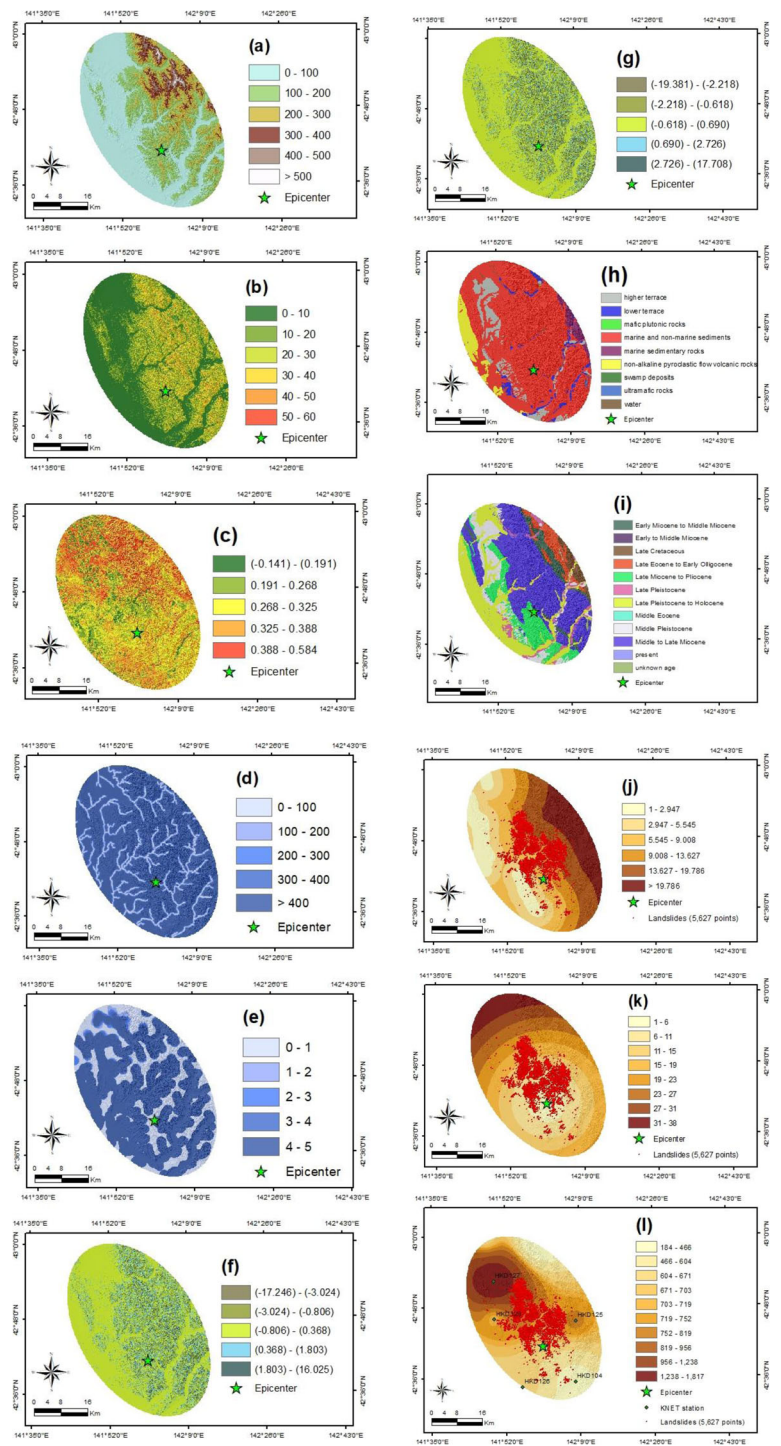


Fig. 2 Thematic maps of landslide-influencing factors in Iburi region of Hokkaido, Northern Japan: **a** elevation [m], **b** slope angle [degree], **c** NDVI, **d** distance to stream [m], **e** stream density [m], **f** plan curvature, **g** profile curvature, **h** lithology, **i** geological age, **j** distance to fault [km], **k** distance to epicenter [km], **l** PGA [gal]

Eeckhaut et al. 2010) and increasing the number of non-events in the non-landslide area (King and Zeng 2001; Raja et al. 2017), have been proposed to improve model's performance, predictive capability, and reduction of

statistical errors. According to King and Zeng (2001), the non-event sample size must not be large but should be two to five times greater than the events because of the disproportionate cost and effort in acquiring data on

