

An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan

Jie Dou¹ · Hiromitsu Yamagishi² · Hamid Reza Pourghasemi³ ·
Ali P. Yunus¹ · Xuan Song⁴ · Yueren Xu⁵ · Zhongfan Zhu⁶

Received: 17 February 2015 / Accepted: 9 May 2015 / Published online: 19 May 2015
© Springer Science+Business Media Dordrecht 2015

Abstract The objective of this study was to select the maximum number of correlated factors with landslide occurrence for slope-instability mapping and assess landslide susceptibility on Osado Island, Niigata Prefecture, Central Japan, integrating two techniques, namely certainty factor (CF) and artificial neural network (ANN), in a geographic information system (GIS) environment. The landslide inventory data of the National Research Institute for Earth Science and Disaster Prevention (NIED) and a 10-m digital elevation model (DEM) from the Geographical Survey of Institute, Japan, were analyzed. Our study identified fourteen possible landslide-conditioning factors. Considering the spatial auto-correlation and factor redundancy, we applied the CF approach to optimize these set of variables. We hypothesize that if the thematic factor layers of the CF values are positive, it implies that these conditioning factors have a correlation with the landslide occurrence. Therefore, based on this assumption and because of their positive CF values, six conditioning factors including slope angle (0.04), slope aspect (0.02), drainage density network (0.34), distance to the geologic boundaries (0.37), distance to fault (0.35), and lithology (0.31) have been selected as landslide-conditioning factors for further analysis. We partitioned the data into two groups: 70 % (520 landslide locations) for model training and the remaining 30 % (220 landslide locations) for validation. Then, a common ANN model, namely the back-propagation neural network (BPNN), was employed to produce the

✉ Jie Dou
douj88@gmail.com; doujie@csis.u-tokyo.ac.jp

¹ Department of Natural Environmental Studies, The University of Tokyo, Kashiwa 277-8568, Japan

² Asian Institute of Space Information, Shiroishi-ku, Hngō-dori 2-choume kita 3-10, Sapporo 003-0025, Japan

³ Department of Natural Resources and Environment, College of Agriculture, Shiraz University, Shiraz, Iran

⁴ Center for Spatial Information Science, The University of Tokyo, 5-1-5, Kashiwa 277-8568, Japan

⁵ Institute of Earthquake Science, China Earthquake Administration, Fuxing Rd. 36, Beijing, China

⁶ College of Water Sciences, Beijing Normal University, Xinjiekouwai Street 19, Beijing 100875, China

landslide susceptibility maps. The receiver operating characteristics including the area under the curve (AUC) were used to assess the model accuracy. The validation results indicate that the values of the AUC at optimized and non-optimized BPNN were 0.82 and 0.73, respectively. Hence, it is concluded that the optimized factor model can provide superior accuracy in the prediction of landslide susceptibility in the study area. In this context, we propose a method to select the factors with landslide occurrence. This work is fundamental for further study of the landslide susceptibility evaluation and prediction.

Keywords Landslide susceptibility · Certainty factor · BPNN · AUC · Osado Island

1 Introduction

Japan is prone to several types of natural disasters that include earthquakes, tsunamis, volcanic eruptions, and typhoons. In addition, the region is also susceptible to many landslides, owing to their topography and position in the world. The report of Fire and Disaster Management Agency of Japan (FMAJ) shows that landslides have caused considerable damage to life and property in the country, more than any other type of natural hazards occurred in recent decades. This trend is expected to continue and may increase in the future with the influence of unplanned urbanization and economic development, deforestation, and increased regional precipitation in landslide-prone areas owing to climate change. To mitigate from serious disasters, landslide susceptibility, hazard, and risk must be predicted (Turner and Schuster 1996). Hence, landslide occurrence prediction has become an important and challenging issue in the risk mitigation research (Chang and Chao 2006; Dou et al. 2014).

Landslide susceptibility is defined as the likelihood of a landslide occurring in an area based on local terrain conditions (Brabb 1984). The landslide susceptibility maps (LSM) delineate the earth's surface into zones of varying degrees of stability based on the evaluated significance of the conditioning factors inducing instability (Gökceoglu and Aksoy 1996; Youssef et al. 2015). LSM plays an important role in risk mitigation. They also provide an important basis for the measure aimed at decreasing the risks associated with landslides (Ayalew et al. 2005; Dou et al. 2014). It involves not only finding where the risk of landslide-related problems is spatially located, but also qualitatively and quantitatively assessing the significance of any such hazards and associated risk factors. Moreover, LSM could be employed to describe known landslides, assist with emergency decisions, and avoid and mitigate future landslide hazards. Moreover, LSM employs handling, processing, and storing an enormous amount of territorial data related to geographic information systems (GIS). GIS has also proved to be a powerful tool in landslide susceptibility evaluation (Aleotti and Chowdhury 1999; Dou et al. 2009; Nefeslioglu et al. 2013; Dou et al. 2015b).

In recent years, landslide susceptibility assessment has attracted the attention of many scholars, and numerous studies have been undertaken for LSM assessment around the world (Vahidnia et al. 2010; Pavel et al. 2011; Park et al. 2012; Tien Bui et al. 2012; Feizizadeh et al. 2014; Dou et al. 2015c). A majority of the research has been based on establishing the relationship between the landslide-conditioning factors and landslide occurrence through spatial data analysis. These relationships can be characterized in terms of rankings or weight. In this context, Chauhan et al. (2010) and Li et al. (2012) studies

classified the data-driven approaches into two general categories: qualitative and quantitative. With the advancement in computer systems and GIS technologies, quantitative methods have become more popular than the qualitative methods, because the latter one is only subjectively valued (Miliaresis and Kokkas 2007; Li et al. 2012; Pourghasemi et al. 2013; Sumer and Turker 2013). Quantitative methods such as artificial neural network (ANN) are useful for problem solving and have been successfully applied in various scientific, engineering, and susceptibility evaluation applications. Such machine learning methods have achieved more satisfactory results of landslide susceptibility than those obtained from other statistical models, e.g., logistic regression (LR) that has the capability of analyzing complicated data at varied scales, such as continuous, binary, and categorical (Ayalew and Yamagishi 2005; Kanungo et al. 2006; Chauhan et al. 2010; Li et al. 2012; Conforti et al. 2014). Moreover, one of the significant advantages of ANN is that it imposes less restrictive requirements with respect to the available information concerning the character of the relationships between the processed data and the type of distribution. It has robustness to deal with noisy or incomplete patterns and high fault tolerance compared with LR (Song et al. 2005; Shi et al. 2012). Furthermore, ANN is a complex and flexible nonlinear statistical method and hence can obtain high prediction accuracy for a classification, particularly with adequate samples (Kanungo et al. 2006).

ANN consists of a number of neurons that perform the desired function in parallel to transform the input layer into output layers. In landslide studies, commonly used input neurons are elevation, slope angle, slope aspect, plan curvature, and distance to drainage networks. The neurons imply conditioning factors, and their selection can also influence the accuracy of the landslide susceptibility maps. Numerous studies have related several of conditioning factors to landslide occurrence (Oguchi 1997; Klimes 2013; Dou et al. 2015a). However, different studies use different conditioning factors to produce the landslide susceptibility maps. Ayalew and Yamagishi (2005) reported in their study that neither universal criteria nor guidelines are followed to select the landslide-conditioning factors. Thus, it is difficult to determine whether the collected factors are appropriate or not for a particular case. Moreover, factor redundancy and a method to quantify these conditioning factors pose a challenge (van Westen et al. 1997). Hence, based on the aforementioned problem, we propose a method to optimize the factors for landslide susceptibility assessment.

The objective of this study was to identify the maximum number of correlated factors for landslide susceptibility mapping. A quantitative method called certainty factor (CF) is applied to select the landslide-conditioning factors relative to landslide occurrence. We believe this step is essential for the further study of landslide susceptibility and prediction. We are not aware of any studies that employed CF to quantify the correlation of the factors with landslide occurrence. And then, a robust model, back-propagation neural network (BPNN), is used to forecast the landslide susceptibility. Hence, the primary difference between the present study and the approaches described in the aforementioned publications is the novel combination of the predictive capabilities of certainty factor and a GIS-based back-propagation neural network for landslide susceptibility assessment in the study area. BPNN is a relatively new model that has been applied previously to landslide susceptibility assessment. Herein, we hypothesize that if the CF value is positive, it means this factor has high relationship with the landslide occurrence. To confirm our assumption, we tested the proposed method in Osado region of Sado Island, Japan. Sado is an independent island offshore of the Niigata area, and many deep-seated landslides occur occasionally due to

heavy rainfall and melting snow. The high-density precipitation and rapid snowmelt can also cause flash floods and shallow landslides. Few landslide studies have been conducted in and around the island (Ayalew et al. 2005; Yamagishi 2008). To our knowledge, there is no reported landslide susceptibility mapping for the Osado region.

