Article
Vol. 23, No. 2, p. 341–355, April 2019
https://doi.org/10.1007/s12303-018-0038-8
pISSN 1226-4806 eISSN 1598-7477
Geosciences Journal GJ

Comparison of landslide susceptibility maps using random forest and multivariate adaptive regression spline models in combination with catchment map units

Lei Chu[1,2], Liang-Jie Wang[1,2]*, Jiang Jiang[1,2], Xia Liu[1,2], Kazuhide Sawada[3], and Jinchi Zhang[1,2]

[1]Collaborative Innovation Center of Sustainable Forestry in Southern China of Jiangsu Province, Nanjing Forestry University, Nanjing 210037, China
[2]Jiangsu Provincial Key Lab of Soil Erosion and Ecological Restoration, Nanjing Forestry University, Nanjing 210037, China
[3]Graduate School of Engineering, Gifu University, Gifu 5011193, Japan

ABSTRACT: Landslide susceptibility mapping (LSM) is a critical tool for mitigating the damages caused by geologic disasters. The selection of map units and mathematical models greatly affects the efficiency of LSM. To obtain the most appropriate combination of map units and mathematical models, four scales of catchment map units (CMUs) were analyzed and random forest (RF) and multivariate adaptive regression spline (MARSpline) models were applied in Gero City, Japan. The percentage of correctly identified landslides and the areas under the relative operating characteristic (ROC) curve were used to evaluate the model performances. The results indicate that the RF model had higher prediction accuracy than the MARSpline model, especially when the size of the CMU was 0.09 km^2. A relatively high percentage of landslides fell into the high and very high landslide susceptibility classes (73%) and the lowest percentage of landslides fell into the very low landslide susceptibility classes (0.82%). The prediction-area (P-A) plots indicated that the prediction rates were higher for the RF model than the MARSpline model. The results of this study also suggest that the model accuracy can be increased if the appropriate CMU size is used. Therefore, the potential benefits of using the RF model in combination with the appropriate CMU size should be further explored using additional landslide-conditioning factors and other models.

Key words: landslide susceptibility, multivariate adaptive regression spline (MARSpline) model, random forest (RF) model, catchment map units (CMUs)

Manuscript received October 26, 2017; Manuscript accepted July 1, 2018

1. INTRODUCTION

Landslides are one of the worst natural disasters, capable of causing massive amounts of human casualties and economic losses (Nefeslioglu et al., 2008). A tool for mapping and tracking landslides could help local governments to mitigate the associated losses. Landslide susceptibility mapping (LSM) technology has been widely applied for the identification of areas susceptible to landslides because of its accessibility and visualization capabilities in a geographic information system (GIS) (Scaioni, 2013; Shahabi and Hashim, 2015). Various quantitative and statistical methods have been used in landslide susceptibility studies. These include the frequency ratio (FR) (Yalcin et al., 2011; Ozdemir and Altural, 2015; Youssef et al., 2015), logistic regression (LR) (Bai et al., 2010; Wang et al., 2013; Althuwaynee et al., 2014; Shahabi et al., 2014), and bivariate statistical (BS) analysis (Nandi and Shakoor, 2010; Xu et al., 2012). Due to improvements in data mining methods, state-of-the-art techniques such as artificial neural networks (ANN) (Bui et al., 2012; Zaren et al., 2012; Conforti et al., 2014), weights of evidence (WOE) (Regmi et al., 2010a, b), fuzzy logic (Lee, 2007; Kanungo et al., 2008; Muthu et al., 2008; Pradhan et al., 2010), neuro-fuzzy systems (Oh and Pradhan, 2011; Sezer et al., 2011; Pradhan et al., 2014), support vector machines (SVM) (Yao et al., 2008; Yilmaz, 2010; Tanner, 2014; Su et al., 2015), and decision trees (DT) (Saito et al., 2009; Yeon et al., 2010; Wang et al., 2016) have been used increasingly to map landslide susceptibility (Su et al., 2015).

In recent studies, researchers have compared different

*Corresponding author:
Liang-Jie Wang
Collaborative Innovation Center of Sustainable Forestry in Southern China of Jiangsu Province, Nanjing Forestry University, Nanjing 210037, China
Tel: +86-25-85428936, E-mail: ljwang0811@gmail.com; pawangwlj@outlook.com

quantitative methods of LSM. Four different statistical methods, including LR, BS, ANN, and SVM with three different kernel functions were compared for LSM for the landslide caused by the 2008 Wenchuan earthquake in China (Xu et al., 2012). Also, a comparative study of the FR, LR, and WOE models was presented by Ozdemir and Altural (2013) for a landslide susceptibility analysis in a mountainous area. Zoglu et al. (2014) compared multi-criteria decision analyses and SVM with LR and found that these approaches generally outperformed the conventional LR method for mapping shallow landslides. Youssef et al. (2015) proposed an ensemble method of FR and LR and compared these approaches to the original LR and FR methods, whereas Chen et al. (2015) compared the FR with a statistical index and an index of entropy models. All these studies were based on the concept that future landslides are likely to occur under the same geologic, geomorphologic, and hydrologic conditions that triggered a landslide event in the past. However, these studies were focused on grid-based (or cell-based) calculations for LSM. The main shortcoming of grid-based calculations is the lack of considering the heterogeneity of the geological and land use factors. The grid-based methods partition the polygon units of the geological and land use factors into raster cells for further calculations.

Although some research on catchment-based (or polygon-based) LSM has been reported in the literature (Guzzetti et al., 2005; Galli et al., 2008; Saito et al., 2009), these studies did not consider the different sizes of the catchment-based (polygon-based) map units (CMUs). The selection of the mapping unit is crucial because it determines how landslides are sampled to obtain the training and prediction (validation) subsets for susceptibility modeling (Hussin et al., 2016) and it affects the accuracy of LSM at the regional scale.

In this study, we used four sizes of CMUs and analyzed the data using a random forest (RF) algorithm for LSM. The RF method has rarely been used in the assessment of landslide susceptibility and has not been compared to the multivariate adaptive regression spline (MARSpline) method in previous studies. The RF and MARSpline models were used to create landslide susceptibility maps in Gero City, Gifu Prefecture, Japan and their performances were compared. The specific objectives were to (1) select the appropriate scale of the CMUs for LSM by considering the heterogeneity of the geological and land use parameters and (2) map the spatial distribution of the landslide susceptibility index using the RF and MARSpline models.

2. MATERIALS AND METHOD

2.1. Study Area

The study area was chosen due to the lack of sufficient landslide-conditioning factors and field observation data of landslides. The study area covers approximately 851 km^2 and is located in the eastern part of Gifu Prefecture, Japan (Fig. 1). The elevation ranges from 220 m to 3052 m. The landforms are hilly with an average slope angle of 24.5°. The lithology in the study area consists primarily of sedimentary and volcanic rocks. The most common land use/land cover classes are plantations and secondary forest. Local government records indicated that the main causes of landslides were prolonged rainfall or snowmelt.

2.2. Datasets

2.2.1. Landslide inventory

A landslide inventory map from 2008 was acquired from the National Research Institute for Earth Science and Disaster Prevention (NIED) (http://www.bosai.go.jp/) to understand the spatial distributions of the landslides in the study area. According to the definition by Varnes (1978), a landslide is defined as mass movements including soil slides, debris slides, rock slides, and debris flows (Fig. 1).

The size of the landslides ranges from 2.5×10^{-3} km^2 to 0.9 km^2. The landslide area percentage (LAP), which is defined as the percentage of the area affected by landslide activity, is 5.48%. The landslide number density (LND) is 0.74 landslides/km^2, which is the number of landslides per sq km.

Different shapes of landslides were chosen to represent the locations where mass movements occurred. Vahidnia et al. (2010) and Yeon et al. (2010) used partial landslide points (grid cells) to define the training dataset and used some of these points to test the dataset. Sterlacchini et al. (2010) and Yalcin et al. (2011) used polygons for the training and test datasets for the landslide areas. Although some researchers prefer to use points to represent landslide locations, others have used polygons to determine the locations of landslides on a regional scale. In this study, the CMUs were derived from an area using appropriately defined terrain mapping units (TMUs) that combined information on the geology and land use/cover. Four different scales of CMUs were used to represent the presence of landslides. It was also necessary to obtain sampling data from areas without landslides for the RF model. For this purpose, four groups of datasets, including the training and testing CMU datasets are used in this study.

2.2.2. Spatial database development

Topographic, hydrologic, lithologic, and land cover datasets were obtained for the study area. The landslide-conditioning factors were extracted from the datasets and stored in a spatial database using a spatial analysis tool (ArcGIS10.1 software package).

