



# A Novel Hybrid Model of Rotation Forest Based Functional Trees for Landslide Susceptibility Mapping: A Case Study at Kon Tum Province, Vietnam

Binh Thai Pham<sup>1</sup> , Viet-Tien Nguyen<sup>2</sup>, Van-Liem Ngo<sup>3</sup>, Phan Trong Trinh<sup>2</sup>, Huong Thanh Thi Ngo<sup>4</sup>, and Dieu Tien Bui<sup>5</sup> 

<sup>1</sup> Department of Geotechnical Engineering, University of Transport Technology, 54 Trieu Khuc, Thanh Xuan, Hanoi, Viet Nam  
binhpt@utt.edu.vn

<sup>2</sup> Institute of Geological Sciences, Vietnam Academy of Sciences and Technology, 84 Chua Lang Street, Dong da, Hanoi, Viet Nam

<sup>3</sup> Faculty of Geography, VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam

<sup>4</sup> University of Transport Technology, 54 Trieu Khuc, Thanh Xuan, Hanoi, Vietnam

<sup>5</sup> Geographic Information System Group, Department of Business and IT, University College of Southeast Norway, Gulbringvegen 36, 3800 Bø i Telemark, Norway  
buitiendieu@gmail.com

**Abstract.** In this study, we proposed a novel hybrid model namely Rotation Forest based Functional Trees (RFFT), which is a hybrid intelligent approach of two state of the art machine learning techniques of Functional Trees (FT) classifier and Rotation Forest (RF) ensemble, for landslide susceptibility mapping at the Kon Tum Province, Viet Nam. Landslide affecting factors (slope angle, slope aspect, elevation, valley depth, land use, NDVI, soil type, lithology, distance to geology boundaries, and distance to faults), and 1404 past and current landslide locations have been first collected from the study area for generating training and testing datasets. Secondly, the hybrid model RFFT has been constructed for landslide susceptibility assessment using training dataset. Performance of the proposed RFFT model has been validated by analysis of the Receiver Operating Characteristic (ROC) curve and statistical indexes, and compared with a well-known landslide models namely Support Vector Machines (SVM) and the single FT. Results show that the proposed RFFT model has good performance for landslide susceptibility assessment. It has better predictive capability compared with well-known SVM model and single FT model. Therefore, it can be concluded that the proposed RFFT model should be used as a great alternative method for better landslide susceptibility assessment in landslide prone area.

**Keywords:** Landslide susceptibility assessment · Ensemble techniques · Rotation Forest · Functional Trees · Viet Nam

## 1 Introduction

Landslide susceptibility mapping is known as a first step in landslide hazard mitigation [1]. It helps in identifying the high and very high prone areas so that the suitable land use planning strategies and decisions can be made in advance to protect the stability of the slopes and prevent the occurrences of future landslides. Landslide susceptibility mapping is carried out on the basis of the analysis of spatial relationship between past and current landslides and a set of influencing factors [2, 3]. Geographic Information Systems (GIS), which has been used widely as a basic tool for data manipulation and spatial management [4], is also a helpful tool for producing the landslide susceptibility map.

In recent years, many literatures have shown landslide susceptibility mapping studies using GIS. Many of these studies have applied machine learning methods such as artificial neural networks [5, 6], decision trees [7, 8], support vector machines [9, 10], logistic regression [11, 12]. More recently, new hybrid techniques have been developed and used for landslide susceptibility mapping. Bui et al. [13] developed a novel hybrid model of least-squares support vector machines and differential evolution optimization for landslide spatial prediction. Pham et al. [14] proposed a novel hybrid machine learning model of rotation forest ensemble and fuzzy unordered rules induction algorithm classifier for landslide prediction. In general, the hybrid models often outperform the single models for landslide susceptibility mapping [15, 16].

In this paper, we proposed a novel hybrid model namely Rotation Forest based Functional Trees (RFFT), which is a hybrid intelligent approach of two state of the art machine learning techniques of Functional Trees (FT) classifier and Rotation Forest (RF) ensemble, for landslide susceptibility mapping at the Kon Tum Province, Viet Nam. Another objective of this study is to compare the novel hybrid RFFT model with another well-known landslide model namely Support Vector Machines (SVM), and the single FT model. To validate and compare the models, the Receiver Operating Characteristic (ROC) curve and statistical indexes analysis have been applied in both training and testing datasets.