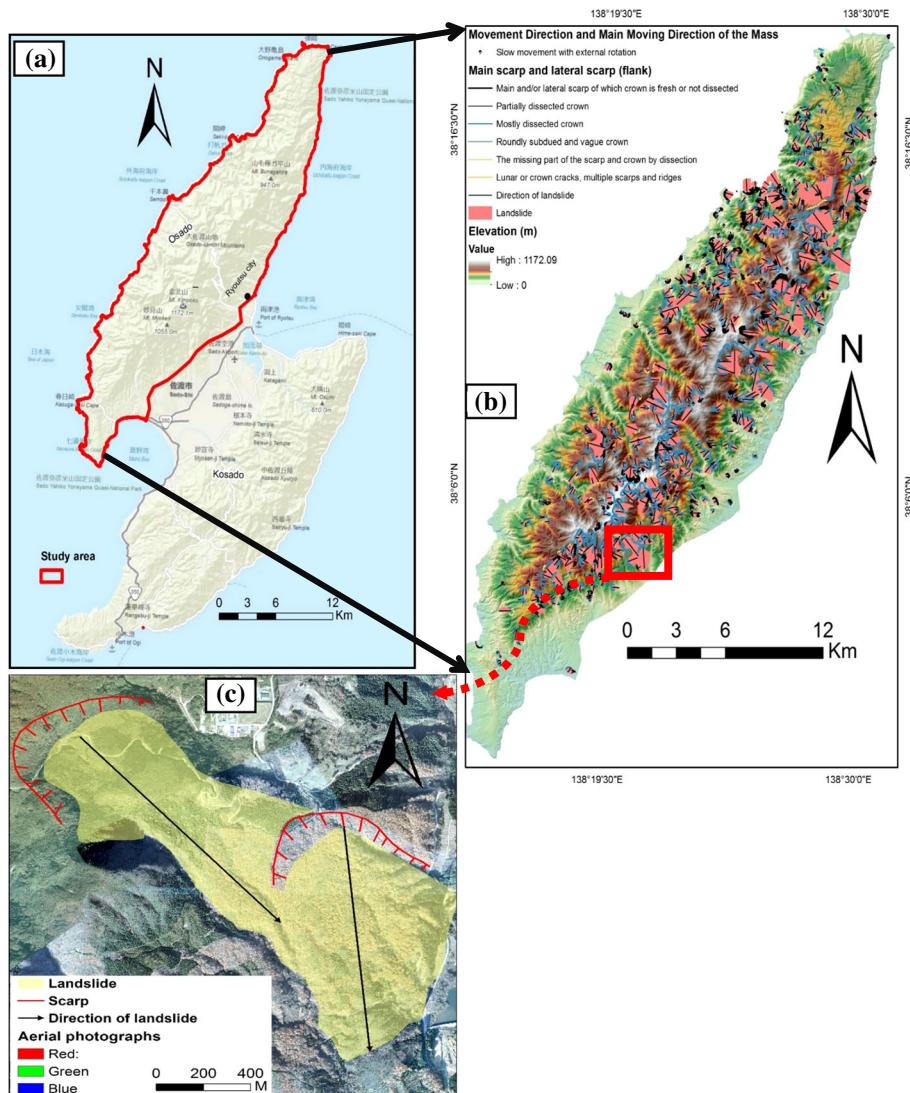


Fig. 1 **a** Location of the study area; **b** distribution of landslides, including the direction of landslide, the scarp (locations of landslide data from NIED); and **c** lower maps are enlargement of distribution of landslides (overlay the aerial photographs which from Midori Niigata and Sado city)

2 Geography and geology of the study area

The study area is located in a mountainous region of Sado Island of Niigata Prefecture, in the northwestern part of Japan (Fig. 1). The Island is composed of two parallel, elongated ridges approximately northeast (NE) to the southwest (SW) as shown in Fig. 1 left, which are locally named Osado and Kosado and cover approximately 856 km². This study concentrated on the Osado region of Sado Island. Osado is approximately 394 km², situated between 138°14'E and 138°32'E longitudes, and 37°57'N and 38°20'N latitudes. Climate is humid, subtropical with warm summers and cold winters. The average temperature is 13.7 °C, and the annual precipitation is approximately 1550 mm (Yamagishi 2008). The elevation varies from sea level to 1172 m with a mean of 333 m. The highest point of the island is Mt. Kimpoku in Osado. Geology of the island is mainly Neogene terrestrial and marine volcanic: rhyolitics (dacites) and andesites, associated with pyroclastics and rhyolitic intrusives. Some of the coastal slopes consist of recently formed semi-consolidated and unconsolidated sand deposits and gravel. Consisting principally of mountains, hills, and tectonic movements evidenced by upheaved benches and active faults, study area is prone to landslide and debris disasters (Yamagishi 2008). In the study area, the landslide types are mostly deep-seated, translational, and rotational slides (Fig. 2). Ayalew et al. (2005) noted that rock falls are seldom occurred in this region. Most of the deep-seated landslides are inventoried in the rhyolitic and dacitic lithologies. The landslide susceptibility may also be influenced by rainfall, snow melting, geology, slope aspect, and slope angle.

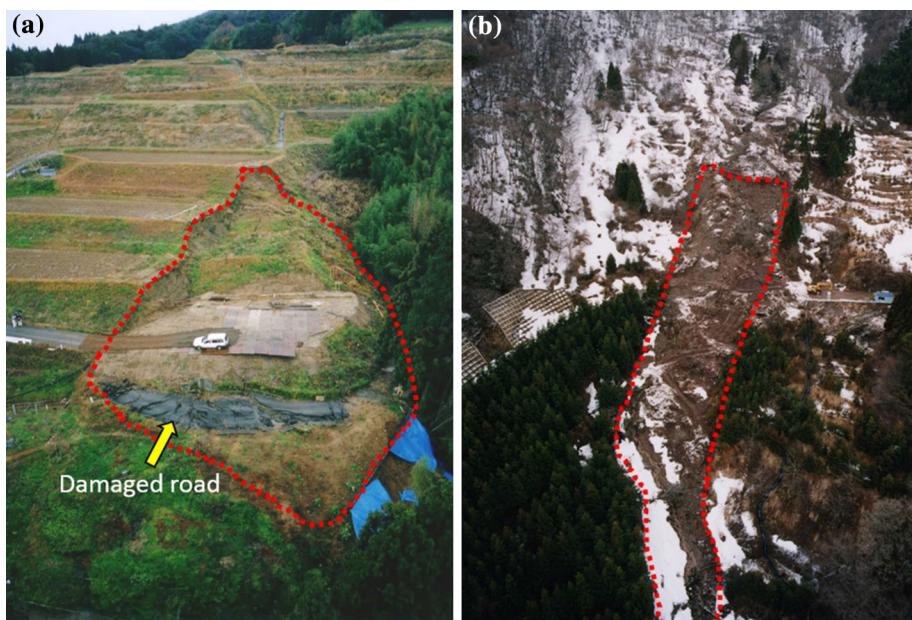


Fig. 2 Examples of identified major landslide types in the Sado Island: **a** rotational slide that severely damaged the roads, **b** translational slide in the hilly terrains that carved the dense vegetation

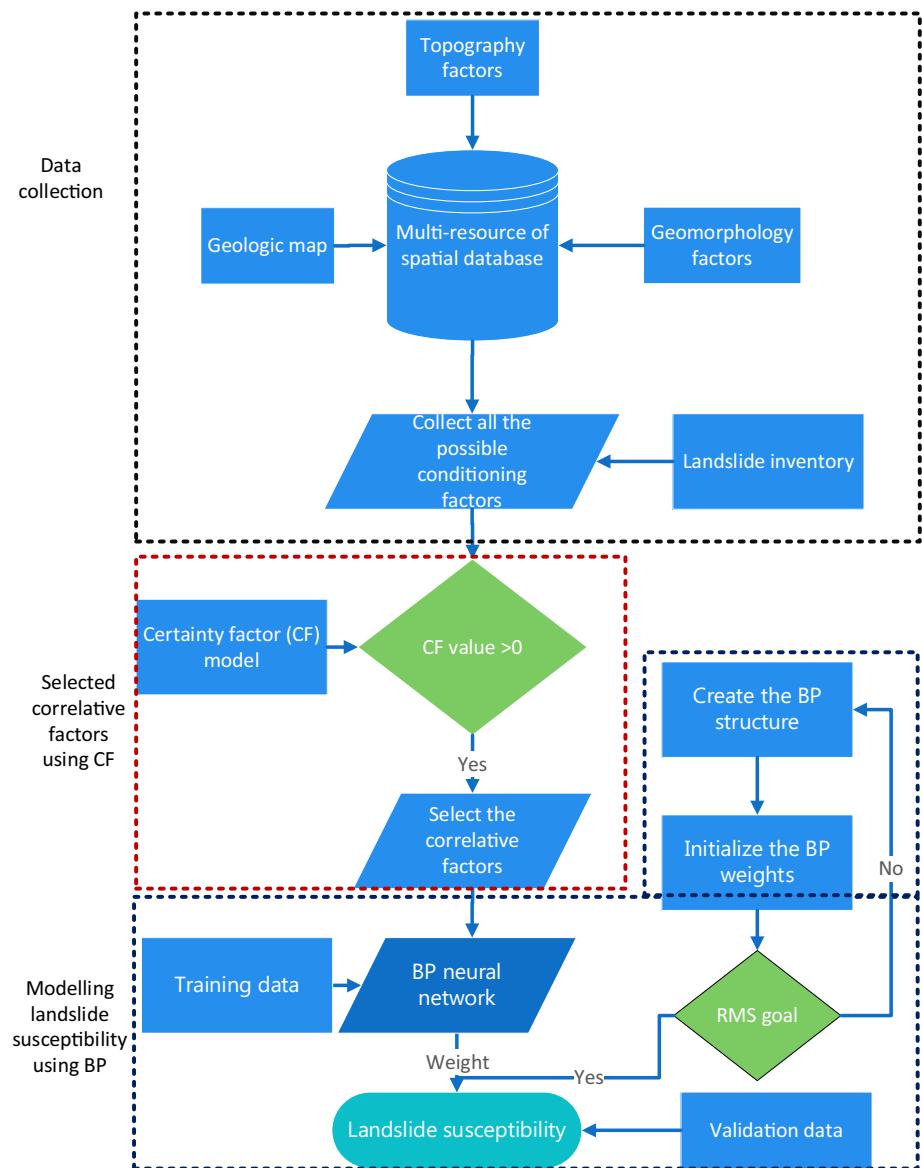


Fig. 3 Outline of hybrid structure of CF and BPNN procedures composed of three phases: (1) data collection and analysis, (2) selection of correlative factors with landslide occurrence by CF model, and (3) BPNN neural networks training and prediction

3 Methodology

Figure 3 is an overview of the approach that was applied for the landslide susceptibility mapping in the study area. The flowchart consists of three phases: (1) data integration and analysis, (2) selection of conditioning factors with landslide occurrence using CF model, and (3) ANN training and prediction. Each method is presented in the following sections.