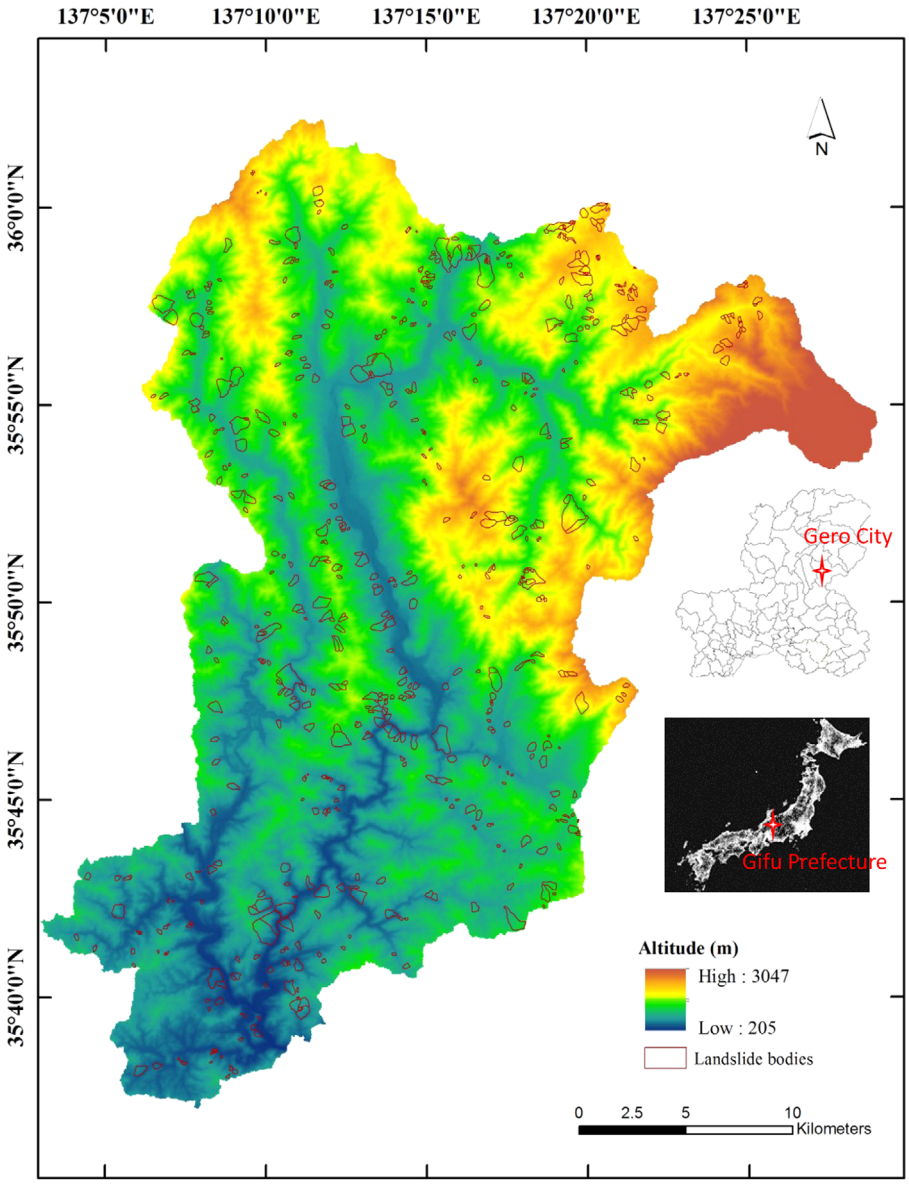

Fig. 1. Geographical location and distribution of landslide bodies in Gero City, Gifu Prefecture, Japan.

Topographic indices: The Advanced Space-borne Thermal Emission and Reflection Radiometer Global Digital Elevation Model Version 2 (ASTER GDEM V2) (http://glcfapp.glcf.umd.edu/data/aster/) was used to determine the topographic characteristics of Gero City in 2013. The spatial resolution of the GDEM is 20 m. We also applied airborne LiDAR data, which were acquired in 2011 to calibrate the ASTER GDEM and to obtain higher resolution DEM data.

The topographic indices of altitude (Fig. 2a), percent slope (Fig. 2b), and relief (Fig. 2c) were derived using the ArcGIS 10.1 software package.

Water-related index: The topographic wetness index (TWI) was derived from the DEM (Fig. 2d); it is used to estimate soil moisture and the erosive power of water flow (Regmi et al., 2010a, b).

$$TWI = log_e(A/btan\,\alpha), \quad (1)$$

where A (m^2) is the upstream area, b (m) is the resolution of the DEM, and α is the slope gradient.

Land use/Land cover: A land use/land cover map is important for predicting landslide occurrences. This data layer (from 2013) was acquired from the Biodiversity Center of Japan (http://www.biodic.go.jp) (Fig. 2e).

Geological parameters: The geological map representing the features of the slope materials was provided by the Geological Survey of Japan, AIST (https://www.gsj.jp/). The lithology map was created in 1992 and was digitized using ArcGIS 10.1 (Fig. 2f). Note that the geology of our study area is very complex and

https://doi.org/10.1007/s12303-018-0038-8

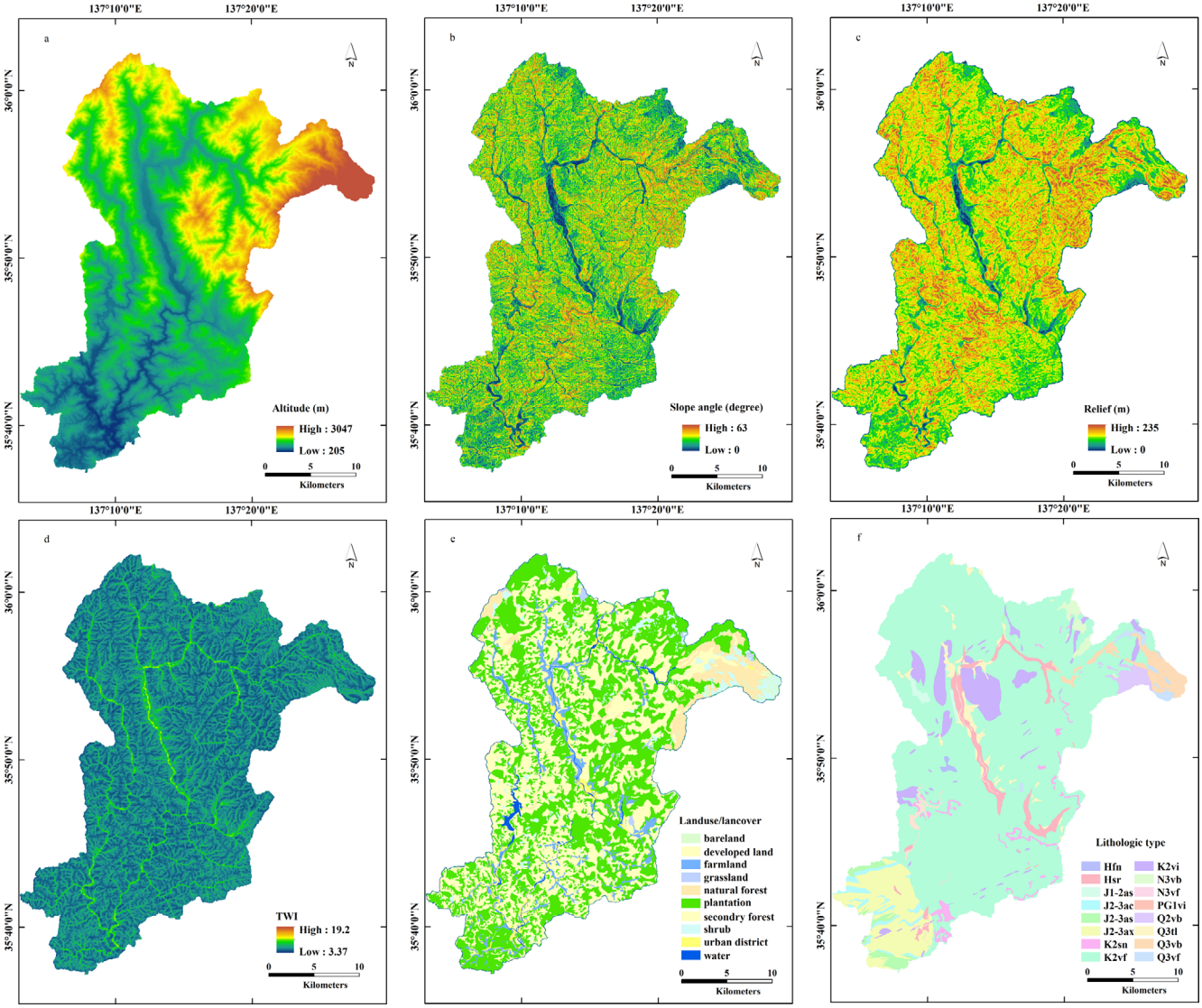

Fig. 2. Landslide conditioning factors: (a) altitude, (b) slope percent, (c) relief, (d) TWI, (e) land use/land cover, and (f) lithology.