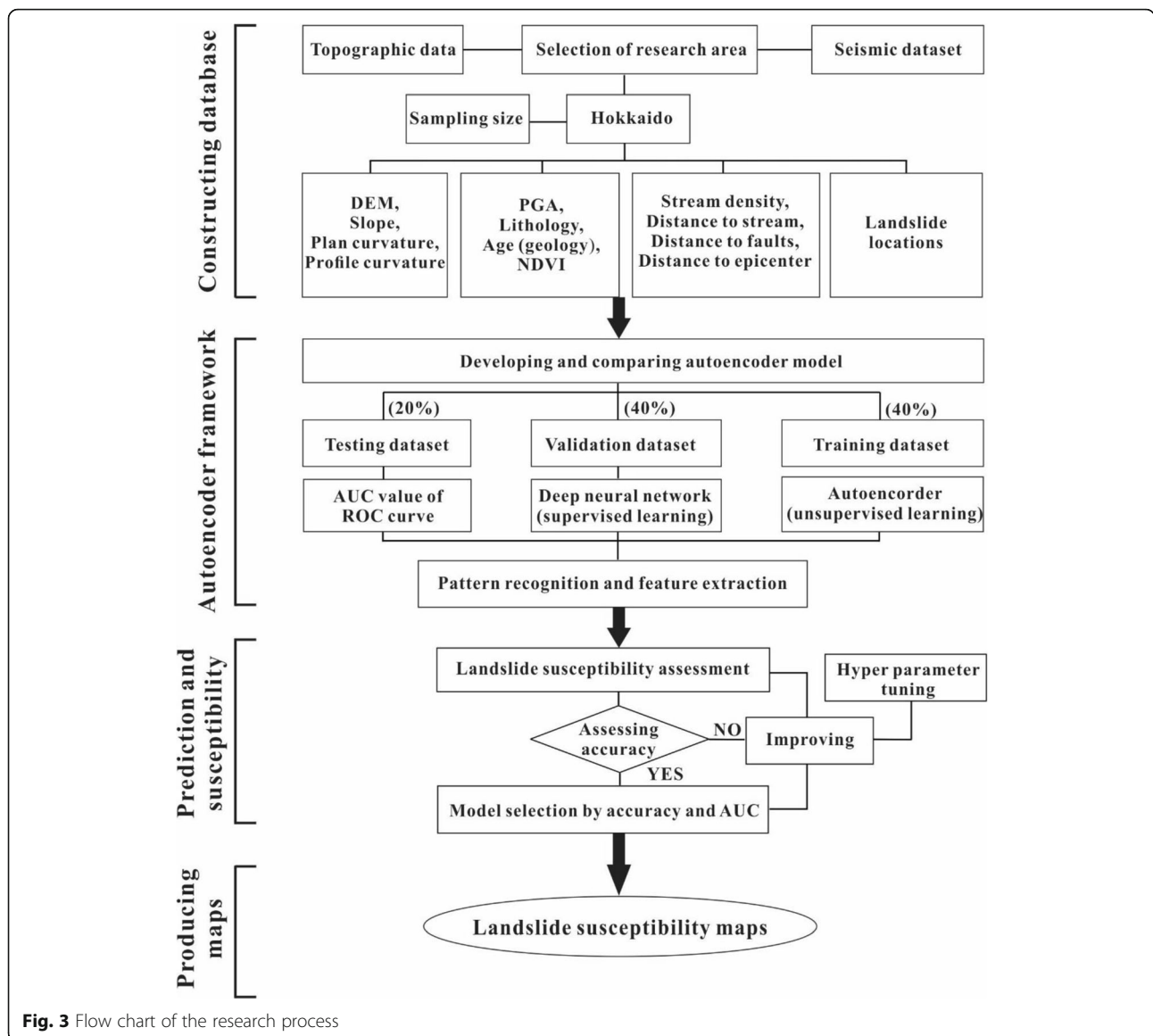


Fig. 3 Flow chart of the research process

many variables, and observations that are not related to the target phenomenon (Heckmann et al. 2014). These studies were conducted mainly to evaluate logistic regression (LR) and rare events LR susceptibility models. Melchiorre et al. (2008) states that un-labeled data sets with a small number of positive examples (events) and a large number of negative examples (non-events) negatively affect the discrimination capabilities of the trained classifier. In this study, to address these issues, 12 different sampling sizes were selected in both the PM and M. Landslide points (event points, 1) and non-landslide points (non-event points, 0) were classified and assigned ratios of approximately 1: 1 (5627: 5627), 1: 2 (5627: 11254), 1: 3 (5627: 16881), 1: 4 (5627: 22508), 1: 5 (5627: 28135), and 1: 10 (5627: 56270) in both areas.

Autoencoder modeling

The autoencoder, which is a special type of multi-layer perceptron, is an artificial neural network. It is a type of unsupervised learning algorithm that has an asymmetric structure, in which the middle layer represents the encoding of the input data in the bottleneck layer (Yu and Principe 2019). The bottleneck constrains the amount of information that can traverse the full network, forcing the learned compression of the input data. The autoencoder is trained to reconstruct input of landslide influencing factors onto the output layer for feature representation, which prevents the simple copying of the data and the network. The middle layer has a lower dimension or a higher dimension based on the desired properties, and it can have as many layers as necessary (Charte et al. 2018). An undercomplete

autoencoder employing a lower dimension learns a nonlinear dimensionality reduction (Hinton and Salakhutdinov 2006) and anomaly detection by taking advantage of the nonlinear dimensionality reduction ability. In this study, an undercomplete autoencoder combined with back propagation neural network was processed for a lower dimension of features than the input data have, which can be used for learning the most important features of the data. Furthermore, in undercomplete representation, an autoencoder with a linear activation function is equivalent to principle component analysis (PCA) in a nonlinear version. The autoencoder models based on the deep neural network were coded in R language on RStudio using H2O packages. These algorithms were performed using hyperbolic tangent function (i.e., the tanh function) in every hidden layer which was used to encode and decode the input to the output in the undercomplete autoencoder. In the H2O library, five hidden layers with encoders and decoders were designed by using the tanh activation function in each layer, which was composed of 10-5-2-5-10 (Fig. 4). The dataset was divided into two separate training sets for unsupervised and supervised learning and one independent test set for the final model comparison. Forty percent of the landslide and non-landslide points were used as training samples for unsupervised learning. The remaining 60% were randomly selected and then used as an independent data set for supervised learning (40%) and for testing the predictive potential (20%) of the autoencoder model to check the performance of the pre-trained model. This study was performed using the following main steps: (1) the unsupervised neural network model was trained based on deep learning autoencoders with the bottleneck algorithm,

where the hidden layer in the middle reduced the dimensionality of the input data; (2) based on the autoencoder model that was previously trained, the input data were reconstructed, and the mean squared error between the actual value and the reconstruction was calculated in each instance; (3) the autoencoder model as pre-training input for the supervised model was performed by using a deep neural network and the weights of the autoencoder for model fitting; (4) to improve the model, different hidden layers were evaluated by performing a grid search by means of hyperparameter tuning, returning to the original features, and trying different algorithms; (5) the area under the curve, such as precision and recall, TPR and TNR, TPR and FPR, and accuracy, were used to measure the model's performance because of the severe bias toward non-event models of randomly generated non-landslide points.