## 2 Methods Used

### 2.1 Rotation Forest

Rotation Forest (RF), which is one of the popular ensemble techniques, was proposed by Rodriguez et al. [17]. It is known as an efficient technique to improve the weaker classifiers [18]. The RF uses Principal Component Analysis (PCA), which is a multivariate technique to analyze the large multivariate datasets for reducing their dimensionality [19], and to separate the original training datasets into subsets which are then used to train the classifiers [20]. Even though the RF is a relatively new ensemble technique, its application has been in many different fields such as medical [21, 22], remote sensing data classification [23, 24]. In landslide susceptibility modeling, the RF has also been applied in few cases [14, 25, 26].

## 2.2 Functional Trees

Functional Trees (FT), which was proposed by Gama [27], is a combination of a discriminant function and a multivariate decision trees by means of constructive induction [28]. It is considered as a generalization of multivariate trees. The FT combines attributes at leaf nodes, decision nodes, or both nodes and leaves for learning classification trees [28] in which decision nodes are constructed as the tree is growing whereas functional leaves are constructed as the tree is pruning [27]. In prediction, the FT can be used to predict the value of the target variables for unclassified samples. Firstly, the sample traverses the tree from the root node to a leaf in which the set of attributes of the sample is extended at each decision node using the constructor function constructed at the node. Thereafter, the decision test of the node is applied to define the path on which the sample will follow. Finally, the sample is classified as a leaf is reached using either the constructor function constructed at the leaf or the constant related with the leaf [28].

## 2.3 Proposed Rotation Forest Based Functional Trees

Rotation Forest Based Functional Trees (RFFT) is a hybrid approach of the RF ensemble and the FT classifier. In landslide modeling, the RFFT is constructed in three main steps as follows:

**Step 1, Optimization:** In this step, the RF ensemble was used to optimize the input data for modeling. Original training dataset was divided into different sub-training datasets which were then used for classification. Optimal sub-training datasets were determined as the best performance of the RFFT model was obtained. Iterations were applied in this step in which the number of iteration was determined as “16” based on a trial-and-error test to give the best for training the RF ensemble.

**Step 2, Classification:** In this step, the FT classifier was used to classify the classes (landslide and non-landslide) for predicting spatially landslides. Different FT classifiers were generated using different sub-training datasets which were generated by the RF ensemble. To train the FT classifier, the parameters of the algorithm were selected using trial-error process such as the minimum number of instances at which a node is considered for splitting is “15”, and the number of iterations for the RF is “15”.

**Step 3, Combination:** In the final step of modeling, the RF ensemble was used to combine the FT classifiers generated to construct the final RFFT model.

## 2.4 Validation Methods

Receiver Operating Characteristic (ROC) curve which is a useful tool to represent the quality of models [29, 30], and the area under the ROC curve (AUC) were selected to validate the overall performance of the hybrid RFFT model in this study. On the ROC curve graph, the “sensitivity” is shown on the y-axis whereas the “100-specificity” is shown on the x-axis. The highest AUC value of “1” indicates the best or perfect performance of the model whereas higher AUC value shows a better model [31–33].

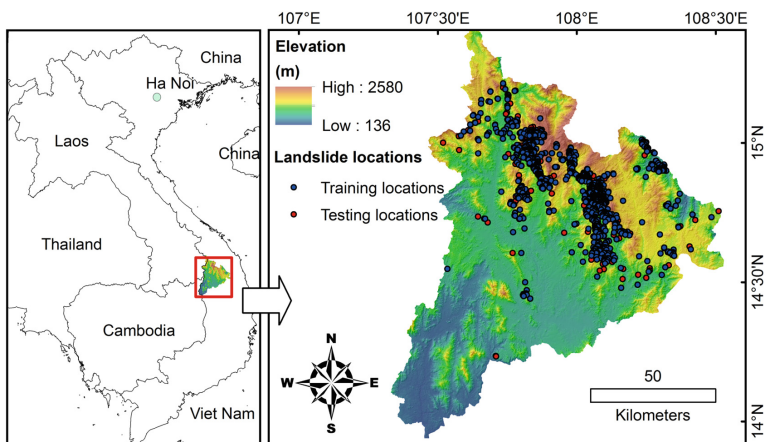
In addition, statistical indexes namely negative predictive value (NPV), positive predictive value (PPV), specificity (SPF), sensitivity (SST), accuracy (ACC), Kappa (K), and Root Mean Squared Error (RMSE), which shows more detail performance of models, were also selected to validate the predictive capability of the proposed hybrid model in the present study. Detail description of these statistical indexes were presented in Bennett et al. [34] and Pham et al. [35].