3.1 Data integration and analysis

Past landslides are vital to predict reactivation in the future (Guzzetti et al. 1999). A landslide inventory map is therefore a basic requirement for landslide analysis. In the present study, we used the landslide inventory data provided by the National Research Institute for Earth Science and Disaster Prevention (NIED), Japan. The inventory was prepared by NIED experts from multi-temporal aerial photographs. Based on the NIED inventory maps, we manually draw the direction of landslides. After careful consideration, a total of 732 deep-seated landslides (debris, not scarps) were selected (Fig. 1) to analyze the relationship between the landslide occurrences and conditioning factors. The landslide inventory was provided as polygons of the debris area, the centroids of which were utilized as landslide points converted from the polygons by geometry conversion as mentioned in Pradhan et al. (2010). The sizes of the smallest and largest landslides are approximately 0.0007 and 1.65 km², respectively. The total area of the 732 landslide debris (including deep-seated and shallow slides) covers 57.33 km² of the study area, which is approximately 14.55 % of the total land area. A digital elevation model (DEM) with 10-m resolution provided by the Geographical Survey Institute (GSI) of Japan and the digital geological map (scale 1:200,000) obtained from GeoNavi of the Geological Survey of Japan (GSJ) was used to prepare the causative factors. The geology map originally included 53 lithological units; however, the study area and its adjacent area are represented primarily by eight units: (a) Metamorphic, (b) Plutonic and intrusives, (c) Sedimentary (mudstone), (d) Sedimentary (sandstone), (e) Sedimentary (slate and sandstone), (f) Volcanic (andesite lava), (g) Volcanic (basalt), (h) Volcanic (dacite lava), (i) Volcanic (dacite), and (j) Volcanic (rhyolite lava) as displayed in Fig. 4e.

The conditioning factors selected in this study were based on the spatial relationships between landslide occurrence and conditioning factors observed in the past literature. It includes topography, hydrology, geology, tectonics, and geomorphology (Kanungo et al. 2006; Dou et al. 2009; Farrokhzad et al. 2011; Klimes 2013). A total of fourteen landslide-conditioning factors were chosen as indicated (Figs. 4, 5): (1) elevation, (2) slope angle, (3) slope aspect, (4) total curvature, (5) plan curvature, (6) profile curvature, (7) compound topographic index (CTI), (8) stream power index (SPI), (9) drainage density networks, (10) distance from drainage network, (11) density of geological boundary, (12) distance to geological boundaries, (13) distance to fault, and (14) lithology. All these factors are processed and investigated with the help of ArcGIS 10.2v software. Among these factors, CTI and SPI, developed by Beven and Kirkby (1979) and Gessler et al. (1995), respectively, are used in this study:

$$\text{CTI} = \ln(\text{As}/\tan \beta) \quad (1)$$

$$\text{SPI} = \text{As} \times \tan \beta \quad (2)$$

where As is the specific catchment area per unit channel width orthogonal to the flow direction (m²/m) and β is the slope angle expressed in degrees.

3.2 Certainty factor model

The certainty factor (CF) is a method for managing uncertainty in rule-based systems and has been widely applied in expert system shell fields, in addition to medical diagnosis studies (Shortliffe 1976; Binaghi et al. 1998; Kanungo et al. 2011; Pourghasemi et al. 2013). The CF approach is one of the most widespread methods for representing and

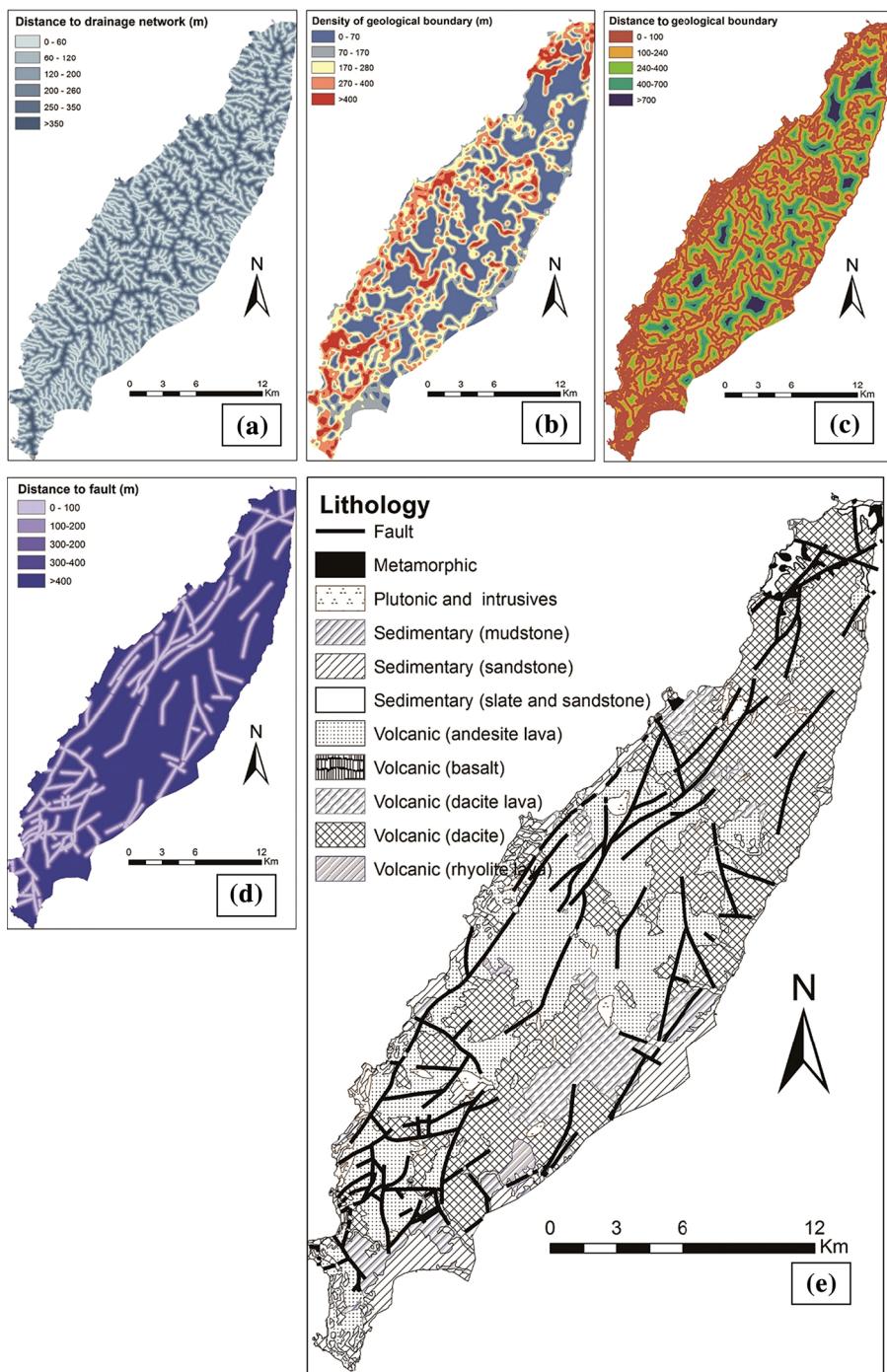


Fig. 4 Landslide-conditioning factors: **a** distance to drainage network, **b** density of geological boundary, **c** distance to geological boundary, **d** distance to fault, and **e** geological map of study area (modified from the Geological Survey of Japan, AIST, 2004)