Modeling validation

Several commonly used classification performances were measured based on the confusion matrix, which is employed to evaluate model performance. Each grid cell of the landslide susceptibility map had a unique value representing the landslide susceptibility value. All grid cells were determined as one of four elements: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Precision gives the percentage of true positives as a ratio over all cases that should have been true (1). Recall or the true positive rate (TPR) measures the number of cases that were predicted as positive that should indeed be positive (2). The true negative rate (TNR) measures the proportion of actual negatives that are correctly identified (3). The false positive rate (FPR) is calculated as the ratio between the number of negative events wrongly categorized as positive

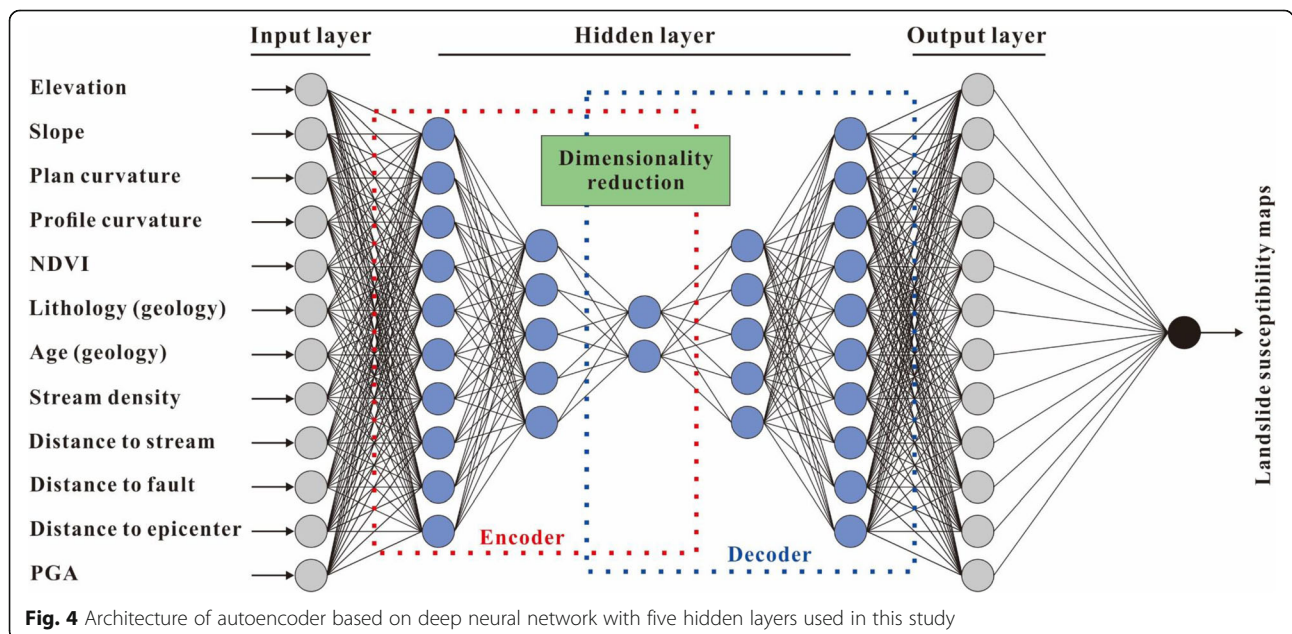


Fig. 4 Architecture of autoencoder based on deep neural network with five hidden layers used in this study

(false positives) and the total number of actual negative events (4). Accuracy is the overall percentage of samples that are correctly predicted as defined in (5).

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \tag{1}$$

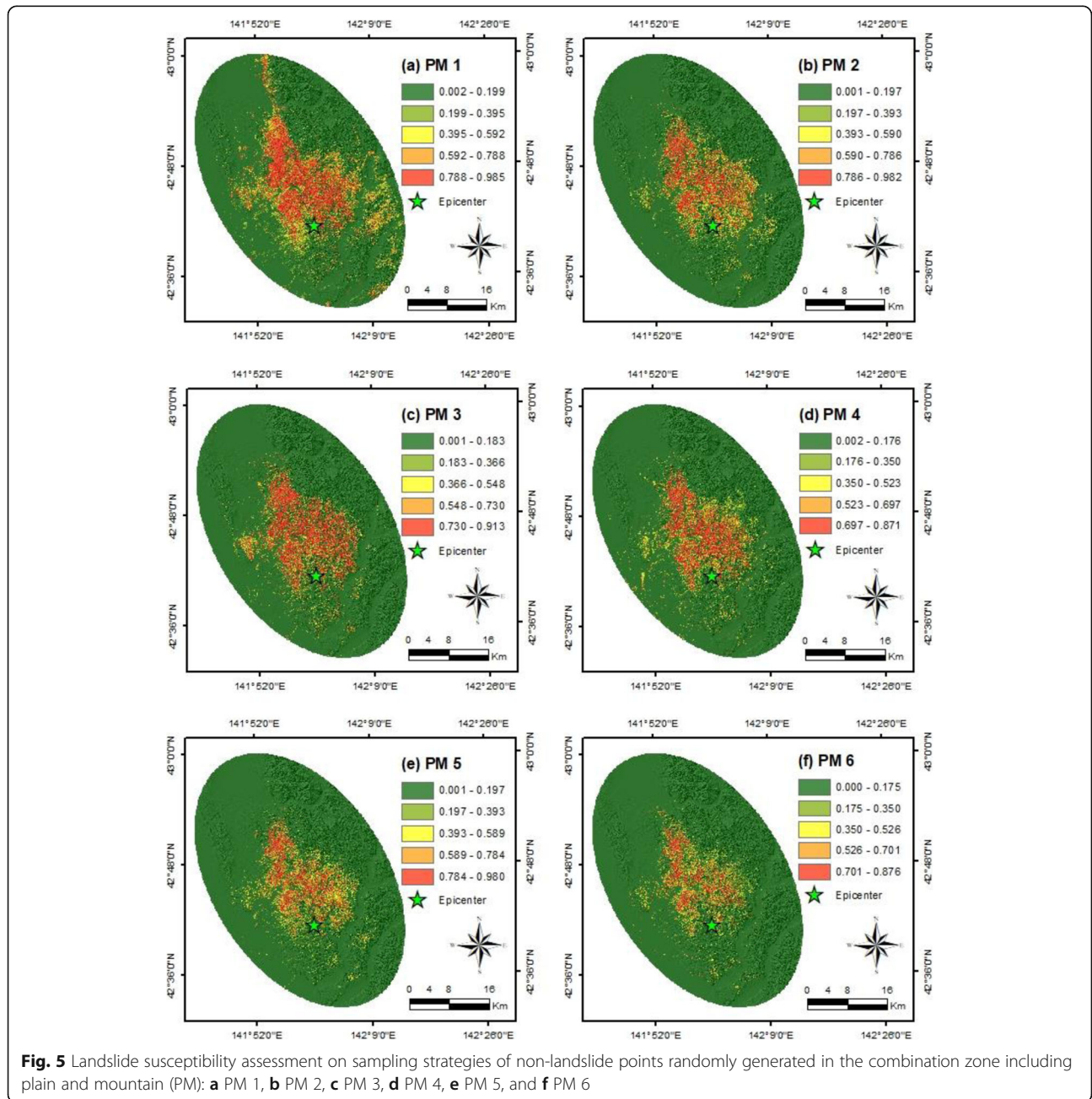
$$\begin{aligned} \text{Recall (true positive rate or sensitivity)} \\ = \text{TP}/(\text{TP} + \text{FN}) \end{aligned} \tag{2}$$

$$\begin{aligned} \text{TNR (true negative rate or specificity)} \\ = \text{TN}/(\text{TN} + \text{FP}) \end{aligned} \tag{3}$$

$$\text{FPR (false positive rate)} = \text{FP}/(\text{FP} + \text{TN}) \tag{4}$$

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{FP} + \text{TN} + \text{FN}) \tag{5}$$