Table 1. Lithologic components of the study area

Code	Lithology	Geological age
Hfn	Fan deposits	Late Pleistocene to Holocene
Hsr	Marine and non-marine sediments	Late Pleistocene to Holocene
J1-2as	Sandstone of accretionary complex	Early to Middle Jurassic
J2-3ac	Chert block of J2-3 accretionary complex	Triassic to Middle Jurassic
J2-3as	Sandstone of J2-3 accretionary complex	Middle to Late Jurassic
J2-3ax	Melange matrix of J2-3 accretionary complex	Middle to Late Jurassic
K2sn	Non-marine sediments	Late Cretaceous
K2vf	Non-alkaline felsic volcanic rocks	Late Cretaceous
K2vi	Non-alkaline felsic volcanic intrusive rocks	Late Cretaceous
N3vb	Non-alkaline mafic volcanic rocks	Late Miocene to Pliocene
N3vf	Non-alkaline felsic volcanic rocks	Late Miocene to Pliocene
PG1vi	Non-alkaline felsic volcanic intrusive rocks	Paleocene to Early Eocene
Q2vb	Non-alkaline mafic volcanic rocks	Middle Pleistocene
Q3tl	Lower terrace	Late Pleistocene
Q3vb	Non-alkaline mafic volcanic rocks	Late Pleistocene
Q3vf	non-alkaline felsic volcanic rocks	Late Pleistocene

the lithological units comprise several formations. The formations were therefore classified into sixteen categories with respect to landslide susceptibility. The descriptions of the geological units are provided in Table 1.

The landslide density defined as "previous knowledge" or "expert knowledge" was used to transform the categorical data into numerical data (Bui et al., 2010). The landslide density (LD) is defined by the following formula:

$$LD = \frac{D_i/A_i}{\sum_i^N D_i/A_i}, \quad (2)$$

where D_i is the area of a landslide of the i-th category, A_i is the area of the i-th category for a certain parameter, and N is the number of the parameter. The lithology and land use types are converted to numerical variables (Table 2).

2.2.3. Descriptive statistics of CMUs and conditioning factors

To select the appropriate CMUs, the study area was divided into four different map unit scales, including 0.01 km^2 (5 cells × 5 cells), 0.02 km^2 (7 cells × 7 cells), 0.04 km^2 (10 cells × 10 cells), and 0.09 km^2 (15 cells × 15 cells). The numbers of CMUs for these scales are 27234, 16163, 7835, and 3147, respectively. Six landslide-conditioning factors, including altitude, slope angle, relief, TWI, land cover, and geological type, are used in the LSM. The standard deviation (STD), maximum (Ma), minimum (Mi), and average (Av) of the altitude, slope angle, relief, and TWI of the grid cells in each CMU were calculated using spatial analysis tools (Table 3).

Table 2. Landslide density values of lithologic components and land use/land cover

Lithologic type	Landslide density	Land cover/use type	Landslide density
Hfn	0.055	Bareland	0.068
Hsr	0.042	Developed land	0.102
J1-2as	0.180	Farmland	0.070
J2-3ac	0.045	Grassland	0.119
J2-3as	0.053	Natural forest	0.059
J2-3ax	0.040	Plantation	0.187
K2sn	0.099	Secondry forestry	0.188
K2vf	0.079	Shrub	0.135
K2vi	0.076	Urban district	0.072
N3vb	0.086		
N3vf	0.000		
PG1vi	0.131		
Q2vb	0.002		
Q3tl	0.062		
Q3vb	0.010		
Q3vf	0.040		

Table 3. CMU characteristics and definitions

Characteristics		Definitions
LS-yes	Landslides	Occurred
LS-no		No occurred
Geo	Geological data	Landslide density of Geologic
LC	Landcover/landuse	Landslide density of landcover
Ele_{ma}	Altitude	Maximum height (m)
Ele_{mi}		Minimum height (m)
Ele_{av}		Average height (m)
Ele_{std}		Standard deviation (m)
$Slope_{ma}$	Slope angle	Maximum (degree)
$Slope_{mi}$		Minimum (degree)
$Slope_{av}$		Average (degree)
$Slope_{std}$		Standard deviation (degree)
$Relief_{ma}$	Relief	Maximum (m)
$Relief_{mi}$		Minimum (m)
$Relief_{av}$		Average (m)
$Relief_{std}$		Standard deviation (m)
TWI_{ma}	TWI	Maximum
TWI_{mi}		Minimum
TWI_{av}		Average
TWI_{std}		Standard deviation

2.2.4. Training and test datasets

The manner in which a training dataset is chosen can affect the LSM results. Nandi and Shakoor (2010) divided a landslide inventory into two subsets: landslides in the left portion of the study area were used as the training dataset and those in the right portion were used as the testing dataset. Vahidnia et al. (2010) used partial landslide points for the training dataset and used a subset of these points for the test dataset. Pradhan (2013) used 50% of the landslide grid cells for training and the other 50% for testing using random sampling. In this study, if the landslides fell entirely within the map units, those units were defined as an occurrence of a landslide. Subsequently, the numbers of landslide map units were 6932, 4913, 3140, and 1774 for the four different map unit sizes of 0.01 km^2, 0.02 km^2, 0.04 km^2, and 0.09 km^2. Seventy percent of the map units were used for training and the other 30% were used for model testing; the sampling was conducted randomly.

2.3. Model Development

2.3.1. RF model

The RF model is a relatively new, tree-based model that has been developed to optimize the predictive performance by combining a large number of simple DTs into a powerful model rather than using a single-tree model based on traditional classification and regression trees (Breiman, 2001; Skurichina and Duin, 2001; Yang et al., 2016).

https://doi.org/10.1007/s12303-018-0038-8

To run the RF model, three a priori parameters are defined, including the number of variables used to grow each tree (mtry), the number of trees in the forest (ntree), and the minimum number of terminal nodes (nodesize). Breiman (2001) and Liaw and Wiener (2002) stated that even one variable/factor (mtry = 1) could generate a good level of accuracy, whereas Gromping (2009) proved the need to include at least two variables/factors (i.e., mtry = 2, 3, 4, ……, m) in order to avoid using the weaker regressors as splitters (Breiman, 2001; Liaw and Wiener, 2002; Gromping, 2009; Rahmati et al., 2016). During the training procedure, each tree is generated by bootstrap samples, leaving approximately one-third of the overall sample for validation and using the out-of-bag (OOB) error (Gromping, 2009). Detailed information of the RF model can be found in Breiman (2001) and Liaw and Wiener (2002).

The aim of the RF model is to analyze the relationship between the independent variables and a dependent variable to determine the weighted value for each factor. To fit an RF model, the default values for mtry (one-third of the total number of predictors), nodesize (5) and ntree (1000) were used.

2.3.2. Multivariate adaptive regression spline (MARSpline)

Multivariate adaptive regression spline (MARSpline) is a technique that combines classical linear regression, mathematical construction of splines, and binary recursive partitioning to produce a local model where the relationships between the response and the predictors are either linear or non-linear (Felicisimo et al., 2013). The MARSpline model is defined as the sum of weighted basis functions:

$$F(x) = a_0 + \sum_{i=1}^{n} a_i f_i(x), \quad (3)$$

where $f_i(x)$ is the basis function, n is the number of basis functions in the model, and $f_0(x)$ is the constant basis function, the coefficient of which is a_0. All of the coefficients are calculated using ordinary least squares (OLS). The basis functions are represented by the following equation:

$$f_i(x) = \prod_{j=1}^{d_i} [S_{ji}(X_{v(j,i)} - t_{ji})]_+, \quad (4)$$

where d_i is the number of variables (interaction order) in the ith basis function S_{ji}, $X_v(j,i)$ is the vth variable, $1 \le v(j,i) \le d$, and t_{ji} is the knot location for each of the corresponding variables.

2.4. Statistical Analyses

The statistical analyses and modeling were performed using R software (Felicisimo et al., 2013). The RF models were developed using the R Random Forest package (Liaw and Wiener, 2002). The implementation of the MARSplines was carried out with SPM software and the training datasets (Salford Systems, http://www.salfordsystems.com/).

2.5. Model Validation

The accuracy of the landslide susceptibility maps was evaluated by calculating the relative operating characteristic curve (ROC) and the percentage of observed landslide points in the various susceptibility categories (Nandi and Shakoor, 2010). Previous studies demonstrated that the area under the ROC curve (AUC) is suitable for quantifying the uncertainty in model predictions and can account for detection biases associated with the estimation (Guzzetti et al., 2005; Wang et al., 2013; Rahmati et al., 2016). The ability and uncertainty of the RF and MARSpline models were investigated using the AUC. An AUC value close to 1 indicates a high level of accuracy, whereas an AUC value close to 0.5 indicates inaccuracy (Wang et al., 2013). In order to assess the relative efficiency of the results, the prediction-area (P-A) plot can be used to evaluate the ability of the different-sized CMUs to predict the landslide susceptibility (Yousefi and Carranza, 2015a, b, c). In this study, the P-A plot was used to determine the relative importance of the different-sized CMUs. More detailed information on the use of the P-A plot can be found in Yousefi and Carranza (2015a, b, c), Yousefi and Carranza (2016, 2017), and Yousefi and Nykänen (2016).