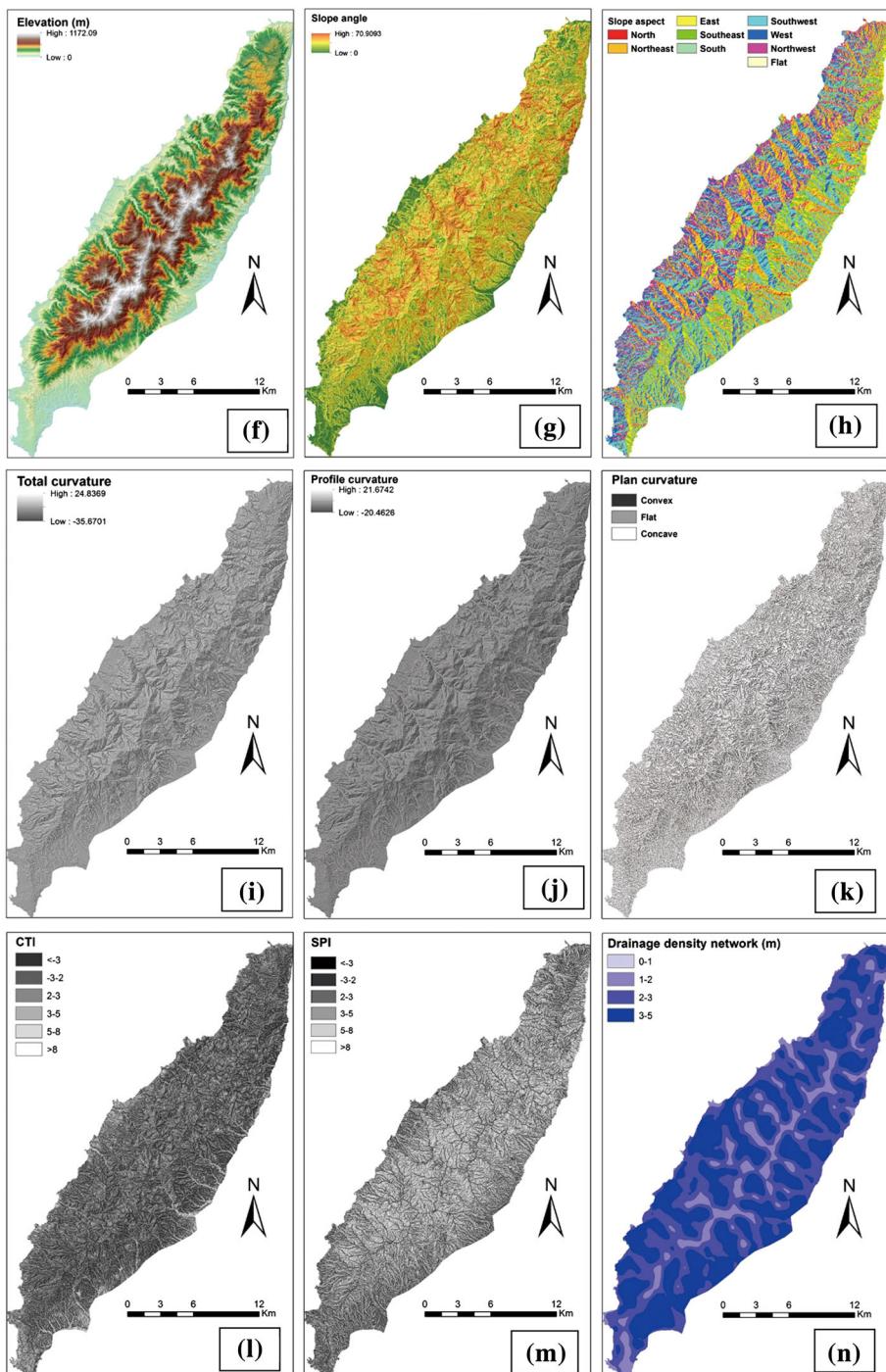


Fig. 5 Landslide-conditioning factors: **f** elevation, **g** slope angle, **h** slope aspect, **j** total curvature, **k** plan curvature, **l** profile curvature, **m** compound topographic index (CTI), **n** stream power index (SPI), **n** drainage density networks

manipulating uncertainty knowledge in the rule-based expert system. It also contains one of the probable favorability functions (FF) to address the problem of integrating heterogeneous data (Chung and Fabbri 1993). The general theory function of CF is given by the following equation:

$$CF = \begin{cases} \frac{PP_a - PP_s}{PP_a * (1 - PP_s)} & \text{if } PP_a \geq PP_s \\ \frac{PP_a - PP_s}{PP_s * (1 - PP_a)} & \text{if } PP_a < PP_s \end{cases} \quad (3)$$

where PP_a is the conditional probability of having a number of landslide events occur in class a and PP_s is the prior probability of having a total number of landslide events in the study area. PP_s for this study was calculated to be 0.015.

The CF value is first calculated for each class layer using Eq. 1. These layers are then integrated pairwise. The integration of two CFs, X and Y , with two layers of information, is expressed as Z based upon the combination rule given in the following equation (Binaghi et al. 1998):

$$Z = \begin{cases} X + Y - X * Y & X, Y \geq 0 \\ X + Y + X * Y & X, Y < 0 \\ \frac{X + Y}{1 - \min(|X|, |Y|)} & X, Y, \text{ opposite signs} \end{cases} \quad (4)$$

Equation 2 is also called the parallel-combination function. The pairwise combination is executed until all the CF layers are grouped together. The CF value ranges between –1 and 1. A positive CF implies an increasing certainty in landslide occurrence, whereas negative values infer a decrease in the certainty. A value close to zero indicates that the prior probability is close to the conditional probability, and it is therefore difficult to determine the certainty of landslide occurrence (Pourghasemi et al. 2012). The preparations of the data layers and their selection for the analysis were performed with the ArcGIS 10.2 software. The favorability values are obtained by overlaying the landslide inventory map and each data layer and computing the landslide frequency.

3.3 Back-propagation for feed forward neural network

Artificial neural networks (ANN) were applied to model the landslide susceptibility on Osado Island. An ANN can be regarded as a quantitative black-box model approach that emulates human pattern recognition functions (Aleotti and Chowdhury 1999). Moreover, earth science's nonlinearity analysis and prediction can be studied applying this efficient tool. ANN has also been successfully implemented for evaluating landslide susceptibility by several researchers (Arora et al. 2004; Falaschi et al. 2009; Pradhan et al. 2010; Zare et al. 2013). Our study focuses on a particular type of ANN model, known as a back-propagation neural network. BPNN algorithm is used in the feed forward ANN. This is typically used to train the network among the different types of ANN models including radial basis function (RBF), general regression neural networks (GRNN), and probabilistic neural networks (PNN). The BPNN algorithm is simply a gradient-descent algorithm (also called a generalized delta rule) that uses to minimize the total error or mean error of target computed by the neural network. This algorithm is a neural network that is composed of three layers: input, hidden, and output. The structure of a typical three-layer BPNN is displayed in Fig. 6. The input layer propagates components of a special input vector after

weighting synaptic weights to each node in the hidden layer. Each hidden layer computes outputs corresponding to these weighted sums through a nonlinear/linear function, e.g., log-sigmoid, purelin, or tan-sigmoid (Yesilnacar and Topal 2005; Prasad et al. 2012). The BPNN algorithm comprises of two paths: feed forward and backward path. The feed forward is expressed as follows (Rumelhart et al. 1986):

$$y_j = F(X_j) = F\left(W_{oj} + \sum_{i=1}^I W_{ij}x_i\right) \quad (5)$$

$$Z_k = F(Y_k) = F\left(W_{ok} + \sum_{j=1}^J W_{jk}y_j\right) \quad (6)$$

where x_i , y_j , and Z_k represent the input, hidden, and output layers, respectively, W_{oj} and W_{ok} are the bias weights for setting the threshold values, X_j and Y_k represent temporarily

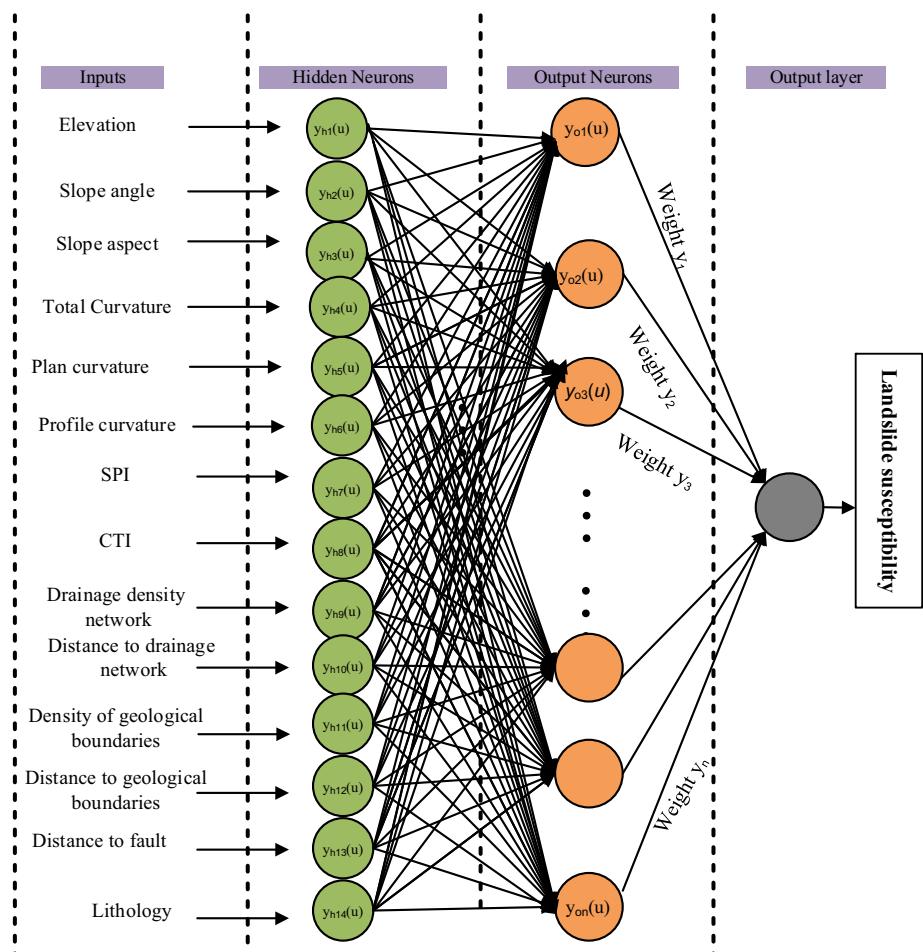


Fig. 6 Structure of a typical three-layer feed forward BPNN (multilayer perceptron) for landslide susceptibility analysis

computing results before using the activation function, and F is the activation function applied in the hidden and output layers. In this study, a sigmoid function or logistic function is chosen as the activation function. Thus, the output y_j and Z_k can be expressed as:

$$y_j = F(X_j) = F\left(\frac{1}{1 + e^{-X_j}}\right) \quad (7)$$

$$Z_k = F(Y_k) = F\left(\frac{1}{1 + e^{-Y_k}}\right) \quad (8)$$

The value of F ranges from 0 to 1.