The precision and recall curve presents the relationship between correct landslide predictions and the proportion of landslides detected. The TPR (sensitivity) and TNR (specificity) curve indicates the relationship between the correctly identified classes in both labels (Zhang and Wang 2019).



Results

The model performance for the landslide prediction and the susceptibility assessment were evaluated according to accuracy and the area under the curve. The predictive capability of all the factors that influenced the landslides was evaluated to better understand the spatial patterns based on a variable importance analysis. In the study area, each grid cell was assigned a susceptibility index using the test dataset. After assigning weights to the factor classes, the landslide susceptibility maps were generated in an ArcGIS environment. Finally, using the equal interval function, the indices were reclassified for better

visualization into five classes: very low, low, moderate, high, and very high.

Landslide susceptibility assessment and validation

The 12 landslide susceptibility maps derived from the autoencoder framework are shown in Figs. 5 and 6. It was observed that the accuracy increased with an increase in the sampling ratio, but the precision and recall curve decreased (Table 1). The 12 susceptibility maps showed different spatial distributions in landslide susceptibility. Both the PM 1 (Fig. 5a) model and the M 1 (Fig. 6a) model were predicted to be prone to landslides

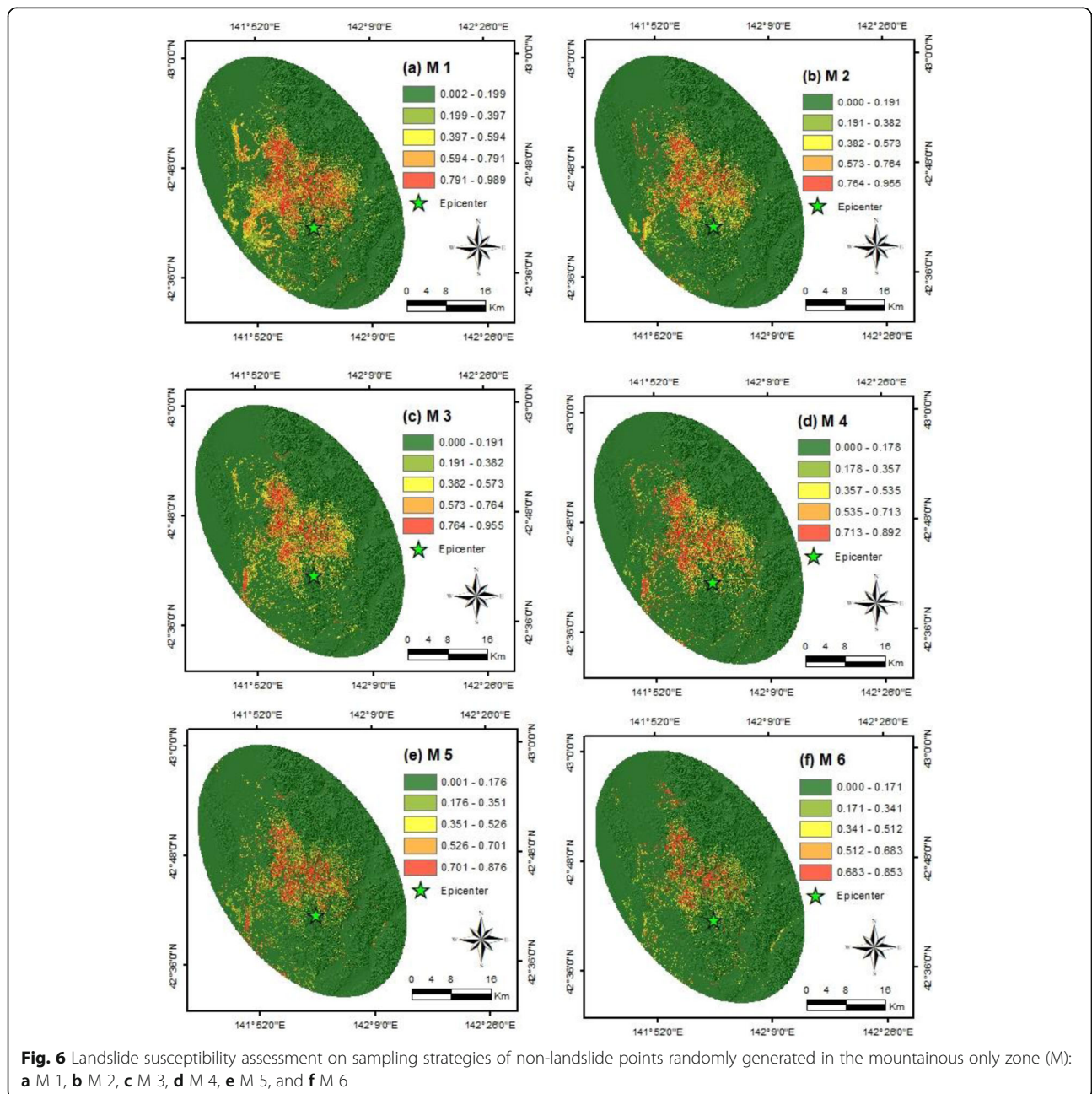


Table 1 The area under the curve and accuracy of models' responses in sampling ratios for the combination zone including plain and mountain (PM) and for the mountainous only zone (M)