3. RESULTS AND DISCUSSIONS

3.1. Model Performance

The OOB error, which is an unbiased procedure, was used to select the optimum parameters of the RF algorithm. The OOB error is a function of the number of trees and is, therefore, lower when more trees are added to the algorithm. As a result, when the value of the OOB error changed slowly, appropriate values of the OOB error and the ntree were obtained. Table 4 shows the predictive performance of the RF models with different values of the OOB error and ntree. As shown in Table 4, the accuracies of the training data decreased from 81.3% to 64.4% as the scale of the map units increased to 0.09 km^2 (RF4). In contrast, the

Table 4. Statistics and accuracies of RF models

	OOB	ntree	Accuracy
RF_1	0.204	50	81.3%
RF_2	0.225	50	79.2%
RF_3	0.263	51	76.3%
RF_4	0.360	27	64.4%

Table 5. Accuracies of MARSpline models

	Accuracy
$MARSplines_1$	79.2%
$MARSplines_2$	75.6%
$MARSplines_3$	75.2%
$MARSplines_4$	68.9%

values of the OOB error and the ntree decreased when the size of the map units increased. These results indicate that a small tree number and OOB error were observed when the size of the map units was large and that the scale of the map unit affects the accuracy of the LSM using the RF algorithm.

Table 5 shows the predictive performance of the MARSpline model with different CMUs. The accuracy of the training data increased from 68.9% to 79.2% as the scale of the map units decreased. When the map units increased to 0.09 km^2 (MARSpline4), the accuracy dropped dramatically. By applying the MARSpline algorithm, the following models based on the training dataset and composed of the basis functions with different CMUs in the study area were obtained:

MARSpline1:
BF3 = max(0, TWIma - 5.515);
BF6 = max(0, 0.0764 - Geo);
BF7 = max(0, Reliefma - 69);
BF11 = max(0, Eleav - 613.1);
BF12 = max(0, 613.1- Eleav);
$$Y = 0.3553 + 0.0855 \times BF3 - 7.0138 \times BF6 + 0.003945 \times BF7 + 0.00069 \times BF11 - 0.0011 \times BF12, \quad (5)$$

MARSpline2:
BF3 = max(0, TWIma - 5.880);
BF6 = max(0, 108.14 - Reliefav);
BF8 = max(0, 0.0764 - Geo);
BF9 = max(0, Elema - 667);
BF13 = max(0, Elema - 926);
$$Y = 0.4371 + 0.0867 \times BF3 - 0.0039 \times BF6 - 7.192 \times BF8 + 0.00185 \times BF9 - 0.0016 \times BF13, \quad (6)$$

MARSpline3:
BF2 = max(0, 1237 - Elema);
BF4 = max(0, 9.576 - TWIma);
BF5 = max(0, Slopemi - 3.338);
BF7 = max(0, Reliefstd - 3.0834);
$$Y = 1.1344 - 0.0011 \times BF2 - 0.1316 \times BF4 + 0.0178 \times BF5 + 0.0081 \times BF7, \quad (7)$$

MARSpline4:
BF2 = max(0, 1170 - Elema);
BF4 = max(0, 9.81006 - TWIma);
BF5 = max(0, Slopeav - 31.7212);
BF12 = max(0, 20.57 - Reliefstd);
$$Y = 0.9958 - 0.00074 \times BF2 - 0.1212 \times BF4 - 0.1646 \times BF5 - 0.013 \times BF12. \quad (8)$$

3.2. Variable Importance

The relative importance of the predictors was determined from the training datasets of the RF models and is shown in Figure 3. The relative importance of the predictors differed between the four models. The important variables were normalized to 100% in the RF models to provide a simple basis for comparison. The five most important predictors in all of the models were Ele_{ma}, Ele_{av}, Ele_{mi}, and Geo, with the exception of the RF_4 model, where Ele_{std} had the fourth highest value. This finding implies that these environmental variables are the primary indicators of

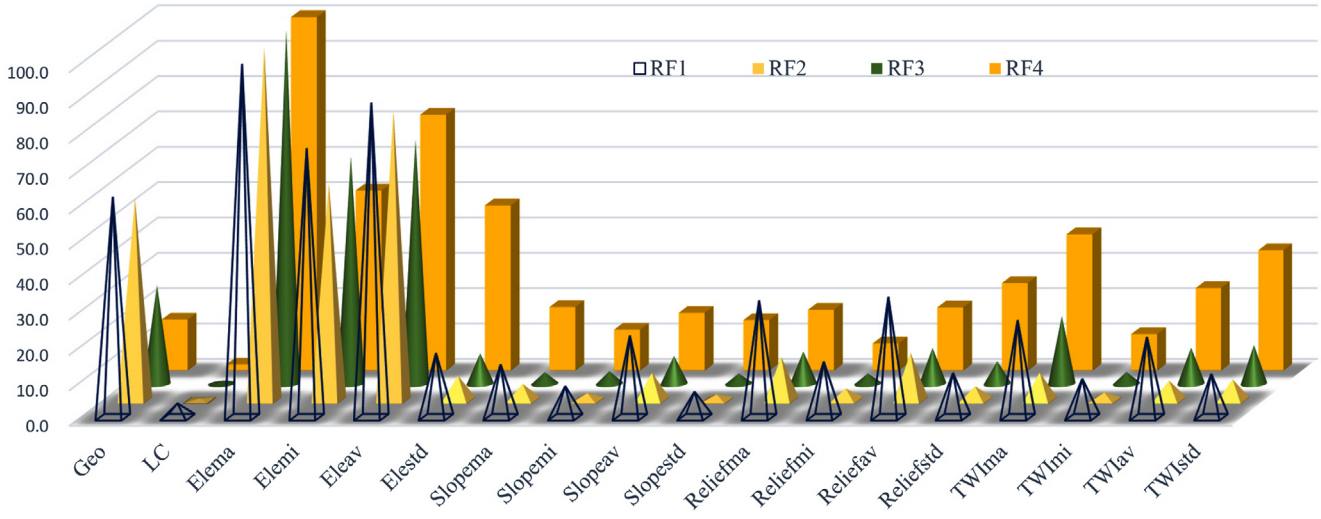

Fig. 3. Relative importance of each variable using random forest (RF) models with different CMUs, which are normalized to 100%.

https://doi.org/10.1007/s12303-018-0038-8

the spatial variability of the landslide susceptibility in Gero City when using the RF algorithm. Figure 3 shows the effects of the landslide-conditioning factors for the different-sized CMUs using the RF model. Ele_{ma}, Ele_{av}, and Ele_{mi} have the most important influences on the landslide occurrence for the different CMUs. For the CMU size of 0.09 km^2, the Ele_{ma} value played a key role in the model performance. $Relief_{ma}$ and $Relief_{av}$ are very important factors for modeling landslide occurrence when the small CMUs are used for training the model. TWI_{ma} and TWI_{std} are more important when using the larger sized CMUs (0.04 km^2 (10 cells × 10 cells) and 0.09 km^2 (15 cells × 15 cells)). Land use/land cover is not a key factor for LSM in this study area using the RF method. However, this finding differs from those of several previous studies (Bui et al., 2005; Nandi and Shakoor, 2010; Wang et al., 2013), indicating that the map unit (cell-based or catchment-based map unit) can be an important factor in landslide susceptibility assessments.

Figure 4 shows the effect of the landslide-conditioning factors for the different training datasets for the MARSpline model. TWI_{ma} has an important influence on landslide occurrence for the different CMUs. For the larger CMUs, Ele_{ma} plays a key role in the model performance. The lithology of the formation is very important for predicting landslide occurrence when the smaller CMUs are used for training the model. $Relief_{std}$ is more important when using the larger CMUs.

3.3. Spatial Distribution of Landslide Susceptibility Index

The landslide susceptibility indices were calculated by applying the results of the RF and MARSpline models for the different scales of map units. Figures 5 and 6 show the spatial distributions of the landslide susceptibility indices derived from the RF and MARSpline model respectively for Gero City, Gifu Prefecture, Japan.