### 3 General Characteristics of the Study Area

The study area is Kon Tum province located in Central Highland plateau where many geo-hazards (i.e. landslides and human induced earthquakes) occurred as the result of the interaction between natural hazards and human activities [36–40]. Geographically, the study area is located on the western part of Truong Son mountain range, between longitudes 107° 20' E to 108° 32' E and latitudes 13° 55' N to 15° 26' N (Fig. 1). It covers an area of approximately 9689.6 km<sup>2</sup>, and has nine administrative units namely Kon Tum city, Dak Glei, Dak To, Kon Ray, Kon Plong, TuMoRong, Sa Thay, Ngoc Hoi, and Dak Ha districts. The study area belongs to the highly precipitation zone of the Central Highland where total annually rainfall varies from 1700 mm to above 3000 mm in which rainfall (90.9%) concentrates strongly in rainy season (April to November). The Kon Tum area has three main large drainages such as the Dak Bla River, the KrongPoKo River, and the Sa Thay River. The drainage density in the study area is measured around 0.45 km/km<sup>2</sup>.

In the study area, the topography feature is aligned in the N - S direction. The elevation varies from 800–1200 m in the north part whereas it ranges from 500–530 m in the southern part, and the highest point of 2598 m is at the Ngoc Linh peak in the southern part of Vietnam.

The geology is complex with 33 geologic formations and complexes from Archean to Cenozoic of ages. Metamorphic rocks are mostly dominated, followed by intrusive



**Fig. 1.** Landslide inventory map and location map of study region

igneous rocks, extrusive igneous rocks, and sedimentary. The study area belongs to the Kon Tum block which is one of the main uplift blocks in the Central Highland affected strongly by the tectonic activities. Fault zones in the Kon Tum province follow the sub-meridian and northwest-southeast directions. Major faults traverse area including Ngoc Linh - Dak To, Po Co River, Xe Ca Man, Dak Sa, and Sa Thay.

## 4 GIS Database

### 4.1 Landslide Inventory Map

Landslide inventory map of this area was constructed from 1404 landslides occurred during the years from 2004 to 2015, data for the inventory has been collected from various sources such as interpretation of Google Earth images and field investigation.

Landslides in this area can be classified into 3 main types in size such as large landslides (above 10000 m<sup>2</sup>), medium landslides (1000–10,000 m<sup>2</sup>), and small landslides (less than 1000 m<sup>2</sup>). Among these landslides, the largest landslide covers an area of approximately 361485 m<sup>2</sup> whereas the smallest one is about 118.4 m<sup>2</sup>.

Landslides, in the study area, have been identified as polygons. However, due to the number of landslides are of big size, thus the center representative landslide pixels converted from polygons have been used for landslide susceptibility analysis. Out of these, a total of 983 landslide pixels have been used for generating training dataset whereas 420 remaining pixels have been utilized for creating testing dataset.

### 4.2 Landslide Affecting Parameters

In the present study, a total of ten parameters (Table 1) namely slope, aspect, elevation, land use, lithology, Normalized Difference Vegetation Index (NDVI), soil type, distance to faults, distance to geology boundaries, and valley depth were selected and used for landslide susceptibility modeling as they are known as landslide influencing factors for landslide spatial prediction [13, 41].

A Digital Elevation Model (DEM) with the spatial resolution of 30 × 30 m was generated from national topographic maps (1:50,000), and then the DEM was used to generate slope map (Fig. 2), aspect map, elevation map, and valley depth map (Fig. 3) using GIS application and SAGA software.

Soil type map (Fig. 4) at a scale of 1:100000 was collected from the Agricultural Planning and Projection of Vietnam in 2000. Geological and mineral resource maps at the scale of 1:200,000 collected from Ministry of Natural Resources and Environment of Vietnam were used to construct lithology map (Fig. 5). These maps have been also used to extract the faults which were then used to generate distance to faults map using the buffering function in ArcGIS application. Landsat 8 OLI (30 m resolution, with three paths/rows 124-50, 125-50 and 125-49) that was acquired on 14 April 2015 was used to generate NDVI map. Land use map at scale of 1:50000 was provided by the local authority of the study area. In addition, distance to geology boundaries was used for landslide susceptibility modeling. Distance to geology boundaries map was constructed using the geology boundaries extracted from geological and mineral resource