Zone	Model	Sampling ratio (Sampling numbers)	Area under the curve on test data set			Accuracy
			Precision & Recall	TPR & TNR	TPR & FPR	
Combination zone including plain and mountain (PM)	PM 1	1:1 (5627:5627)	0.937	0.948	0.891	0.892
	PM 2	1:2 (5627:11254)	0.938	0.968	0.907	0.911
	PM 3	1:3 (5627:16881)	0.903	0.967	0.915	0.920
	PM 4	1:4 (5627:22508)	0.833	0.957	0.905	0.916
	PM 5	1:5 (5627:28135)	0.854	0.967	0.911	0.932
	PM 6	1:10 (5627:56270)	0.754	0.964	0.910	0.943
Mountainous zone (M)	M 1	1:1 (5627:5627)	0.941	0.951	0.891	0.892
	M 2	1:2 (5627:11254)	0.897	0.952	0.886	0.894
	M 3	1:3 (5627:16881)	0.872	0.952	0.887	0.894
	M 4	1:4 (5627:22508)	0.845	0.949	0.881	0.910
	M 5	1:5 (5627:28135)	0.800	0.949	0.883	0.918
	M 6	1:10 (5627:56270)	0.710	0.956	0.893	0.940

in most of the study area, indicating the over-estimation of landslide susceptibility and low capacity in distinguishing landslide-prone areas from stable areas. The PM 1 and M 1 models had the same accuracy of 89.2% by means of confusion matrix, while the precision and recall curve showed that the M 1 model had a performance of 94.1%, which was higher than the 93.7% of the PM 1 model. Regarding the PM 2 (Fig. 5b) and M 2 (Fig. 6b) models, the PM 2 model had an accuracy of 91.1% and a precision and recall curve of 93.8%, which was better than the M 2 model, indicating an accuracy of 89.4% and a precision and recall curve of 89.7%. In the PM 2 model and the M 1 model, the area under the curve of TPR and TNR, and TPR and FPR showed good performance. The PM 2 model with a sampling ratio of 1:2 had the best performance in both distinguishing landslide-prone areas and producing sound information on landslide susceptibility values. The M 1 model sampled on the mountainous zone showed lower accuracy performance compared with the PM 2 model sampled on the combination including plain and mountainous zone. However, the precision and recall curve had the best performance with high accuracy. Regarding the area under the curve on TPR and TNR, and TPR and FPR, the PM 3 model (Fig. Fig. 5c), the PM 4 model (Fig. 5d), the PM 5 model (Fig. 5e), and the PM 6 model (Fig. 5f) distinguished the potential landslide areas from the stable areas without expanding the spatial extent of the potential landslide areas. The M 3 model (Fig. 6c), the M 4 model (Fig. 6d), the M 5 model (Fig. 6e), and the M 6 model (Fig. 6f) generated in the mountainous zone tended to detect stable areas as landslide susceptibility areas even though there were no source areas that caused landslides to be triggered by earthquakes. As shown in Fig. 7, the

final landslide susceptibility index mapped five categories for the PM 2 model (Fig. 7a) and the M 1 model (Fig. 7b), which were the best models selected regarding accuracy and the area under the curve in precision and recall, TPR and TNR, and TPR and FPR.

Variable importance analysis

The predictive capability of all factors that influenced the landslides was evaluated using the test dataset based on H2O's deep learning algorithm (Gedeon 1997), which is a methodology for computing variable importance. Tables 2 and 3 lists the results of the analysis of the variable importance of the factors that influenced the landslides in H2O's deep neural network. In general, the results showed that the earthquake dataset, such as distance to fault, distance to epicenter, and PGA was of high importance to the models, whereas the geomorphology, including slope, plan curvature, profile curvature, stream density, and distance to stream, had lower predictive capability in both areas. Furthermore, in the PM 1 model, the lithology of the geology dataset as categorical variables indicated the highest importance in the models.

Discussion

The 12 landslide susceptibility maps produced by the autoencoder framework were evaluated by the area under the curve in precision and recall, TPR and TNR, and TPR and FPR. In the PM 1 and M 1 models, the spatial distributions of landslide susceptibility were much higher than in the other models. The susceptibility values were mainly distributed around the two opposite extremes between 0 and 1. The PM 2 model had better precision, recall, sensitivity, specificity, and overall