3.4. Evaluation of Landslide Susceptibility Maps

The accuracy of the landslide susceptibility maps was evaluated by calculating the AUC for the ROC and the percentage of observed landslides in the various susceptibility categories. The AUC values showed that the RF_1 model had the highest success rate (AUC = 0.88) and the best prediction rate (AUC = 0.89) when used with the smallest map unit (Table 6). As the catchment size increased,

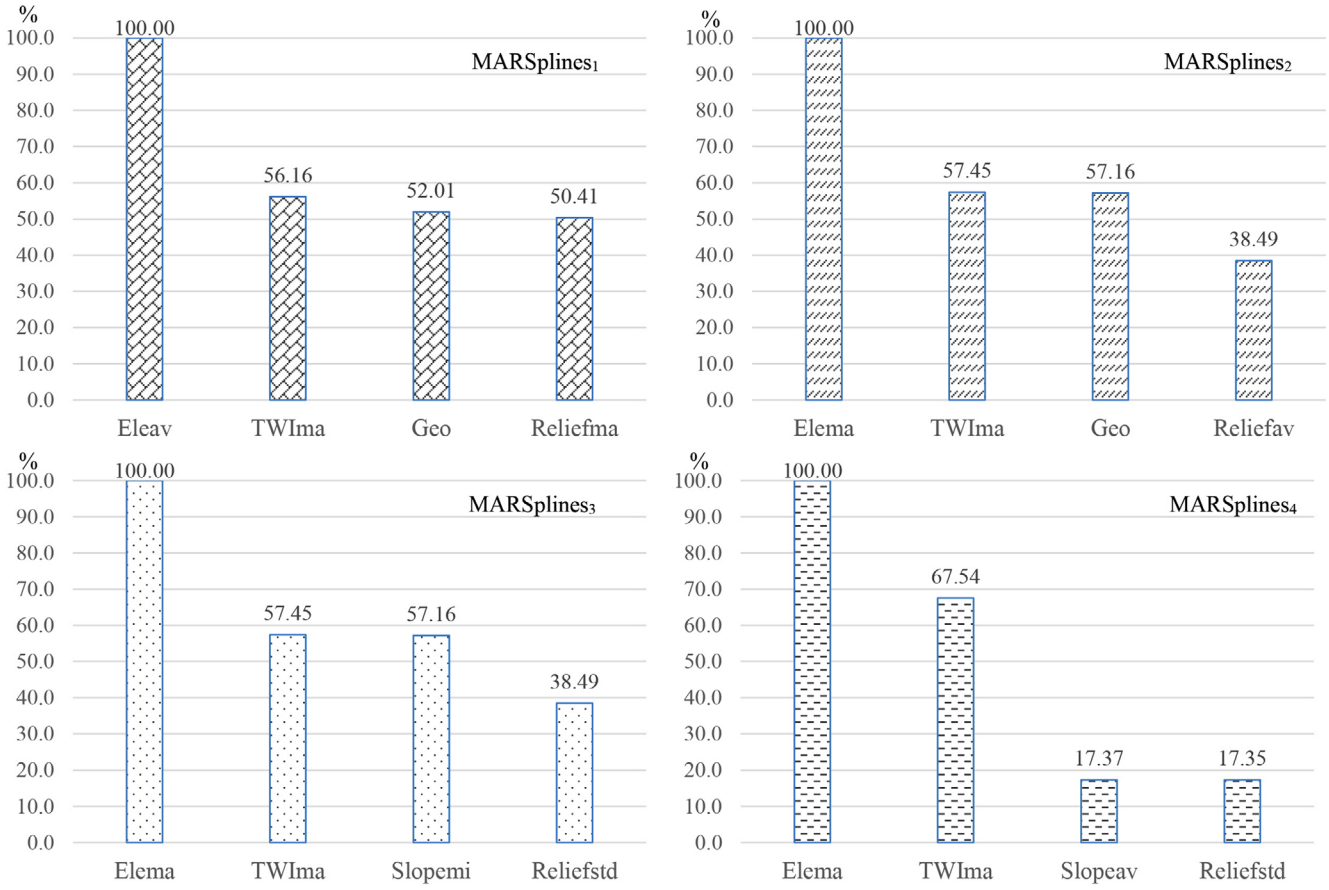

Fig. 4. Parameter importance map obtained from MARSplines model. The horizontal axis is the factor importance.

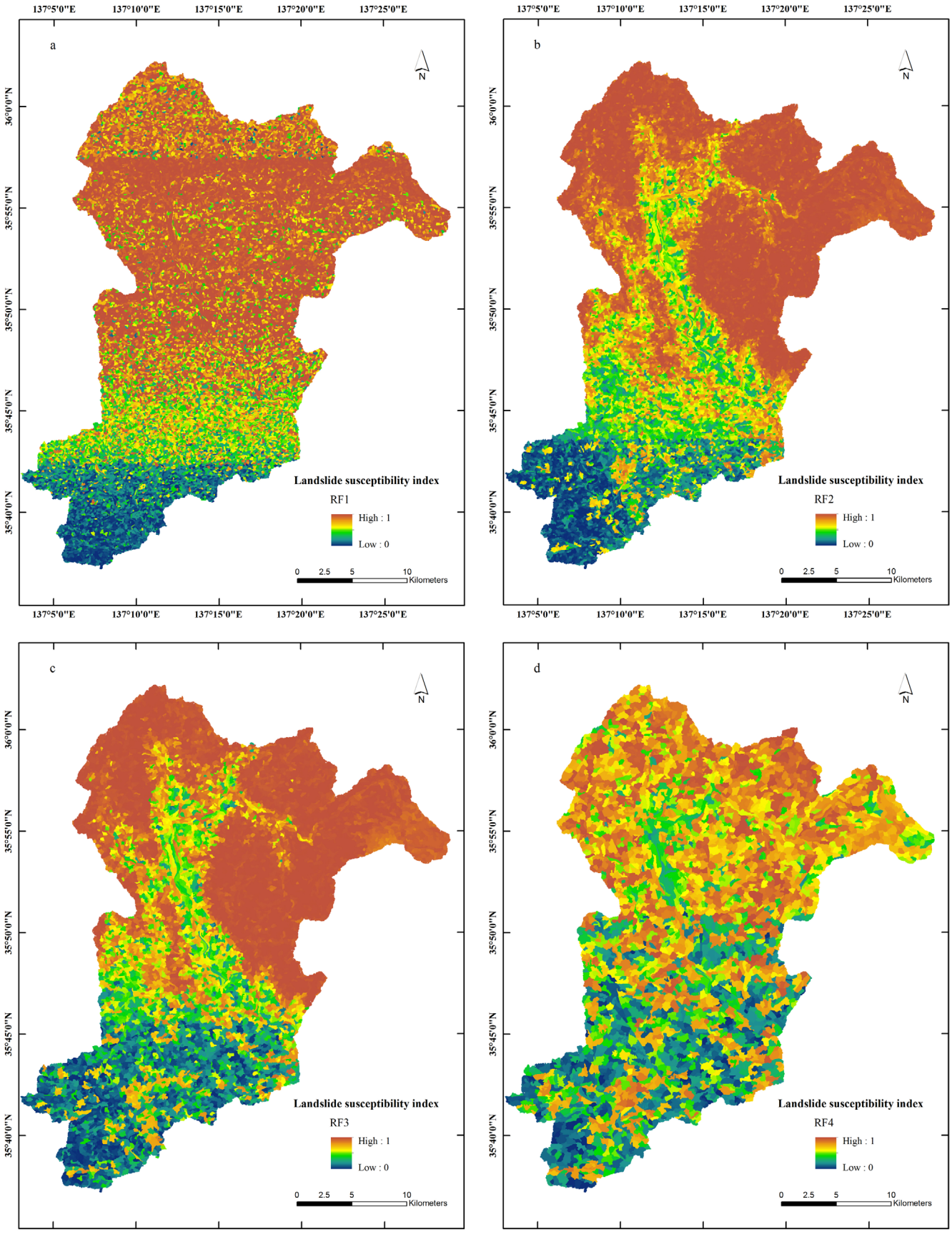

Fig. 5. Landslide susceptibility maps derived from RF models: (a) RF_1, (b) RF_2, (c) RF_3, (d) RF_4.