For error back-propagation weight training, the error function can be defined as (Rumelhart et al. 1986):

$$E = \frac{1}{2} \sum_{k=1}^K e_k^2 = \frac{1}{2} \sum_{k=1}^K (t_k - z_k)^2 \quad (9)$$

where t_k and e_k are the predefined target value and error in each output node, respectively. The goal was to minimize E , the error between the desired and actual output values of the network. To adjust the weight, a gradient-descent strategy was used. The weight between the hidden and output layers can be expressed as follows:

$$\frac{\partial E}{\partial w_{jk}} = -e_k \frac{\partial F(Y_k)}{\partial Y_k} y_j = -e_k F'(Y_k) y_j = -\delta_k y_j \quad (10)$$

$$\delta_k = e_k F'(Y_k) = (t_k - z_k) F'(Y_k) \quad (11)$$

Therefore, the weight adjustment in the link can be computed by:

$$\Delta w_{jk} = \eta * y_j * \delta_k \quad (12)$$

where η is the learning rate with value ranges between 0 and 1. If the learning rate is relatively small, the BPNN is slow to converge the network. Conversely, a learning rate that is overly large can lead to a widely oscillating network. Thus, it is preferable to choose a single value throughout the experiment. The new weight herein is updated by the following equation:

$$w_{jk}(n+1) = w_{jk}(n) + \Delta w_{jk}(n) \quad (13)$$

Here n is the number of iterations in the network.

Similarly, the error gradient in links between the input and hidden layers can be derived from the partial derivative with respect to w_{ij} ,

$$\frac{\partial E}{\partial w_{ij}} = \left[\sum_{k=1}^K \frac{\partial E}{\partial z_k} \frac{\partial z}{\partial Y_k} \frac{\partial Y_k}{\partial y_j} \right] * \left(\frac{\partial y_j}{\partial X_j} \right) * \left(\frac{\partial X_j}{\partial w_{ij}} \right) = -\Delta_j x_i \quad (14)$$

$$\Delta_j = F'(X_j) \sum_{k=1}^K \delta_k w_{jk} \quad (15)$$

The new weight in the hidden and input links can be updated as:

$$\Delta w_{ij} = \eta * x_i * \Delta_j \quad (16)$$

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n) \quad (17)$$

In this study, a three-layer BPNN is employed as the network structure of 14–15–1, which composed of an input layer (14 neurons), hidden layer (15 neurons), and an output layer. To determine the number of hidden layer nodes, arbitrary squashing function randomly selected the number of hidden neurons (Hornik et al. 1989). Coefficient of determination (R^2) was used to evaluate the hidden layers. The equation of R^2 can be expressed:

$$R^2 = \left(n \sum_{i=1}^n \hat{y}_i y_i - \sum_{i=1}^n \hat{y}_i * \sum_{i=1}^n y_i \right)^2 / \left(n \sum_{i=1}^n \hat{y}_i^2 - \left(\sum_{i=1}^n \hat{y}_i \right)^2 \right) * \left(n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right) \quad (18)$$

where \hat{y}_i and y_i ($i = 1, 2, 3, \dots, n$) are the i th predicted vale and true value, respectively. R^2 value ranges from 0 to 1.

The input factors were normalized to eliminate the dominating effect of large values mainly due to the fragmented functions such that their values ranged from zero to one using the equation:

$$Y = \frac{X - X_{\text{Min}}}{X_{\text{Max}} - X_{\text{Min}}} \quad (19)$$

where Y is the normalized data value and X is the original data value.

All the input data in the BPNN model were normalized in the range [0, 1]. The landslides were divided into two groups; (a) for training the network and (b) for validating the model.

The initial weights were automatically assigned to the random values between 0.1 and 0.25. The parameters were adjusted as follows (Table 1): (1) initial learning rate (LR) for influencing the convergence of the network: 0.1, (2) number of epochs: 1500 iterations, (3) momentum parameters: 0.9 (to prevent instabilities caused by an excessively high LR), (4) activation (transfer) function for layers: transig for hidden layer, purelin for output layer, and (5) training function of networks: variable LR with momentum (traingdx). The value of the root mean square error (RMSE) goal for terminating the criterion was set to 0.001.

Table 1 Setting the parameters in the BPNN model

Parameters	Parameters setting
Structure of BPNN networks	–1
Initial weights	0.1–0.25
Learning rate	0.1
Number of epochs	1500 iterations
Momentum parameters	0.9
Activation (transfer) function for layers	Transig for hidden layer, purelin for output layer
Training function of networks:	Variable LR with momentum (traingdx)
Root mean square error (RMSE)	0.001

Table 2 continued

Conditioning factors	Class	No. of pixels in domain	Percentage of domain	Percentage of landslides	No. of landslides	PP _a	CF	Z
Lithology	Sedimentary (sandstone)	359,589	10.1774	3.9063	20	0.0056	-0.6196	0.3059
	Sedimentary (mudstone)	122,249	3.4600	2.1484	11	0.0090	-0.3825	
	Plutonic and intrusives	76,237	2.1577	2.3438	12	0.0157	0.0805	
	Volcanic (basalt)	11,143	0.3154	0.7813	4	0.0359	0.6051	
	Volcanic (rhyolite lava)	24,774	0.7012	0.3906	2	0.0081	-0.4465	
	Volcanic (dacite)	1,481,407	41.9280	47.8516	245	0.0165	0.1256	
	Volcanic (dacite lava)	224,403	6.3512	10.1563	52	0.0232	0.3802	
	Volcanic (andesite lava)	1,184,335	33.5201	29.8828	153	0.0129	-0.1099	
	Sedimentary (slate and sandstone)	31,106	0.8804	0.9766	5	0.0161	0.0999	
	Metamorphic	17,408	0.4927	1.5625	8	0.0460	0.6947	

4 Results and discussion

4.1 Resultant CF approaches

According to the range of CF values, the landslide susceptibility index (the LSI is defined as the sum of the ratio calculated relative to factors) can be classified into six susceptibility classes : (1) (-1.0)–(-0.09) (extremely low certainty), (2) (-0.09)–(0.09) (uncertainty), (3) (0.09–0.2) (low certainty), (4) (0.2–0.5) (medium certainty), (5) (0.5–0.8) (high certainty), and (6) (0.8–1.0) (extremely high certainty). The LSI can provide a measure of certainty in predicting landslides (Binaghi et al. 1998). The relationship between the spatial landslide locations and landslide-conditioning factors was analyzed in GIS (Table 2). Based on the calculation of the Z values, we determined that the six conditioning factors, slope angle (0.04), slope aspect (0.02), drainage density networks (0.34), distance to the geologic boundaries (0.37), distance to fault (0.35), and lithology (0.31), were positive. The results indicate that these six landslide-conditioning factors have a certainty correlation with landslide occurrence. Thus, we can conclude these factors are significantly correlated with the landslide occurrence. This result is the fundamental basis for the following study of the landslide susceptibility evaluation and prediction.

Of these six conditioning factors, slope angle is an important parameter influencing slope stability (Chauhan et al. 2010; Conforti et al. 2014). A slope angle of the isotropic material aggravated slope correlates with increased likelihood of a slope failure (Conforti et al. 2014). In the case of slope angle, the CF values were strongly positive from 10° to 30° with the maximum CF value 0.18. The percentage of landslide occurrence in the slope classes 10°–20° and 20°–30° was 21.11 and 23.7 %, respectively. The results indicate that the occurrence of landslides gradually decreases with an increasing slope angle until it becomes insignificant after a 30° slope angle. Gentler slopes have a relatively low frequency of landslide occurrences because of the lower shear stress corresponding to the low gradient; steep slope angles lead to outcropped bedrock that is less susceptible to landslides. Slope aspect also plays a significant role in landslide susceptibility assessment. Slope aspect influences several microclimatic factors, such as exposure to sunlight and winds, soil moisture, snow thickness, and rainfall intensity (Yilmaz 2009). With respect to the slope aspect, most of the landslides occurred along the east, southeast, and south-facing slopes with a positive CF values. The highest percentage of landslides (12.98 %) occurred along the northeastern slopes, followed by the southern slopes (12.63 %). Slope aspect of the Osado is featured by the northwest slopes, which have less thickness of snow than the southeast direction. The reason is that along the northeast slopes, snow usually was blown out by the wind from the northwest; consequently, the southeast slopes have much amount of snow that influences the snow melting leading to slide. Landslides in this hilly region may also occur because of the erosion activity associated with the drainage systems. The drainage density network (DR) showed a positive value for classes 2–3 and 3–5 m. The positive maximum CF value of 0.11 was observed with the 3–5 m DR class. The highest percentage of landslide occurrences was 41.4 %.

The distance to geological boundary showed positive values for classes 100–240, 240–400, 400–700, and more than 700 m. Of the positive CF values observed, the highest percentage of landslide occurrence was 24.22 %. Among, the maximum percentage of landslide occurrence was found at classes less than 100 m (38.4 %). This indicates that the closer the geological boundary, the maximum the landslide instability occurs. These phenomena can be attributed to the lithological resistance of rock types and the

environment of their emplacement. The inherent characteristics of rock types such as their structure and composition might result in the differential resistance and are widely regarded as a conditioning factor for landslide (Chauhan et al. 2010). With the respect to distance to fault, the positive CF values range from 0 to 300 m. It has the largest CF value (0.17) at 0–100 m. With increasing the distance, the CF value declines. The result of this study also indicates that the six lithology classes are highly prone to landslides, namely, (1) Plutonic and intrusives, (2) Volcanic (rhyolite lava), (3) Volcanic (dacite lava), (4) Volcanic (dacite), (5) Sedimentary (slate and sandstone), and (6) Metamorphic and have CF values of 0.08, 0.61, 0.13, 0.1, and 0.69, respectively. We noticed a higher frequency of landslides along the margins of the rhyolite lavas, dacite lavas, and dacites (>50%). These lavas were once emplaced in the submarine environment, and hence, we believe that their margins may be transformed into pelitic rocks (glass with much water), which were further altered to be abundant in smectite clay that are easily prone to sliding.