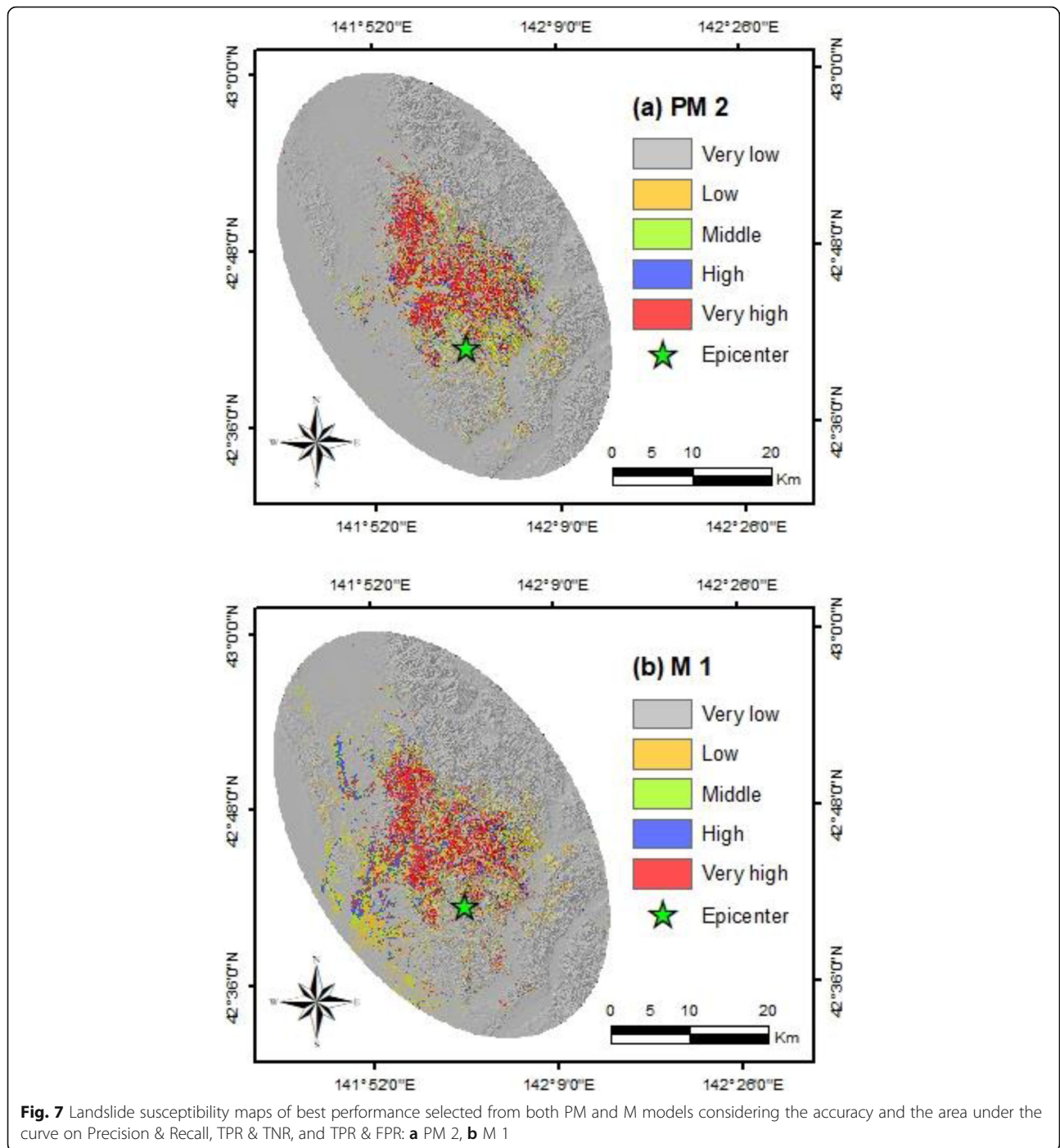


Fig. 7 Landslide susceptibility maps of best performance selected from both PM and M models considering the accuracy and the area under the curve on Precision & Recall, TPR & TNR, and TPR & FPR: **a** PM 2, **b** M 1

accuracy than other models did. In both regions, the models with sampling sizes greater than 1:3 showed poor classification performance (Table 1). The landslide susceptibility maps were produced differently depending on the sampling size used and the area selected. The sampling size and the area selected in PM and M may have affected various prediction abilities. In this study, the sampling size and the area resulted in different contributions to the models. In the autoencoder method,

the sampling ratio of 1:2 in the non-landslide points generated in the PM and M improved the prediction accuracy of landslide susceptibility mapping. The autoencoder effectively extracted a feature selection of spatial patterns using dimensionality reduction, and it significantly reduced the number of network parameters. Deep learning techniques could be used to explore the representation needed for making predictions based on raw data. Therefore, a promising avenue of research is to

Table 2 Variable importance and ranking for the combination zone including plain and mountain (PM)

Variable factors	Variable importance of the combination zone including plain and mountain (PM)											
	PM 1	Ranking	PM 2	Ranking	PM 3	Ranking	PM 4	Ranking	PM 5	Ranking	PM 6	Ranking
Elevation	0.065	9	0.104	4	0.098	5	0.058	9	0.073	7	0.057	9
Slope	0.076	7	0.088	5	0.068	8	0.072	8	0.071	8	0.089	5
Plan curvature	0.028	12	0.027	12	0.016	12	0.022	11	0.023	12	0.024	12
Profile curvature	0.041	11	0.040	11	0.031	11	0.020	12	0.034	11	0.024	11
NDVI	0.067	8	0.074	9	0.081	7	0.053	10	0.100	4	0.077	7
Lithology (geology)	0.263	1	0.080	6	0.121	3	0.113	4	0.082	5	0.108	4
Age (geology)	0.096	3	0.077	7	0.084	6	0.118	3	0.041	10	0.087	6
Stream density	0.058	10	0.075	8	0.066	9	0.103	5	0.082	6	0.072	8
Distance to stream	0.085	5	0.045	10	0.042	10	0.085	6	0.056	9	0.049	10
Distance to fault	0.138	2	0.117	3	0.108	4	0.138	2	0.157	1	0.132	2
Distance to epicenter	0.082	6	0.135	2	0.144	1	0.141	1	0.140	2	0.171	1
PGA	0.096	4	0.139	1	0.140	2	0.077	7	0.140	3	0.108	3

explore the probability of applying powerful deep learning methods to landslide susceptibility mapping (Wang et al. 2019). Moreover, various sampling strategies present varying accuracy in landslide susceptibility mapping, and it is recommended that sampling strategies be considered in applying new statistical techniques.

Conclusion

This research investigated the application of an autoencoder framework and 12 sampling strategies to landslide susceptibility mapping in Iburi region of Hokkaido in Northern Japan. The validation of the results was conducted based on the objective measures of the area under the curve in precision and recall, TPR and TNR, TPR and FPR, and accuracy. The experimental results led to the following conclusions. First, various sampling

strategies showed improved accuracy in landslide susceptibility assessment, which should be analyzed and compare with precision and recall curve in imbalanced data. Second, the prediction results obtained using the proposed autoencoder framework had good performance in terms of the area under the curve in precision and recall, TPR and TNR, TPR and FPR, and accuracy. Third, in the various models, the best performance by means of confusion matrix, especially in areas where landslides were triggered by the earthquake, was to utilize a sampling ratio of 1:2 of landslides and non-landslides generated in the PM. Finally, the prediction accuracies of landslide susceptibility mapping using the autoencoder model can be effectively improved by two strategies: (1) hyperparameter tuning in constructing the autoencoder architecture; (2) the selection of the tanh activation

Table 3 Variable importance and ranking for the mountainous only zone (M)