**Table 1.** Landslide affecting parameters and their classes

No	Landslide affecting parameters	Classes
1	Slope (degrees)	(1) 0–3.69, (2) 3.69–9.95, (3) 9.95–15.92, (4) 15.92–21.32, (5) 21.32–26.44, (6) 26.44–31.55, (7) 31.55–38.09, and (8) 38.09–72.49
2	Aspect	(1) flat; (2) south; (3) southeast; (4) southwest; (5) east; (6) west; (7) north; (8) northeast; and (9) northwest
3	Elevation	(1) 136.85–433.86, (2) 433.86–673.38, (3) 673.38–855.42, (4) 855.42–1037.46, (5) 1037.46–1200.34, (6) 1200.34–1372.8, (7) 1372.9–1583.58, (8) 1583.58–1880.59, and (9) 1880.59–2580
4	Valley depth (m)	(1) 0–5; (2) 5–30; (3) 30–60; (4) 60–100; (5) 100–150; and (6) >150
5	Lithology	(1) quaternary sediments; (2) sedimentary classic rocks; (3) extrusive mafic - ultramafic igneous rocks; (4) intrusive mafic - ultramafic igneous rocks; (5) extrusive acid - neutral igneous rocks; (6) intrusive acid - neutral igneous rocks; and (7) metamorphic rock with rich quartz components
6	Soil type	(1) water: (lakes and rivers); (2) Fluvisols: Humi-Gleyic Fluvisols, Dystric-Arenic-Fluvisols; (3) Gleysols: Hapli-Umbric Gleysols; (4) Cambisols: Anthraqui-Fluvic Cambisols; (5) Haplic Acrisols; (6) Chromi-Hyperdystric Acrisols, Chromi-Arenic Acrisols, Hapli-Hyperdystric Acrisols; (7) Hapli-Chromic Acrisols; (8) Alumi-Humic Acrisols; (9) Ferralsols: Humi-Acric Ferralsols, Verti-Acric Ferralsols; (10) Humi-Endoleptic Acrisols; (11) Leptosols: Dystric-lithic Leptosols; and (12) Anthraqui-Plinthic Acrisols
7	Land use	(1) wet rice land; (2) upland rice; (3) grass land; (4) annual crop land; (5) perennial crops; (6) productive forest land; (7) protective forest land; (8) Specialized use forest land; (9) rivers, lakes; (10) non-agricultural land; and (11) unused land
8	Distance to faults (m)	(1) 0–200; (2) 200–400 m; (3) 400–600; (4) 600–800; (5) 800–1000; (6) 1000–1200; and (7) >1200
9	Distance to geology boundaries (m)	(1) 0–200; (2) 200–400 m; (3) 400–600; (4) 600–800; (5) 800–1000; (6) 1000–1200; and (7) >1200
10	NDVI	(1) [(-0.23)–(0.02)]; (2) 0.02–0.12; (3) 0.12–0.19; (4) 0.19–0.25; (5) 0.25–0.3; (6) 0.3–0.35; (7) 0.35–0.4; (8) 0.4–0.45; and (9) 0.45–0.63

maps, and buffered into the study area. For landslide susceptibility assessment, different classes of these factors have been classified on the basis of the study of literatures [13, 42] and the analysis of the characteristics of the study area in relation with landslide occurrences.