4.2 Resultant BPNN models

When applying an ANN-multilayer model with a feed forward error back-propagation learning algorithm, the data should be divided into two subsets: training data and testing data. Approximately 70 % of the entire dataset is considered sufficient for the training network. The remaining is normally reserved for testing the final structure of the ANN network (Rumelhart et al. 1986; Chauhan et al. 2010). In this study, the three layers of the BPNN were simulated using the Matlab R2012a software. To select the number of hidden neurons, we randomly set the neurons and run the 10 times of network to choose the average value of R^2 for reducing the effect of initial value in the BP network. From Fig. 7, it can be seen that when the number of hidden neurons is 10, R^2 has the highest values (0.96) at the six optimized factors as shown in Table 3, whereas at the fourteen factors, the number of hidden neurons (=5), R^2 has the largest value (0.94). Hence, the number of hidden neurons was selected as 15 and 10. The structures of the networks required in the BPNN were constructed as 14–15–1 and 6–10–1, respectively (input–hidden–output). The weights between the layers were obtained by training the BPNN network to compute the contribution of each landslide-conditioning factor. To test the stability of the BPNN model, it was iterated ten times, each with a random set of landslide data selected from the entire data pool. The results were found to be similar. The standard deviation was 0.0032. Hence,

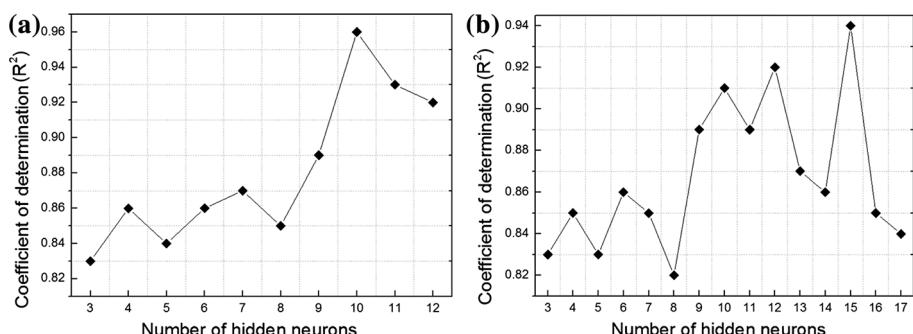


Fig. 7 Number of hidden neurons influences the back-propagation network: **a** Coefficient of determination R^2 has the largest value (0.96) when the number of hidden neurons is 10 at six conditioning factors; **b** R^2 has the largest value (0.94) when the number of hidden neurons is 15 at fourteen conditioning factors

the random samples did not have a significant influence on the results. In this case, the average values were computed only to interpret the results.

When the BPNN reached the minimum RMSE values (0.001), the entire study area was fed into the network to assess the landslide susceptibility mapping. The weights for minimum error were recorded in the process, and weights of each factor were determined for the study area. The set of landslide index values obtained in each pixel were then converted to raster in GIS environment. We used the original fourteen landslide-conditioning factors (elevation, slope angle, slope aspect, total curvature, plan curvature, profile curvature, compound topographic index (CTI), stream power index (SPI), drainage density network, distance to drainage network, distance to geological boundary, density of geological boundary, distance to fault, and lithology) and the selected six factors (slope angle, slope aspect, drainage density network, distance from the geologic boundary, distance to fault and lithology) based on an optimization to separately produce the landslide susceptibility mappings as shown in Fig. 8b. There are several classification methods available including natural break, quantiles, equal intervals, and standard deviations (Ayalew and Yamagishi 2005; Zare et al. 2013). In this study, after completing several trials, we employed natural break for the classification because natural break is suitable for big jumps in the data value and also identified the best group similar values that can maximize the dissimilarity between each classes for forming the homogeneity of the classification. The landslide susceptibility index values were classified into six classes: extremely low, low, moderate, high, very high, and extremely high for easier visual interpretation of the LSM. Figure 8 shows the spatial probability of landside occurrence, in the six classes, from extremely low (dark green) where landslides are not expected to extremely high (red) where landslides are probable. The lower maps are enlargements of a portion of the LSM maps obtained for the Osada Island, with six and fourteen factors. It can be observed that the landslide susceptibility rates increase with the increasing number of the indices. In the two susceptibility maps, there are locations where the differences are subtle, whereas other areas have noticeable differences in the distribution of the index values. There are additional red- or reddish-colored areas in the LSM map produced from the fourteen factors compared with the six factors; very high susceptibility occurs more frequently in the latter as shown in Fig. 9. Furthermore, the lower susceptibility areas (dark green) are more extensive in the map using six factors. Figure 10 indicates that 77.19 % of the total landslides occurred in the 59.11 % of the area classified as high (high, very high, and

Table 3 Results of 10 times running the back-propagation network model for selection of the number of hidden neurons inputting the six optimized factors

Number of hidden neurons	Coefficient of determination (R^2)		
	Minimum	Maximum	Average
3	0.67	0.93	0.83
4	0.73	0.95	0.86
5	0.68	0.92	0.84
6	0.64	0.91	0.86
7	0.76	0.94	0.87
8	0.69	0.92	0.85
9	0.63	0.94	0.89
10	0.84	0.97	0.96
11	0.78	0.96	0.93
12	0.74	0.95	0.92

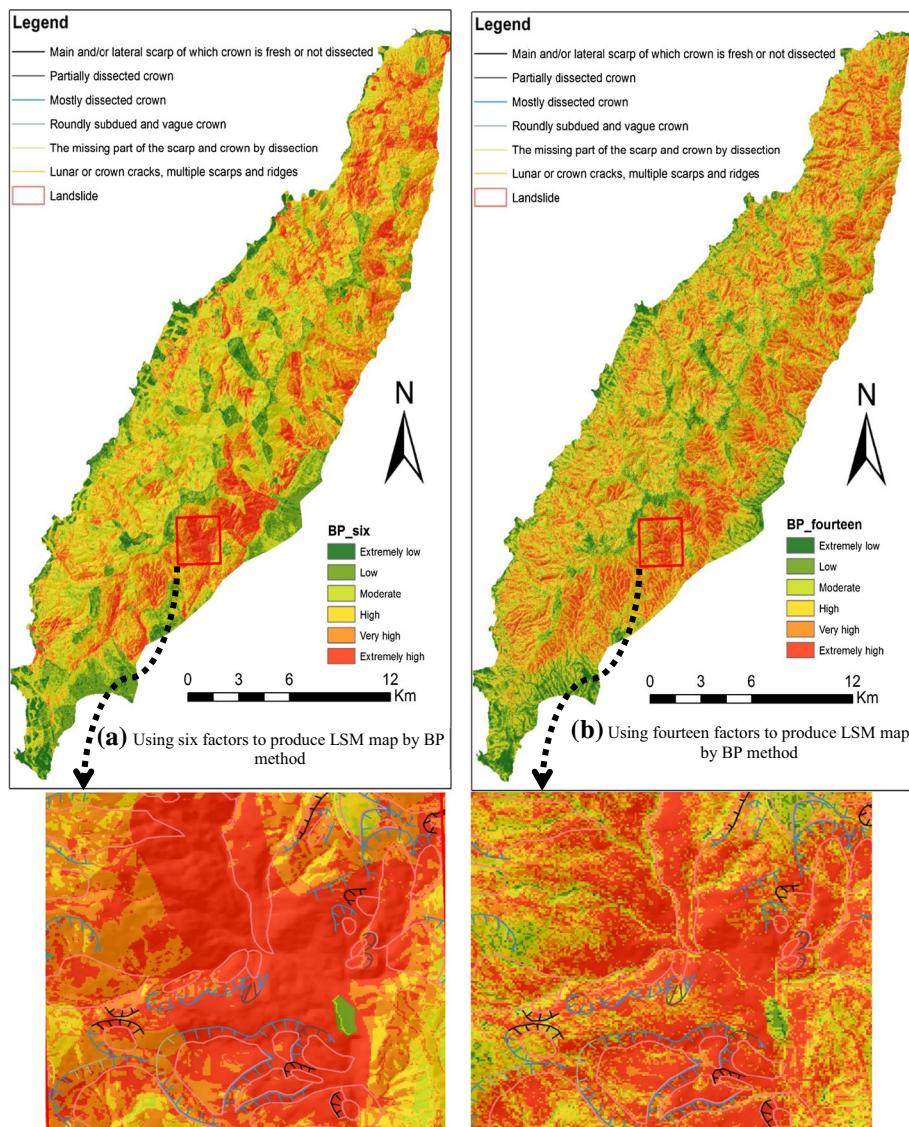


Fig. 8 LSM maps produced by the BPNN method: **a** six factors and **b** fourteen factors. Maps indicate the spatial probability of landslide occurrence in six classes. Lower maps are enlargements of the LSM maps. These two susceptibility maps are similar; however, there are significantly more red color areas in the latter, whereas there are more dark green areas in the former. The black is the main scarp, and the blue line is the almost dissected crown

extremely high) according to their LSI values (Table 4) using the six-factor BPNN. For the fourteen-factor susceptibility map, Fig. 11 shows that 85.93 % of the total landslides occurred in 62.84 % of the area classified as high (high, very high, and extremely high) based on Table 5. The results reveal that there were more areas classified as the susceptible zones using fourteen factors rather than six factors. The landslide susceptibility map in both the