Variable factors	Variable importance of the mountainous only zone (M)											
	M 1	Ranking	M 2	Ranking	M 3	Ranking	M 4	Ranking	M 5	Ranking	M 6	Ranking
Elevation	0.080	7	0.075	6	0.067	7	0.090	5	0.065	9	0.071	7
Slope	0.068	9	0.070	7	0.039	10	0.049	8	0.037	10	0.034	10
Plan curvature	0.021	12	0.032	12	0.025	11	0.034	11	0.027	11	0.031	11
Profile curvature	0.044	11	0.033	11	0.021	12	0.033	12	0.021	12	0.024	12
NDVI	0.099	5	0.113	4	0.074	6	0.111	4	0.129	2	0.095	4
Lithology (geology)	0.085	6	0.104	5	0.105	4	0.082	6	0.091	7	0.095	5
Age (geology)	0.114	2	0.069	8	0.091	5	0.072	7	0.095	5	0.093	6
Stream density	0.072	8	0.058	9	0.058	8	0.045	10	0.089	8	0.067	8
Distance to stream	0.063	10	0.057	10	0.057	9	0.047	9	0.095	5	0.044	9
Distance to fault	0.109	4	0.124	2	0.170	1	0.166	1	0.107	4	0.162	2
Distance to epicenter	0.134	1	0.144	1	0.167	2	0.150	2	0.127	3	0.103	3
PGA	0.110	3	0.120	3	0.126	3	0.121	3	0.149	1	0.181	1

function. The landslide susceptibility maps produced in this study could be useful for decision-makers, planners, and engineers in disaster planning to mitigate economic losses and casualties. In future research, the accuracy of the landslide susceptibility maps in this study could be enhanced by selecting the optimal sampling strategy and investigating highly efficient deep learning techniques.

Authors' contributions

KN and FW conducted the field investigation in the study area. FW provided guidance for the spatial relationship between non-landslides and landslides triggered by earthquakes, where landslides intensively occurred, or not, between the epicenter and the highest peak ground acceleration recorded station. KN carried out the landslide susceptibility assessment and produced landslide susceptibility maps using an autoencoder framework.

Funding

This study was financially supported by the fund "Initiation and motion mechanisms of long runout landslides due to rainfall and earthquake in the falling pyroclastic deposit slope area" (JSPS-B-19H01980, Principal Investigator: Fawu Wang).

Availability of data and materials

The DEM data utilized in this work is freely available from the Geospatial Information Authority of Japan (<https://fgd.gsi.go.jp/download/menu.php>). The landslide inventory was mapped based on the landslides published by the Geospatial Information Authority of Japan (<http://www.gsi.go.jp/BOUSAI/H30-hokkaidoiburi-east-earthquake-index.html>).

Competing interests

The authors declare that they have no competing interests.

Received: 19 October 2019 Accepted: 18 November 2019

Published online: 11 December 2019

References

- Aditian A, Kubota T, Shinohara Y (2018) Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. *Geomorphology* 318:101–111
- Alessandro T, Carla I, Carlo E, Gabriele SM (2015) Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). *Geomorphology* 249:119–136
- Althuwaynee OF, Pradhan B, Lee S (2016) A novel integrated model for assessing landslide susceptibility mapping using CHAID and AHP pair-wise comparison. *Int J Remote Sens* 37(5):1190–1209
- Arnone E, Francipane A, Scarbaci A, Puglisi C, Noto LV (2016) Effect of raster resolution and polygon-conversion algorithm on landslide susceptibility mapping. *Environ Model Softw* 84:467–481
- Ayalew L, Yamagishi H (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* 65:15–31
- Bai S, Wang J, Lü G, Zhou P, Hou S, Xu S (2010) GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the three gorges area, China. *Geomorphology* 115:23–31
- Ballabio C, Sterlacchini S (2012) Support vector machines for landslide susceptibility mapping: the Staffora River Basin case study, Italy. *Math Geosci* 44(1):47–70
- Charte D, Charte F, García S, Jesus MJ, Herrera F (2018) A practical tutorial on autoencoders for nonlinear feature fusion: taxonomy, models, software and guidelines. *Inf Fusion* 44:78–96
- Chen H, Zeng Z, Tang H (2015) Landslide deformation prediction based on recurrent neural network. *Neural Process Lett* 41(2):169–178
- Chen W, Panahi M, Tsangaratos P, Shahabi H, Ilija I, Panahi S, Li S, Jaafari A, Ahmadg BB (2019) Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility. *Catena* 172:212–231
- Colkesen I, Sahin EK, Kavzoglu T (2016) Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression. *J Afr Earth Sci* 118:53–64
- Dagdelenler G, Nefeslioglu HA, Gokceoglu C (2016) Modification of seed cell sampling strategy for landslide susceptibility mapping: an application from the eastern part of the Gallipoli peninsula (Canakkale, Turkey). *Bull Eng Geol Environ* 75(2):575–590
- Devkota KC, Regmi AD, Pourghasemi HR, Yoshida K, Pradhan B, Ryu IC, Dhital MR, Althuwaynee OF (2013) Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya. *Nat Hazards* 65(1):135–165
- Fujiwara S, Nakano T, Morishita Y, Kobayashi T, Yurai H, Une H, Hayashi K (2019) Detection and interpretation of local surface deformation from the 2018 Hokkaido Eastern Iburu Earthquake using ALOS-2 SAR data. *Earth Planets Space* 71:64
- Gedeon TD (1997) Data mining of inputs: analyzing magnitude and functional measures. *Int J Neural Syst* 8(2):209–218
- Guo C, David RM, Zhang Y, Wang K, Yang Z (2015) Quantitative assessment of landslide susceptibility along the Xianshuihe fault zone, Tibetan Plateau, China. *Geomorphology* 248:93–110
- Heckmann T, Gegg K, Gegg A, Becht M (2014) Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows. *Nat Hazards Earth Syst Sci* 14:259–278
- Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. *Science* 313:504–507
- Hong H, Pourghasemi HR, Pourtaghi ZS (2016) Landslide susceptibility assessment in Lianhua County (China): a comparison between a random forest data mining technique and bivariate and multivariate statistical models. *Geomorphology* 259:105–118
- Huang L, Xiang LY (2018) Method for meteorological early warning of precipitation-induced landslides based on deep neural network. *Neural Process Lett* 48(2):1243–1260
- Kimura G (1994) The latest Cretaceous-early Paleogene rapid growth of accretionary complex and exhumation of high pressure series metamorphic rocks in Northwestern Pacific margin. *J Geophys Res Solid Earth* 99(B11):22147–22164
- King G, Zeng L (2001) Logistic regression in rare events data. *Polit Anal* 9:137–163
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436
- Lee S (2005) Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *Int J Remote Sens* 26:1477–1491
- Lee S, Talib JA (2005) Probabilistic landslide susceptibility and factor effect analysis. *Environ Geol* 47(7):982–990
- Liu Y, Wu L (2016) Geological disaster recognition on optical remote sensing images using deep learning. *Proc Comput Sci* 91:566–575
- Lu P, Qin Y, Li Z, Mondini AC, Casagli N (2019) Landslide mapping from multi-sensor data through improved change detection-based Markov random field. *Remote Sens Environ* 231:1–17
- Mabu S, Fujita K, Kuremoto T (2019) Disaster area detection from synthetic aperture radar images using convolutional autoencoder and one-class SVM. *J Robot Network Artif Life* 6(1):48–51
- Melchiorre C, Matteucci M, Azzoni A, Zanchi A (2008) Artificial neural networks and cluster analysis in landslide susceptibility zonation. *Geomorphology* 94(3–4):379–400
- Meten M, Prakash B, Yatabe R (2015) Effect of landslide factor combinations on the prediction accuracy of landslide susceptibility maps in the Blue Nile gorge of Central Ethiopia. *Geoenviron Disaster* 2(1):9
- Meusburger K, Alewell C (2009) On the influence of temporal change on the validity of landslide susceptibility maps. *Nat Hazards Earth Syst Sci* 9:1495–1507
- Park I, Choi J, Lee M, Lee S (2012) Application of an adaptive neuro-fuzzy inference system to ground subsidence hazard mapping. *Comput Geosci* 48:228–238
- Pradhan B, Lee S, Buchroithner MF (2010) GIS-based back-propagation neural network model and its cross-application and validation for landslide susceptibility analyses. *Comput Environ Urban Syst* 34:216–235
- Raja NB, Cicek I, Turkoglu N, Aydin O, Kawasaki A (2017) Landslide susceptibility mapping of the Sera River Basin using logistic regression model. *Nat Hazards* 85:1323–1346
- Romero A, Gatta C, Camps-Valls G (2016) Unsupervised deep feature extraction for remote sensing image classification. *IEEE Trans Geosci Remote Sens* 54(3):1349–1362
- Roy J, Saha S (2019) Landslide susceptibility mapping using knowledge driven statistical models in Darjeeling District, West Bengal, India. *Geoenviron Disaster* 6:1–18
- Saito H, Nakayama D, Matsuyama H (2009) Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: the Akaishi Mountains, Japan. *Geomorphology* 109(3):108–121