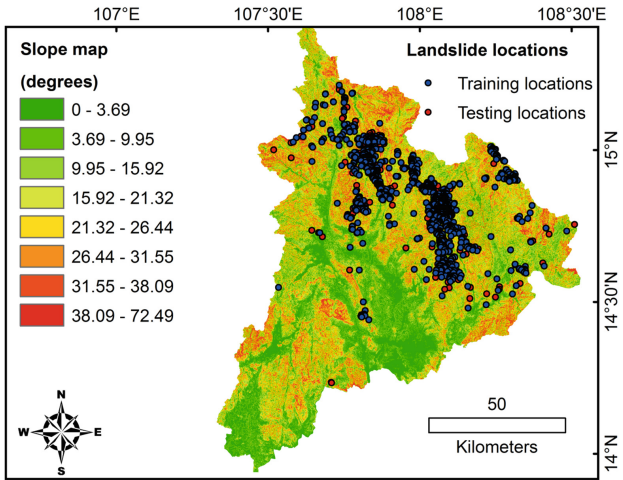


Fig. 2. Slope map of the study area

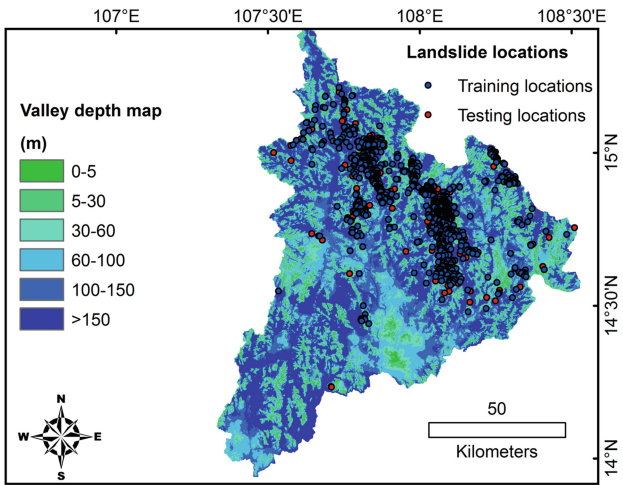


Fig. 3. Valley depth map of the study area



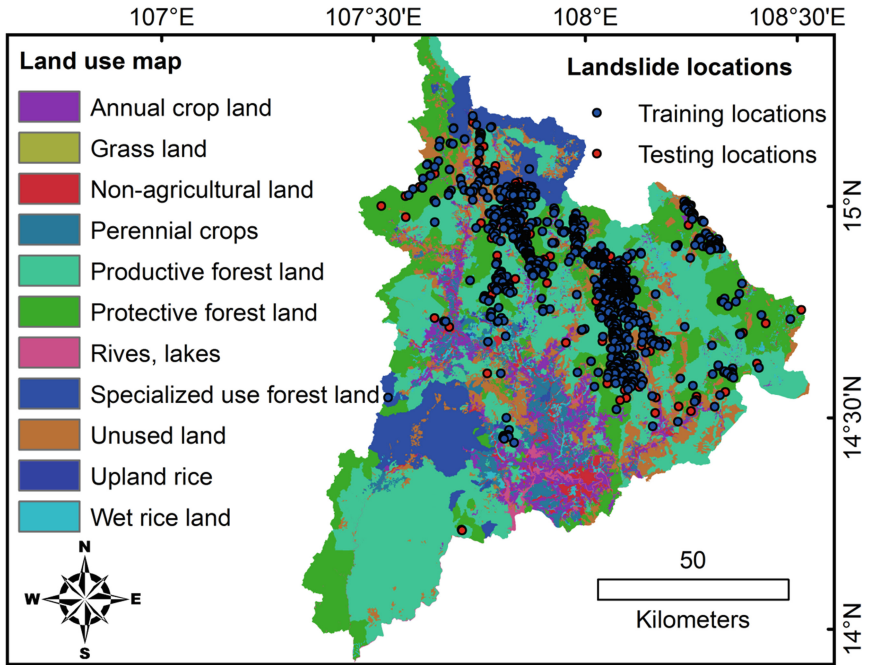


Fig. 4. Land use map of the study sarea

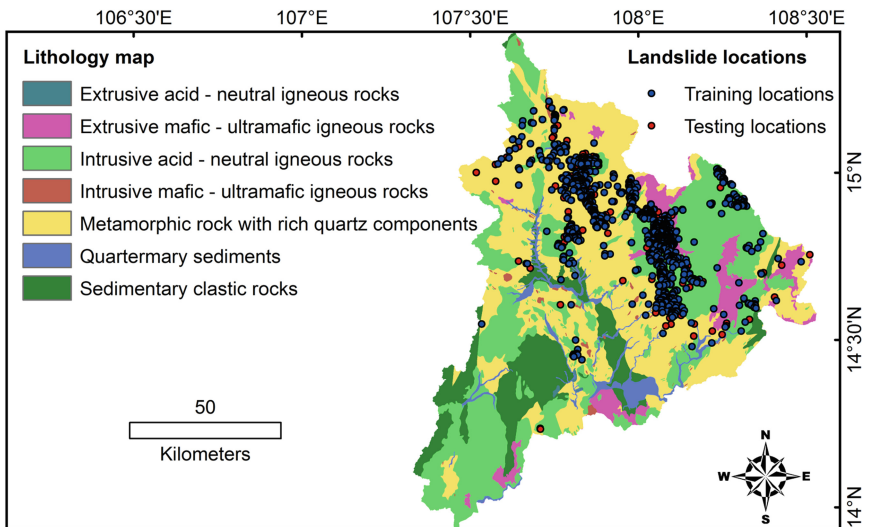


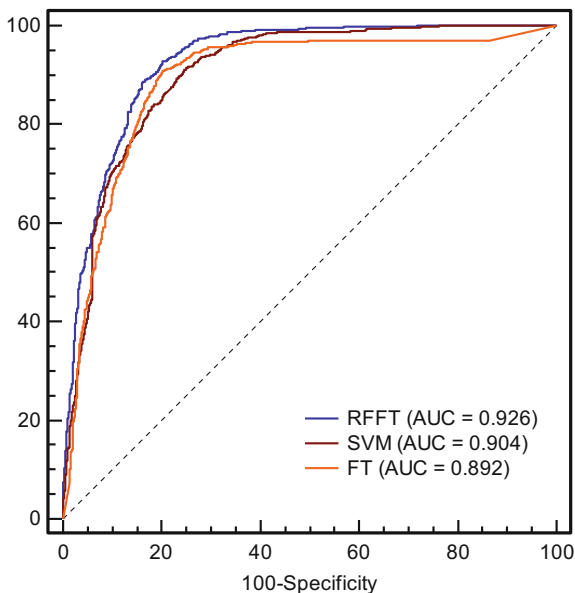
Fig. 5. Land use/ land cover map of the study area



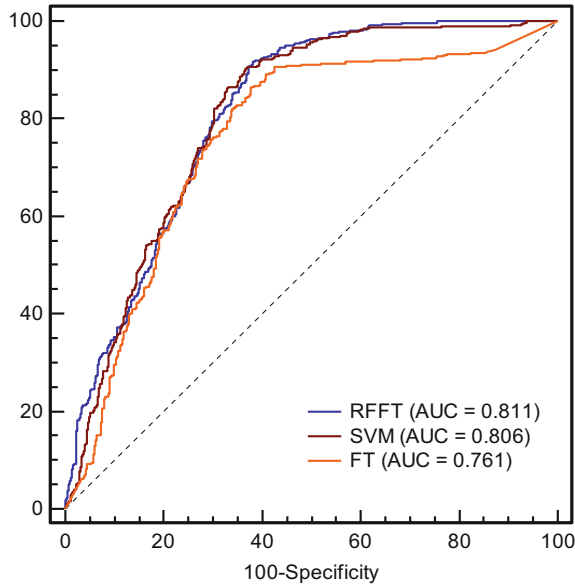
## 5 Results and Discussion

Landslide susceptibility model using the novel hybrid RFFT model has been constructed and validated using training and testing datasets, respectively, and the results are shown in Figs. 6, 7, 8, and 9. According to the ROC curve analysis (Figs. 6 and 7), it can be observed that the RFFT model has high AUC values for training (0.926), and testing (0.811) datasets. These values are higher than those of the SVM model (0.904 for training dataset and 0.806 for testing dataset), and the single FT model (0.892 for training dataset and 0.761 for testing dataset).

Likewise, according to the statistical indexes analysis (Figs. 8 and 9), the RFFT model has the highest values of PPV (80.57% for training dataset and 70.31% for testing dataset), NPV (90.05% for training dataset and 78.15% for testing dataset), SST (90.00% for training dataset and 76.29% for testing dataset), SPF (82.41% for training dataset and 72.47% for testing dataset), ACC (85.81% for training dataset and 74.23% for testing dataset), and K (0.716 for training dataset and 0.485 for testing dataset), followed by the SVM model with PPV (80.16% for training dataset and 70.30% for testing dataset), NPV (90.74% for training dataset and 77.67% for testing dataset), SST (89.65% for training dataset and 75.90% for testing dataset), SPF (82.06% for training dataset and 72.35% for testing dataset), ACC (85.45% for training dataset and 73.99% for testing dataset), and K (0.710 for training dataset and 0.480 for testing dataset), and the single FT model with PPV (79.86% for training dataset and 68.65% for testing dataset), NPV (90.54% for training dataset and 76.96% for testing dataset), SST