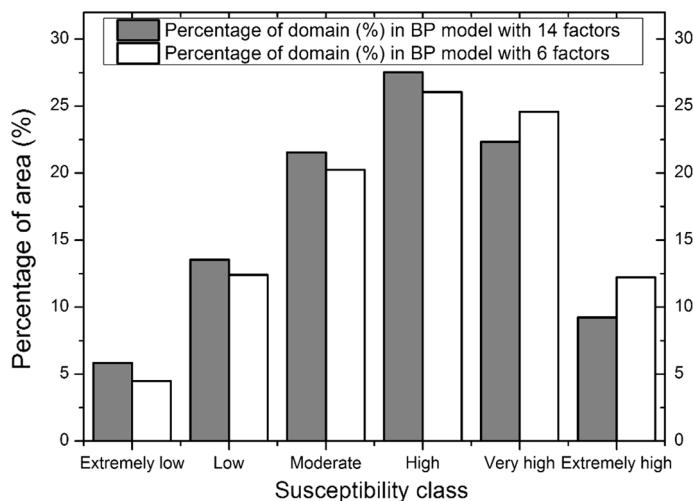


Fig. 9 Comparison with the relative distribution of susceptibility classes using the six and fourteen factors, respectively

cases reveals well matching between the forecasted susceptibility class and the absence or presence of landslides in each mapping unit. Furthermore, these observations indicate that the map using six factors can separate the different susceptibility areas more clearly and hence is considered to be better than the map using the fourteen factors.

4.3 The validation of the landslide susceptibility models

It is critical to verify the accuracy of any prediction model. Without the validation, the predicted model and LSM maps lack the necessary scientific significance (Chung and Fabbri 1993). For the verification, the total landslides were divided into two groups: training data (512 landslide locations) and validation data (220 landslide locations). We verified the accuracy of the BPNN model using a receiver operating characteristic (ROC) curve.

A ROC plot of sensitivity (true-positive rate is the portion of false-positives out of the total actual positives) and 1-specificity (false-positive rate is the portion of false-positives out of the total actual negatives) was made for the model validation (Swets 1988). The area under the ROC curve (AUC) can characterize the quality of a forecasting system by describing the system's ability to correctly predict the occurrence or non-occurrence of a predefined event (Swets 1988). The AUC value varies from 0.5 to 1.0. An ideal model has an AUC value of 1.0. For the proposed model, the AUC value (0.82) was higher using six factors than 0.73 using fourteen factors (Fig. 12). Moreover, the AUC is significantly different from 0.5 because the p value is 0.00, implying that the BPNN classified the group significantly better than by chance. Thus, it is concluded that the optimized selected factors can produce higher model of accuracy than all the collected factors because the latter can suffer from data redundancy or multi-collinearity.

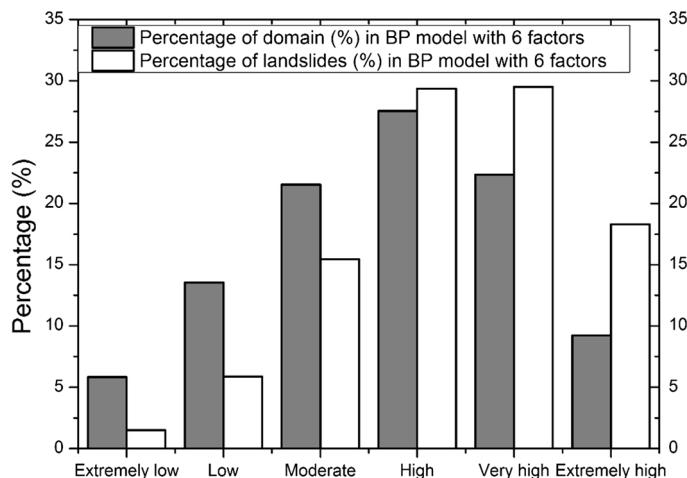


Fig. 10 Distribution of the study area and the occurrence of landslides according to the classification scheme used for LSM with six factors

Table 4 Statistical analysis result of landslide susceptibility in BPNN with six factors

Susceptibility class	Area of each class	Percentage of domain (%)	No. of landslides	Percentage of landslides (%)
Extremely low	205,502	5.82	11	1.50
Low	477,698	13.53	43	5.87
Moderate	760,493	21.54	113	15.44
High	972,034	27.53	215	29.37
Very high	788,560	22.34	216	29.51
Extremely high	325,975	9.23	134	18.31

5 Discussion

The predication of the precise locations of the instabilities and susceptibility assessment is rather difficult owing to the uncertainty of the spatial and temporal distribution of the rainstorms. The susceptibility assessment may also be influenced by other important factors such as geology, slope aspect, and slope angle. This issue is commonly addressed by GIS-based landslide susceptibility studies. Different researchers utilize various landslide-conditioning factors to produce the landslide susceptibility mapping. There is no critical standard or universal rule to select the conditioning factors (Ayalew and Yamagishi 2005). Selecting factors is the fundamental step for susceptibility evaluation and influences the result of LSM. To address this problem, we propose CF method to select the factors. After process, six factors were selected, such as slope angle, slope aspect, distance to fault, and lithology. All these conditioning factors are regarded as high relationship with the landslide occurrence. The results of AUC also prove our assumption. The optimized factors have higher accuracy than all possible factors (improve 8 %) due to the spatial autocorrelation and data redundancy among the conditioning factors.

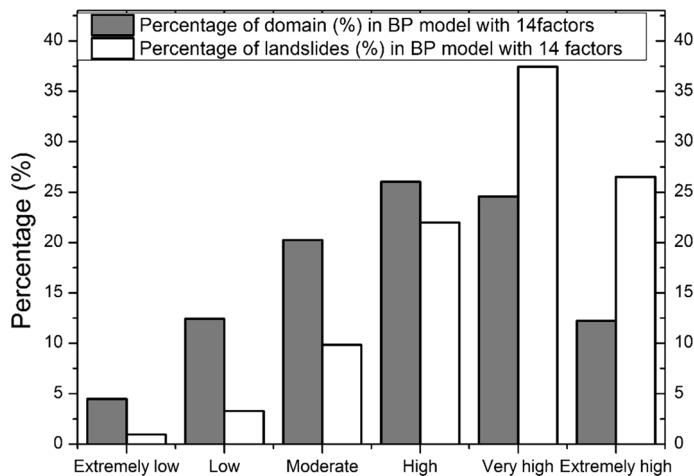


Fig. 11 Distribution of the study area and the occurrence of landslides according to the classification scheme used for LSM with fourteen factors

Table 5 Statistical analysis result of landslide susceptibility in BPNN with fourteen factors

Susceptibility class	Area of each class	Percentage of domain (%)	No. of landslides	Percentage of landslides (%)
Extremely low	106,324	4.49	7	0.96
Low	294,359	12.42	24	3.28
Moderate	480,058	20.25	72	9.84
High	617,344	26.04	161	21.99
Very high	582,619	24.58	274	37.43
Extremely high	289,692	12.22	194	26.50

In this study, we also draw the direction of landslide. The slope aspect from DEM and direction of landslides were compared with each other as displayed in Fig. 13. We can observe that landslides mostly occur at SE, SW, and S from slope aspect direction (Fig. 13a) and at SW, NE, S, and SE from direction of landslide (Fig. 13b), respectively. The results of Fig. 12 were slightly different; however, both of them suggest a similar conclusion, for instance, landslides frequently occurred at SW and S direction. Because the western part of the ridge (NNE-SSW) of Osado is characterized by less snow cover due to strong western wind, however, much more snow existed in the eastern side of the ridge, and snow melted quickly in the southeast area. This may increase the soil moisture and aggravate the slope instability. Therefore, the likely occurrence of the landslide direction is relatively higher in the southeast part. Additionally, vegetation data are available for the study area. Yet, the vegetation cover was not considered as conditioning factors due to the type of triggered landslides—deep-seated—because the root of vegetation was very relatively shallow. Hence, they do not influence the landslide occurrence much.