- Tajika J, Ohtsu S, Inui T (2016) Interior structure and sliding process of landslide body composed of stratified pyroclastic fall deposits at the Apporo 1 archaeological site, southeastern margin of the Ishikari Lowland, Hokkaido, Northern Japan. *J Geol Soc Jpn* 122(1):23–35
- Tamaki M, Kusumoto S, Itoh Y (2010) Formation and deformation processes of late Paleogene sedimentary basins in southern Central Hokkaido, Japan: paleomagnetic and numerical modeling approach. *Island Arc* 19(2):243–258
- Tien Bui D, Pradhan B, Lofman O, Revhaug I (2012) Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naïve Bayes models. *Math Probl Eng* 2012:1–26
- Trigila A, Iadanza C, Esposito C, Scarascia-Mugnozza G (2015) Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in Giampileri (NE Sicily, Italy). *Geomorphology* 249(15):119–136
- Tsangaratos P, Ilia I (2016) Comparison of a logistic regression and naïve Bayes classifier in landslide susceptibility assessments: the influence of models complexity and training dataset size. *Catena* 145:164–179
- Tsangaratos P, Ilia I, Hong H, Chen W, Xu C (2017) Applying information theory and GIS-based quantitative methods to produce landslide susceptibility maps in Nancheng County, China. *Landslides* 14(3):1091–1111
- Van Den Eeckhaut M, Marre A, Poesen J (2010) Comparison of two landslide susceptibility assessments in the Champagne-Ardenne region (France). *Geomorphology* 115(1–2):41–155
- Vijith H, Dodge D (2019) Modelling terrain erosion susceptibility of logged and regenerated forested region in northern Borneo through the Analytical Hierarchy Process (AHP) and GIS techniques. *Geoenviron Disaster* 6:1–18
- Wang Y, Fang Z, Hong H (2019) Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci Total Environ* 666:975–993
- Westen CJV, Rengers N, Soeters R (2003) Use of geomorphological information in indirect landslide susceptibility assessment. *Nat Hazards* 30(3):399–419
- Xiao L, Zhang Y, Peng G (2018) Landslide susceptibility assessment using integrated deep learning algorithm along the China-Nepal highway. *Sensors* 18:1–13
- Xu C, Dai F, Xu X, Lee YH (2012) GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. *Geomorphology* 145–146:70–80
- Yalcin A (2008) GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. *Catena* 72:1–12
- Yang BB, Yin KL, Lacasse S, Liu ZQ (2019) Time series analysis and long short-term memory neural network to predict landslide displacement. *Landslides* 16(4):677–694
- Yao X, Tham LG, Dai FC (2008) Landslide susceptibility mapping based on support vector machine: a case study on natural slopes of Hong Kong, China. *Geomorphology* 101:572–582
- Yeon Y, Han J, Ryu K (2010) Landslide susceptibility mapping in Injae, Korea, using a decision tree. *Eng Geol* 116:274–283
- Yilmaz I (2010) Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environ Earth Sci* 61:821–836
- Yilmaz C, Topal T, Süzen ML (2012) GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak-Turkey). *Environ Earth Sci* 65(7):2161–2178
- Yu S, Príncipe JC (2019) Understanding autoencoders with information theoretic concepts. *Neural Netw* 117:104–123
- Zhang S, Wang FW (2019) Three-dimensional seismic slope stability assessment with the application of Scoops3D and GIS: a case study in Atsuma, Hokkaido. *Geoenviron Disaster* 6:1–14
- Zhang S, Li R, Wang FW, Iio A (2019) Characteristics of landslides triggered by the 2018 Hokkaido Eastern Iburi earthquake, Northern Japan. *Landslides* 16(9):1691–1708
- Zhou S, Fang L (2015) Support vector machine modeling of earthquake-induced landslides susceptibility in central part of Sichuan province, China. *Geoenviron Disaster* 2:1–12
- Zhu X, Miao Y, Yang L, Bai S, Liu J, Hong H (2018) Comparison of the presence-only method and presence-absence method in landslide susceptibility mapping. *Catena* 171:222–233

Publisher's Note

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

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)