https://doi.org/10.1007/s12303-018-0038-8

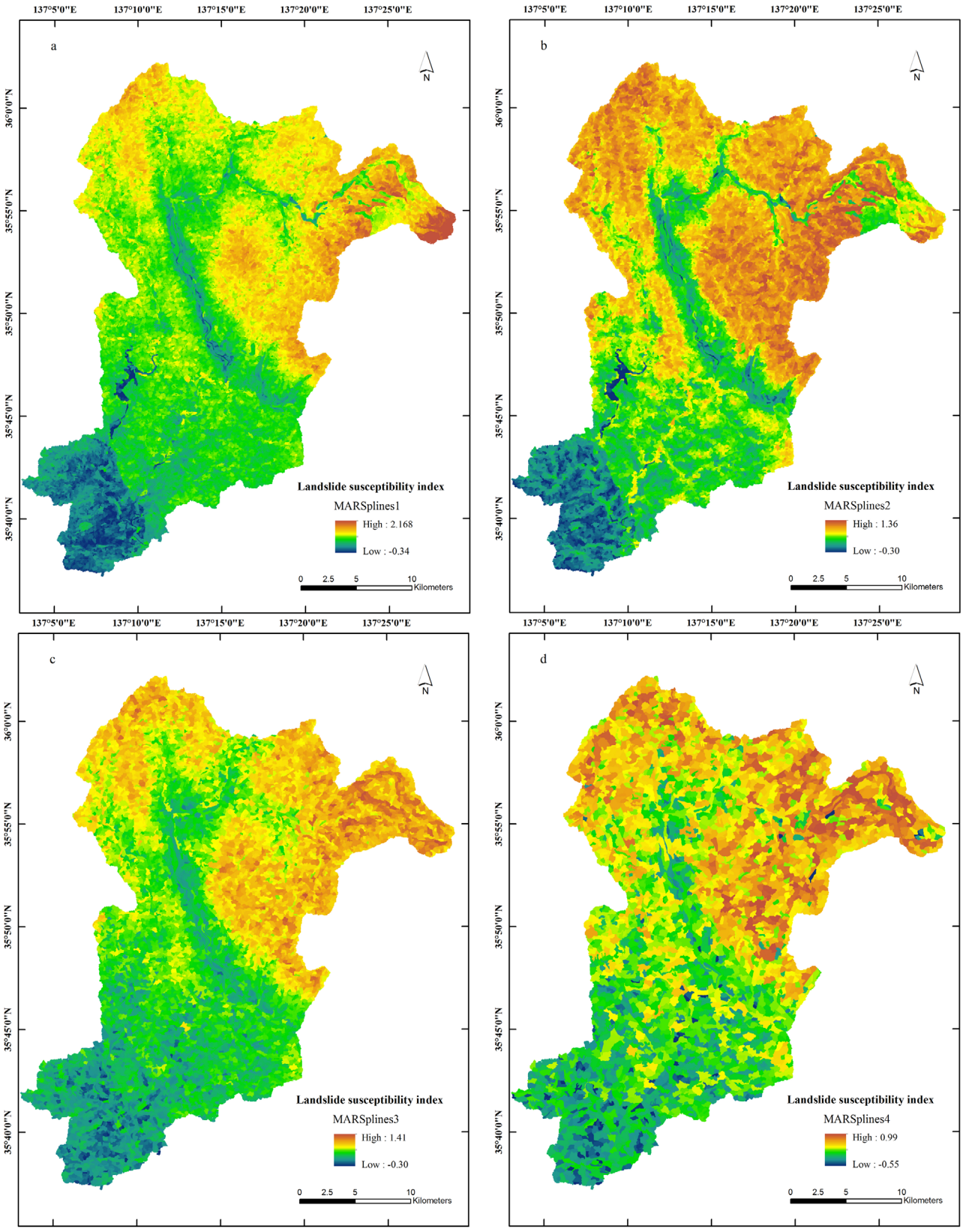

Fig. 6. Landslide susceptibility maps derived from MARSpline models: (a) MARSplines1, (b) MARSplines2, (c) MARSplines3, (d) MARSplines4.

the AUC values of the success rate and prediction rate decreased. The AUC values showed a similar trend for the MARSpline algorithm. The RF model with the smallest CMU resulted in the second highest success and prediction rates (AUC = 0.88 and

Table 6. Success rates and prediction rates of RF and MARSpline models

	Success rate (AUC)	Prediction rate (AUC)
RF_1	0.88	0.89
RF_2	0.87	0.86
RF_3	0.82	0.83
RF_4	0.70	0.73
$MARSplines_1$	0.88	0.87
$MARSplines_2$	0.85	0.84
$MARSplines_3$	0.81	0.82
$MARSplines_4$	0.72	0.74

AUC = 0.87, respectively), which were close to the RF_1 values. The RF model provided higher AUC values than the MARSpline model for the same size of CMUs with the exception of the CMU with a size of 0.09 km^2, where the RF_4 model resulted in success and prediction rates with lower AUC values (AUC = 0.70 and AUC = 0.73, respectively).

In addition, the test landslides were used to assess the accuracy of the landslide susceptibility maps. There are several mathematical methods for classifying the degrees of susceptibility (Wang et al., 2013). In this research, the quantile method was used in the ArcGIS platform. The quantile method equally divides the landslide susceptibility areas into equal area classes (ESRI, http://www.esri.com/). The susceptibility maps were then divided into five susceptibility categories (Fig. 7): very low, low, medium, high, and very high. If the maximum number of landslides possible were designated very high susceptibility zones and the minimum number of landslides possible were designated very low susceptibility zones, this would indicate that the classification system for the susceptibility maps is appropriate. Figure 7 shows the

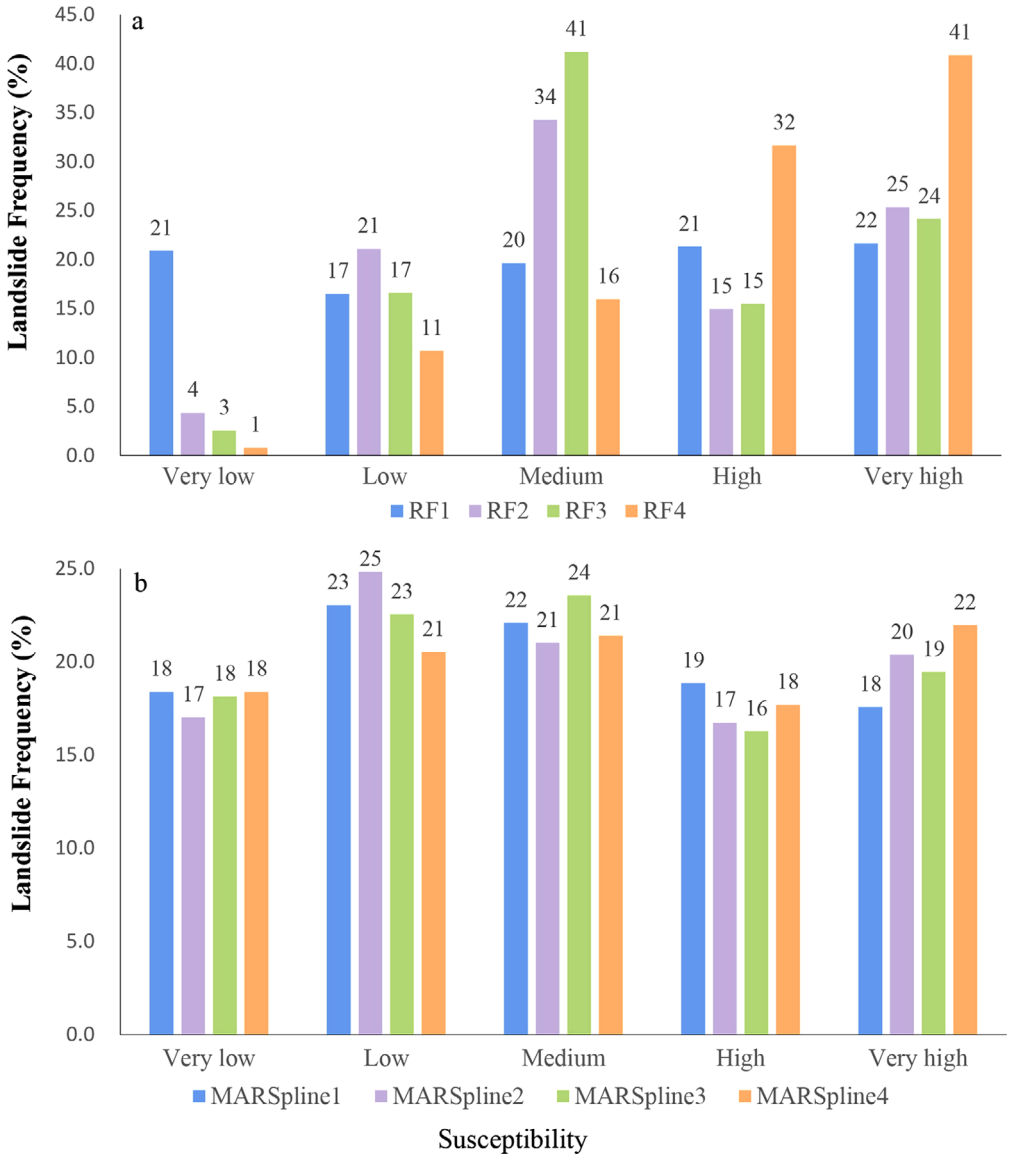

Fig. 7. Percentages of test landslide bodies falling into different susceptibility categories: (a) RF, (b) MARSplines.

https://doi.org/10.1007/s12303-018-0038-8

percentage of the observed landslides in the various susceptibility categories.

As shown in Figure 7a, less than 1% (0.82%) of the observed landslides fell into the very low susceptibility class using the RF method with the largest map unit. Although the RF_1 model result indicates that 21% of the landslides fell into the very low susceptibility class, the other results show that less than 5% of landslides fell into the very low class using the RF algorithms. Over 22%, 25%, 24%, and 41% of the landslides occurred in the areas with very high susceptibility classes based on the RF method with different CMUs. The largest proportion of landslides in the medium to very high susceptibility classes was 88% based on the RF_4 model. The RF_1 model had the lowest value of 63%. Nearly 74% of the landslides fell into the high and very high classes according to the RF_4 model, which indicates that this CMU (0.09 km^2) was suitable for assessing the landslide susceptibility. A comparison of the results of the RF_1, RF_2, and RF_3 models indicates that the larger CMUs provided better results.