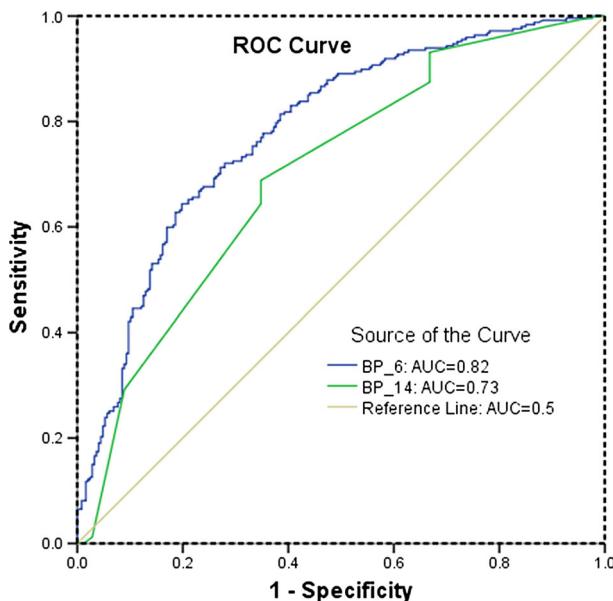


Fig. 12 ROC curves for landslide susceptibility maps produced using the BPNN with six and fourteen factors in this study

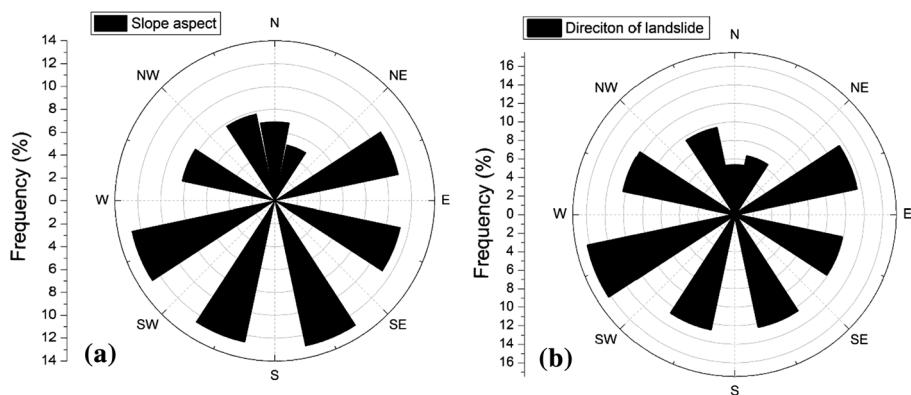


Fig. 13 Frequency of slope aspect (a) and direction of landslides (b)

6 Conclusions

In this research, the landslide susceptibility models with the hybrid CF and artificial neural network-BPNN models were applied to Osado in Sado Island, Japan. Employing a certainty factor method, this study selected six significant factors including slope angle, slope aspect, drainage density network, distance from the geologic boundary, distance to fault, and lithology from fourteen available landslide-conditioning factors. Identifying the major conditioning factors is a fundamental requirement for further study of landslide

susceptibility and prediction. We believe that research differs from the others and provides a method to select and qualify the landslide-conditioning factors.

The selected factors were then input to the BPNN model. The application of the BPNN model was partitioned into two phases: the training network and the validation phase for the landslide susceptibility maps. The results indicate that the BPNN model was successful at identifying landslide susceptible areas because of its flexibility, power, and ease of use. The derived landslide susceptibility maps highlight that a substantial percentage of the areas fell into high susceptibility classes (from high to extremely high), more than 46.77 % using six factors, and 63.78 % using fourteen factors. The models were verified using AUC values obtained from ROC plots. The AUC value for the model with six factors (0.82) is slightly higher than that with fourteen factors (0.73). Moreover, the map using the six factors more clearly indicates separate areas with different landslide susceptibility levels. These results indicate that inputting more factors into a model does not necessarily produce a better result because of the redundancy and multi-collinearity among the factors. The proposed method can provide a method to select the effective factors.

Furthermore, the susceptibility maps represent areas that are prone to landslides and provide the baseline information for further assessment of landslide susceptibility and the related risk. Thus, the prepared LSMs are of great importance for controlling damage and sustainable urban development in the Island.

Acknowledgments We would like to express our gratitude to Midori NET Niigata and Sado City for providing the ortho photographs of Sado Island and the NIED for providing the landslide data.

References

- Aleotti P, Chowdhury R (1999) Landslide hazard assessment: summary review and new perspectives. *Bull Eng Geol Environ* 58:21–44. doi:[10.1007/s100640050066](https://doi.org/10.1007/s100640050066)
- Arora MK, Das Gupta AS, Gupta RP (2004) An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. *Int J Remote Sens* 25:559–572. doi:[10.1080/014311603100156819](https://doi.org/10.1080/014311603100156819)
- 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. doi:[10.1016/j.geomorph.2004.06.010](https://doi.org/10.1016/j.geomorph.2004.06.010)
- Ayalew L, Yamagishi H, Marui H, Kanno T (2005) Landslides in Sado Island of Japan: part I. Case studies, monitoring techniques and environmental considerations. *Eng Geol* 81:419–431
- Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. *Hydrol Sci J* 24:43–69
- Binaghi E, Luzi L, Madella P et al (1998) Slope instability zonation: a comparison between certainty factor and fuzzy Dempster–Shafer approaches. *Nat Hazards* 17:77–97. doi:[10.1023/A:1008001724538](https://doi.org/10.1023/A:1008001724538)
- Brabb EE (1984) Innovative approaches to landslide hazard mapping. *Proc 4th Int Symp Landslides* 1:307–324
- Chang T-C, Chao R-J (2006) Application of back-propagation networks in debris flow prediction. *Eng Geol* 85:270–280. doi:[10.1016/j.enggeo.2006.02.007](https://doi.org/10.1016/j.enggeo.2006.02.007)
- Chauhan S, Sharma M, Arora MK, Gupta NK (2010) Landslide susceptibility zonation through ratings derived from artificial neural network. *Int J Appl Earth Obs Geoinf* 12:340–350. doi:[10.1016/j.jag.2010.04.006](https://doi.org/10.1016/j.jag.2010.04.006)
- Chung C-J, Fabbri AG (1993) The representation of geoscience information for data integration. *Nonrenewable Resour* 2:122–139. doi:[10.1007/BF02272809](https://doi.org/10.1007/BF02272809)
- Conforti M, Pascale S, Robustelli G, Sdao F (2014) Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy). *Catena* 113:236–250. doi:[10.1016/j.catena.2013.08.006](https://doi.org/10.1016/j.catena.2013.08.006)
- Dou J, Qian J, Zhang H et al (2009) Landslides detection: a case study in Conghua city of Pearl River delta. *Second Int Conf Earth Obs Glob Chang*. doi:[10.11117/12.836328](https://doi.org/10.11117/12.836328)

- Pradhan B, Lee S, Buchroithner MF (2010) A GIS-based back-propagation neural network model and its cross-application and validation for landslide susceptibility analyses. *Comput Environ Urban Syst* 34:216–235. doi:[10.1016/j.compenvurbsys.2009.12.004](https://doi.org/10.1016/j.compenvurbsys.2009.12.004)
- Prasad R, Pandey A, Singh KP et al (2012) Retrieval of spinach crop parameters by microwave remote sensing with back propagation artificial neural networks: a comparison of different transfer functions. *Adv Sp Res* 50:363–370. doi:[10.1016/j.asr.2012.04.010](https://doi.org/10.1016/j.asr.2012.04.010)
- Rumelhart D, Hinton G, Williams R (1986) Learning representations by back-propagating errors. *Nature* 323:533–536
- Shi H-Y, Lee K-T, Lee H-H et al (2012) Comparison of artificial neural network and logistic regression models for predicting in-hospital mortality after primary liver cancer surgery. *PLoS ONE* 7:e35781. doi:[10.1371/journal.pone.0035781](https://doi.org/10.1371/journal.pone.0035781)
- Shortliffe EH (1976) Computer-based medical consultations: MYCIN. North-Holland, New York
- Song JH, Venkatesh SS, Conant EA et al (2005) Comparative analysis of logistic regression and artificial neural network for computer-aided diagnosis of breast masses. *Acad Radiol* 12:487–495. doi:[10.1016/j.acra.2004.12.016](https://doi.org/10.1016/j.acra.2004.12.016)
- Sumer E, Turker M (2013) An adaptive fuzzy-genetic algorithm approach for building detection using high-resolution satellite images. *Comput Environ Urban Syst* 39:48–62. doi:[10.1016/j.compenvurbsys.2013.01.004](https://doi.org/10.1016/j.compenvurbsys.2013.01.004)
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293
- Tien Bui D, Pradhan B, Lofman O et al (2012) Landslide susceptibility assessment in the Hoa Binh province of Vietnam: a comparison of the Levenberg–Marquardt and Bayesian regularized neural networks. *Geomorphology* 171–172:12–29. doi:[10.1016/j.geomorph.2012.04.023](https://doi.org/10.1016/j.geomorph.2012.04.023)
- Turner AK, Schuster R (1996) Landslides: investigation and mitigation. Transp. Res. Board, Washington, DC
- Vahidnia MH, Alesheikh AA, Alimohammadi A, Hosseinali F (2010) A GIS-based neuro-fuzzy procedure for integrating knowledge and data in landslide susceptibility mapping. *Comput Geosci* 36:1101–1114. doi:[10.1016/j.cageo.2010.04.004](https://doi.org/10.1016/j.cageo.2010.04.004)
- Van Westen CJ, Rengers N, Terlien MTJ, Soeters R (1997) Prediction of the occurrence of slope instability phenomena through GIS-based hazard zonation. *Geol Rundschau* 86:404–414. doi:[10.1007/s005310050149](https://doi.org/10.1007/s005310050149)
- Yamagishi H (2008) GIS mapping of landscape and disasters of Sado Island, Japan. *Int Arch Photogramm Remote Sens Spat Inf Sci* XXXVII:1429–1432 (Part B7, Beijing 2008)
- 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* 79:251–266. doi:[10.1016/j.enggeo.2005.02.002](https://doi.org/10.1016/j.enggeo.2005.02.002)
- Yilmaz I (2009) Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). *Comput Geosci* 35:1125–1138. doi:[10.1016/j.cageo.2008.08.007](https://doi.org/10.1016/j.cageo.2008.08.007)
- Youssef AM, Pradhan B, Pourghasemi HR, Abdullahi S (2015) Landslide susceptibility assessment at Wadi Jawrah Basin, Jizan region, Saudi Arabia using two bivariate models in GIS. *Geosci J* doi:[10.1007/s12303-014-0065-z](https://doi.org/10.1007/s12303-014-0065-z)
- Zare M, Pourghasemi HR, Vafakhah M, Pradhan B (2013) Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arab J Geosci* 6:2873–2888. doi:[10.1007/s12517-012-0610-x](https://doi.org/10.1007/s12517-012-0610-x)