**Fig. 6.** The ROC curve analysis of landslide models using training dataset

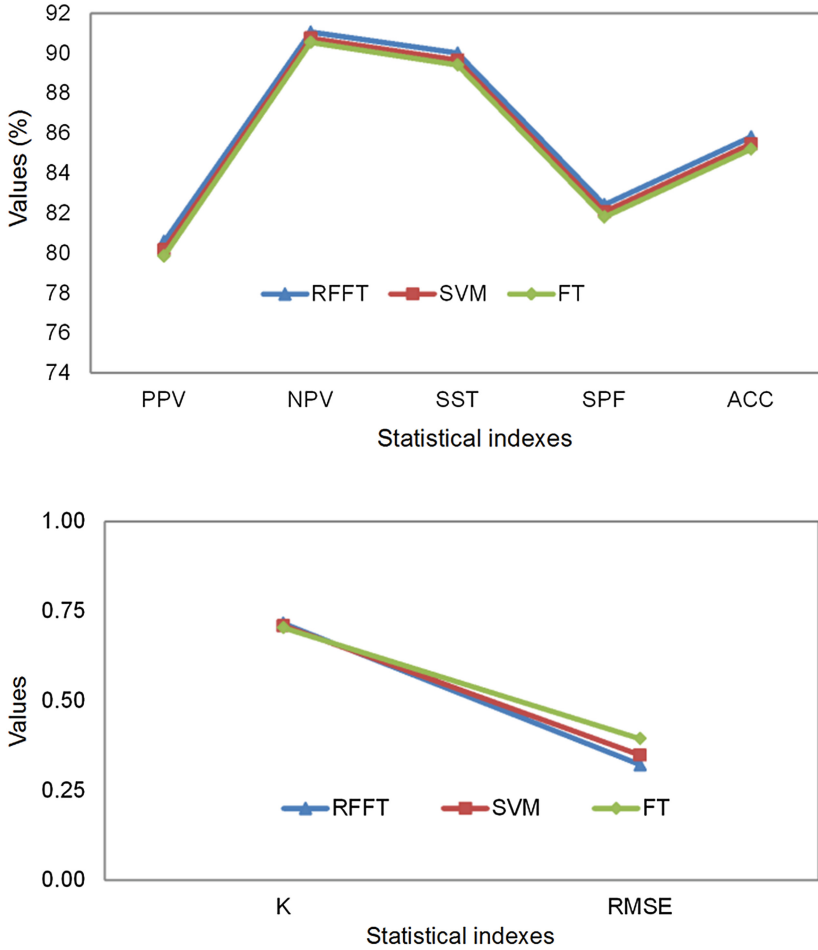


**Fig. 7.** The ROC curve analysis of landslide models using testing dataset

(89.41% for training dataset and 74.87% for testing dataset), SPF (81.80% for training dataset and 71.05% for testing dataset), ACC (85.20% for training dataset and 72.80% for testing dataset), and K (0.704 for training dataset and 0.456 for testing dataset), respectively. In contrast, the RFFT model has the lowest value of RMSE (0.322 for training dataset and 0.412 for testing dataset), followed by the SVM model (0.349 for training dataset and 0.419 for testing dataset), and the single FT model (0.395 for training dataset and 0.454 for testing dataset), respectively.

In landslide susceptibility modeling, the hybrid model is known as a great alternative tool [16]. In this study, the RFFT model, which is a hybrid approach of the RF ensemble and the FT classifier, also performed very well for landslide susceptibility modeling based on the analysis of above results as it outperforms both the well-known SVM model and single FT model. It can be reasonably explained that the RFFT model takes advantage of machine learning ensemble method of the RF ensemble for improving the prediction ability of FT classifier [17]. In the RF, the diversity is guaranteed by the rotation strategy, thus the classification accuracy of the RFFT model is better than single FT classifier [43]. In addition, the PCA used in the RF is a good filter method which will help in reducing the learning errors [17].

Landslide susceptibility map has been constructed at the Kontum province, Viet Nam using the novel hybrid model RFFT in this study. For this, a total of 24021009 pixels of the study area has been assigned the susceptibility indexes which have been generated during training the RFFT model. Thereafter, susceptibility classes namely



**Fig. 8.** Model performance using training dataset

very low, low, moderate, high, and very high were determined to construct the landslide susceptibility map (Fig. 10) based on the classification of the generated susceptibility indexes using the natural breaks method [44] in ArcGIS environment. Furthermore, landslide pixels frequency analysis on the map has been carried out to validate the reliability of the landslide susceptibility map, and the results are shown in Fig. 11. It can be observed that the highest frequency of landslides is on the very high class (69.82%), followed by the high class (15.63%), the moderate class (9.47%), the low class (3.71), and the very low class (1.37%), respectively. Based on the frequency analysis results, it is reasonably to state that the produced landslide susceptibility map is much reliable for landslide hazard management in the study area.