Figure 7b shows that the high and very high categories only contain 37%, 37%, 35%, and 40% of the landslides using the MARSpline method. It is also not apparent if the percentage of the landslides in the very high category is higher than the percentage of the landslides in the other categories in all the models. The percentages of each class are very similar. For the MARSpline algorithm, the best model results are obtained for the largest CMU.

A comparison of the results from the RF and MARSpline analyses for the same size of CMU shows that the RF model provides slightly better results. A lower percentage of landslides fell in the very low and low classes and a greater percentage of landslides fell into the very high class, especially when the size of the CMU was 0.09 km^2. Over 41% of the landslides fell into the very high class. These results indicate that the RF model is relatively successful in determining the landslide susceptibility in the study area.

The P-A plots for the different CMUs are shown in Figure 8. The intersection point of the P-A plots for the RF4 is higher than the intersection points of the P-A plots for the other models. Based on the intersection points in Figure 8h, 70.1% landslide occurrences are delineated in 38.1% of the study area according to RF4 model. It is the highest percentage landslide occurrences felling the lowest percentage of study area, which indicates that the RF4 model is relatively successful for landslide susceptibility mapping in the study area. The ranking of the prediction rates of the RF and MARSpline models from highest to lowest is (1) RF_4, (2) RF_3, (3) RF_2, (4) RF_1, (5) MARSpline4, (6) MARSpline3 (7) MARSpline2 and (8) MARSpline1. These results were consistent with the results shown in the previous sections. The RF models had higher prediction rates than the MARSplines algorithms. the larger CMUs provided better results.

4. CONCLUSIONS

In this study, the RF and MARSpline models were used for the assessment of landslide susceptibility at four scales of CMUs and for the creation of landslide susceptibility maps that can be used by local authorities for land use and environmental protection plans. The RF4 model was the best model with the highest predictive ability for the study area. A susceptibility level at which an active landslide will occur is assumed in landslide susceptibility analyses. If only areas with very high landslide susceptibility were at risk, the RF method would offer the best results. If landslides were found in areas with high levels of susceptibility, the RF model would also yield better results because a higher percentage of landslides were observed in the area of very high susceptibility (40%).

The results obtained from the RF and MARSpline models show good prediction abilities with AUC values higher than 0.7. Although the AUC values were lower for the larger CMUs, the differences in the results were relatively small. It should also be noted that the performance of the landslide susceptibility maps depends not only on the mathematical model but also on the CMUs. Therefore, if the appropriate CMUs are selected, the performance of the RF and MARSpline models may improve.

Furthermore, the main disadvantages of data-driven methods such as RF and MARSpline models are system bias and uncertainty. Although the RF model performed better than the MARSpline model in this study area, the latter model may perform well in other study areas. It would be of interest to compare the RF model with other methods, e.g., ANN, BS, and WOE. In many cases, it is not appropriate to conclude that a particular method performs better because it is more powerful (Yousefi and Carranza, 2015c) since the performance of landslide susceptibility maps depends not only on the mathematical model but also on the selected datasets and the conditioning factors.

To some extent, the landslide susceptibility maps were not able to provide more detailed information on the landslide-prone areas at the medium scale for Gero City but the maps can provide the local authorities with the spatial distributions of potential future landslides. Additionally, the factors relating with the landslide mechanisms are very complex in this area. In this study, we used the CMUs to create landslide susceptibility maps because of a lack of conditioning factors. In future studies, more detailed information and high-resolution datasets should be used to create landslide susceptibility maps. The catchment scale-based landslide susceptibly maps provide only basic information for quantifying the risk of landslides.

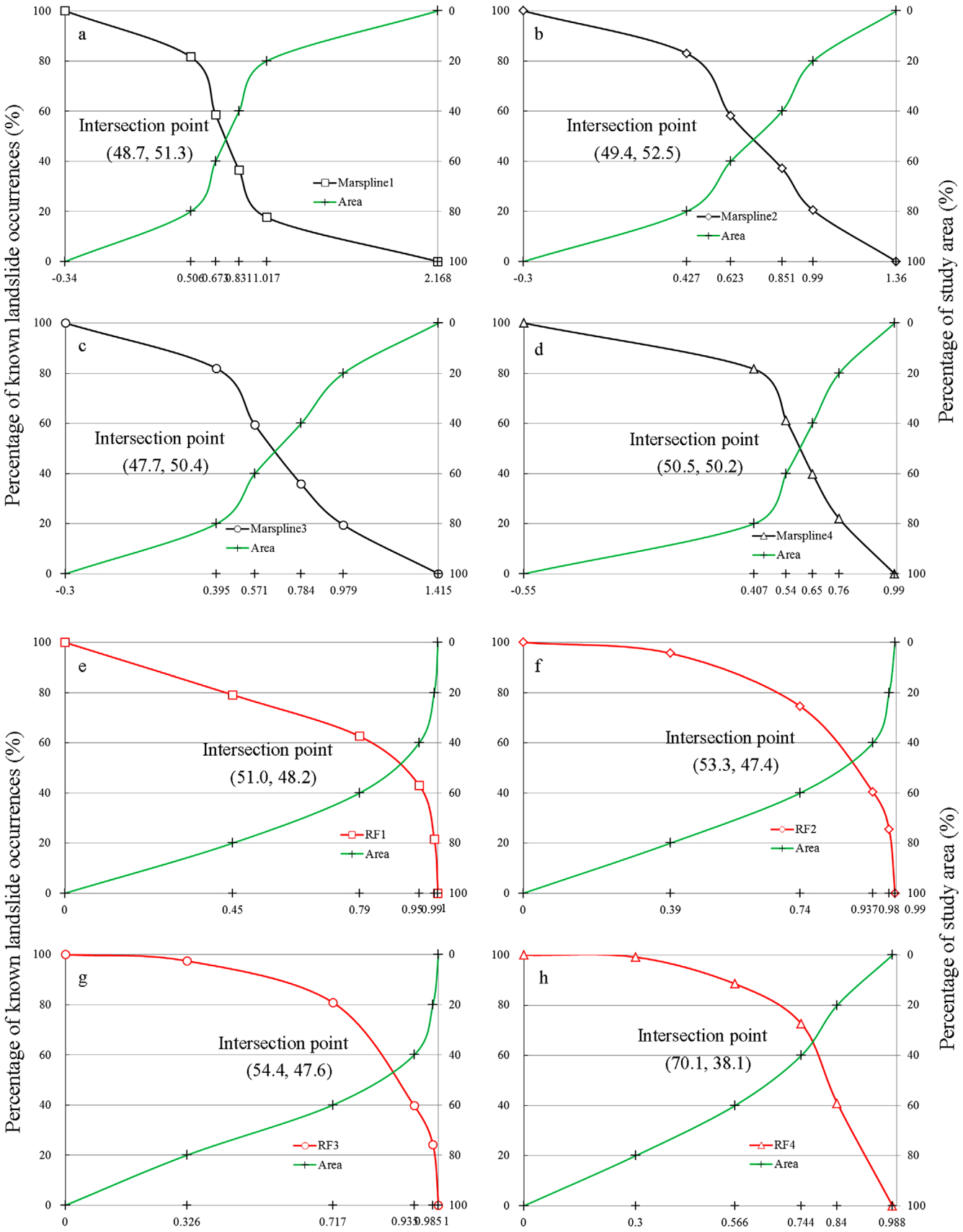

Fig. 8. P-A plot for the different-sized CMUs for the RF and MARSpline models with intersection point values (percentage of known landslide occurrences, percentage of study area): (a) MARSpline1, (b) MARSpline2, (c) MARSpline3, (d) MARSpline4, (e) RF_1, (f) RF_2, (g) RF_3, (h) RF_4.

https://doi.org/10.1007/s12303-018-0038-8

ACKNOWLEDGMENTS

This research is supported by National Natural Science Foundation of China (Grant No.41601209), Natural Science Foundation for colleges and universities in Jiangsu Province (Grant No.16KJB220004), Agricultural science and technology innovation foundation of Jiangsu Province (Grant No.CX(17)1004), High Academic Talent Foundation Nanjing Forestry University (Grant No.GXL2014037), Priority Academic Program Development of Jiangsu High Education Institutions (PAPD). Data were supported by Gifu University and National Research Institute for Earth Science and Disaster Prevention of Japan.

REFERENCES

Althuwaynee, O.F., Pradhan, B., Park, H.J., and Lee, J.H., 2014, A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping. Cantena, 114, 21–36.

Bai, S.B., Wang, J., Guo, N.L., Zhou, P.G., Hou, S.S., and Xu, S.N., 2010, GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorgesarea, China. Geomorphology, 115, 23–31.

Breiman, L., 2001, Random forests. Machine Learning, 45, 5–32.

Bui, D.T., Pradhan, B., Lofman, O., Revhaug, I., and Dick, O.B., 2012, Landslide susceptibility assessment in the Hoa Binh province of Vietnam using artificial neural network. Geomorphology, 171–172, 12–29.

Chen, W., Li, W., Hou, E., Bai, H., Chai, H., Wang, D., Cui, X., and Wang, Q., 2015, Application of frequency ratio, statistical index, and index of entropy models and their comparison in landslide susceptibility mapping for the Baozhong Region of Baoji, China. Arabian Journal of Geosciences, 8, 1829–1841.

Conforti, M., Pascale, S., Robustelli, G., and 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.

Felicisimo, A., Cuartero, A., Remondo, J., and Quiros, E., 2013, Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides, 10, 175–189.

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., and Reichenbach, P., 2008, Comparing landslide inventory maps. Geomorphology, 94, 268–289.

Gromping, U., 2009, Variable importance assessment in regression: linear regression versus random forest. American Statistical Association, 63, 308–319.

Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M., and Ardizzone, F., 2005, Probabilistic landslide hazard assessment at the basin scale. Geomorphology, 72, 272–299.

Hussin, H.Y., Zumpano, V., Reichenbach, P., Sterlacchini, S., Micub, M., Westen, C., and Bălteanu, D., 2016, Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. Geomorphology, 253,508–523.

Kanungo, D.P., Arora, M.K., Gupta, R.P., and Sarkar, S., 2008, Landslide risk assessment using concepts of danger pixels and fuzzy set theory in Darjeeling Himalayas. Landslides, 5, 407–416.

Kavzoglu, T., Sahin, K.E., and Colkesen, I., 2014, Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. Landslides, 11, 425–439.

Lee, S., 2007, Application and verification of fuzzy algebraic operators to landslide susceptibility mapping. Environmental Geology, 52, 615–623.

Liaw, A. and Wiener, M., 2002, Classification and regression by random forest. R News, 2, 18–22.

Muthu, K., Petrou, M., Tarantino, C., and Blonda, P., 2008, Landslide possibility mapping using fuzzy approaches. IEEE Transactions on Geoscience and Remote Sensing, 46, 1253–1265.

Nandi, A. and Shakoor, A., 2010, A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses. Engineering Geology, 110, 11–20.

Nefeslioglu, H., Gokceoglu, C., and Sonmez, H., 2008, An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. Engineering Geology, 97, 171–191.

Oh, H.J. and Pradhan, B., 2011, Application of a neuro-fuzzy model to landslide susceptibility mapping in a tropical hilly area. Computers & Geosciences, 37, 1264–1276.

Ozdemir, A. and Altural, T., 2013, A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. Journal of Asian Earth Sciences, 64, 180–197.

Pradhan, B., 2010, Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches. Journal of the Indian Society of Remote Sensing, 38, 301–320.

Pradhan, B., 2013, A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. Computers & Geosciences, 51, 350–365.

Rahmati, O., Pourghasemi, R.H., and Melesse, M.A., 2016, Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: A case study at Mehran Region, Iran. Catena, 137, 360–372.

R Development Core Team, 2009, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, 409 p.

Regmi, N.R., Giardino, J.R., and Vitek, J.D., 2010a, Assessing susceptibility to landslide: using models to understand observed changes in slopes. Geomorphology, 122, 25–38.

Regmi, N.R., Giardino, J.R., and Vitek, J.D., 2010b, Modeling susceptibility to landslides using the weight of evidence approach: western Colorado, USA. Geomorphology, 115, 172–187.

Saito, H., Nakayama, D., and Matsuyama, H., 2009, Comparison of landslide susceptibility based on a decision tree model and actual landslide occurrence: The Akaishi Mountains, Japan. Geomorphology, 109, 108–121.

Scaioni, M., 2013, Remote sensing for landslide investigations: from

research into practice. Remote Sensing, 5, 5488–5492.
Sezer, E., Pradhan, B., and Gokceoglu, C., 2011, Manifestation of an adaptive neuro-fuzzy model on landslide susceptibility mapping: Klang valley, Malaysia. Expert Systems with Applications, 38, 8208–8219.
Shahabi, H. and Hashim, M., 2015, Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment. Scientific Reports, 5, 9899. https://doi.org/10.1038/srep09899
Shahabi, H., Hhezri, S., Ahmhad, B.B., and Hashim, M., 2014, Landslide susceptibility mapping at central Zab basin, Iran: a comparison between analytical hierarchy process, frequency ratio and logistic regression models. Catena, 115, 55–70.
Skurichina, M. and Duin, R.P., 2001, Bagging, boosting and the random subspace method for linear classifiers. Pattern Analysis and Applications, 5, 121–135.
Sterlacchini, S., Ballabio, C., Blahut, J., Masetti, M., and Sorichetta, A., 2010, Spatial agreement of predicted patterns in landslide susceptibility maps. Geomorphology, 125, 51–61.
Su, C., Wang, L., Wang, X., Huang, Z., and Zhang, X., 2015, Mapping of rainfall-induced landslide susceptibility in Wencheng, China, using support vector machine. Nature Hazards, 76, 1759–1779.
TannerSan, B., 2014, An evaluation of SVM using polygon-based random sampling in landslide susceptibility mapping: the Candir catchment area (western Antalya, Turkey). International Journal of Applied Earth Observation and Geoinformation, 26, 399–412.
Vahidnia, M.H., Alesheikh, A.A., Alimohammadi, A., and Hosseinali, F., 2010, A GIS-based neuro-fuzzy procedure for integrating knowledge and data in landslide susceptibility mapping. Computers & Geosciences, 36, 1101–1114.
Varnes, D.J., 1978, Slope movement types and processes. Transportation Research Board Special Report, 176, 11–33.
Wang, L.J., Guo, M., Sawada, K., Lin, J., and Zhang, J., 2016, A comparative study of landslide susceptibility maps using logistic regression, frequency ratio, decision tree, weights of evidence and artificial neural network. Geosciences Journal, 20, 117–136.
Wang, L.J, Sawada, K., and Moriguchi, S., 2013, Landslide susceptibility analysis with logistic regression model based on FCM sampling strategy. Computers & Geosciences, 57, 81–92.
Xu, C., Xu, X., Dai, F., and Saraf, A., 2012, Comparison of different models for susceptibility mapping of earthquake triggered landslides related with the 2008 Wenchuan earthquake in China. Computers & Geosciences, 46, 317–329.
Yalcin, A., Reis, S., Aydinoglu, A.C., and Yomralioglu, T., 2011, A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. Catena, 85, 274–287.
Yang, R.M., Zhang, G.L., Liu, F., Lu, Y.Y., Yang, F., Yang, F., Yang, M., Zhao, Y.G., and Li, D.C., 2016, Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. Ecological Indicators, 60, 870–878.
Yao, X., Tham, L.G., and Dai, F.C., 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.-K., Han, J.-G., and Ryu, K.H., 2010, Landslide susceptibility mapping in Injae, Korea, using a decision tree. Engineering Geology, 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. Environmental Earth Sciences, 61, 821–836.
Youssef, A.M., Al-Kathery, M., and Pradhan, B., 2015, Landslide susceptibility mapping at Al-Hasher Area, Jizan (Saudi Arabia) using GIS-based frequency ratio and index of entropy models. Geosciences Journal, 19, 113–134.
Youssef, A.M., Pradhan, B., Jebur, M.N., and El-Harbi, H.M., 2015, Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. Environmental Earth Sciences, 73, 3745–3761.
Yousefi, M. and Carranza, E.J.M., 2015a, Fuzzification of continuous-value spatial evidence for mineral prospectivity mapping. Computers & Geosciences, 74, 97–109.
Yousefi, M. and Carranza, E.J.M., 2015b, Prediction-area (P-A) plot and C-A fractal analysis to classify and evaluate evidential maps for mineral prospectivity modeling. Computers & Geosciences, 79, 69–81.
Yousefi, M. and Carranza, E.J.M., 2015c, Geometric average of spatial evidence data layers: a GIS-based multi-criteria decision-making approach to mineral prospectivity mapping. Computers & Geosciences, 83, 72–79.
Yousefi, M. and Carranza, E.J.M., 2016, Data-driven index overlay and Boolean logic mineral prospectivity modeling in greenfields exploration. Natural Resources Research, 25, 3–18.
Yousefi, M. and Carranza., E.J.M., 2017, Union score and fuzzy logic mineral prospectivity mapping using discretized and continuous spatial evidence values. Journal of African Earth Sciences, 128, 47–60.
Yousefi, M. and Nykänen, V., 2016. Data-driven logistic-based weighting of geochemical and geological evidence layers in mineral prospectivity mapping. Journal of Geochemical Exploration, 164, 94–106.
Zarea, M., Pourghasemi, H.R., Vafakhah, M., and Pradhan, B., 2012, Landslide susceptibility mapping at Vaz watershed (Iran) using an artificial neural network model: a comparison between multi-layer perceptron (MLP) and radial basic function (RBF) algorithms. Arabian Journal of Geosciences, 6, 2873–2888.