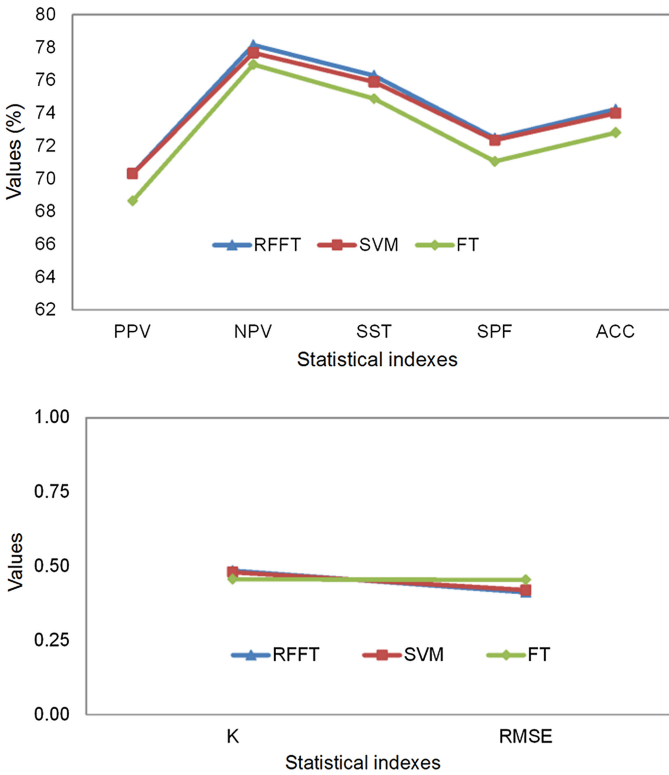


Fig. 9. Model performance using testing dataset

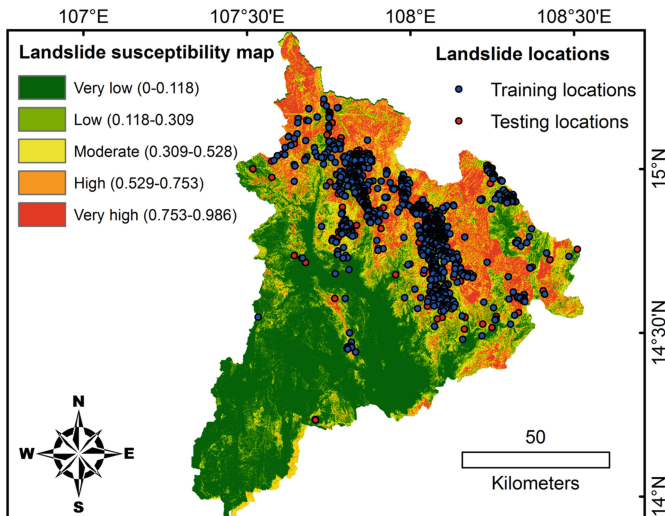
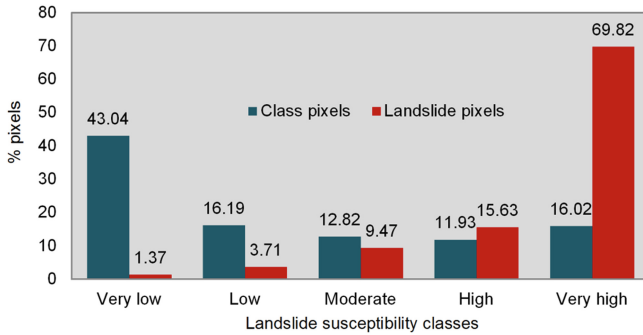


Fig. 10. Landslide susceptibility map using the proposed hybrid model RFFT



**Fig. 11.** Pixels distribution analysis on landslide susceptibility map

## 6 Concluding Remarks

Landslide susceptibility mapping has been conducted at the Kontum province, Viet Nam using the novel hybrid model named RFFT which is a combination of two state of the art machine learning methods of the RF ensemble and the FT classifier. Performance of the proposed RFFT model has been validated by analysis of the ROC curve and statistical indexes, and compared with a well-known landslide model namely Support Vector Machines (SVM), and the single FT model.

Results show that the proposed RFFT model has very good performance for landslide susceptibility assessment. It has better predictive capability compared with well-known SVM model and single FT model. Therefore, it can be concluded that the proposed RFFT model should be used as a great alternative method for better landslide susceptibility assessment in landslide prone areas. Results of this study also reconfirmed that the hybrid model using ensemble techniques such as the RF is a promising approach for landslide susceptibility modeling.

**Acknowledgement.** Authors are thankful to the Director, Bhaskarcharya Institute for Space Applications and Geo-Informatics (BISAG), Department of Science & Technology, Government of Gujarat, Gandhinagar, Gujarat, India for providing facilities to carry out this research work.

## References

1. Ercanoglu, M., Gokceoglu, C., Van Asch, T.W.: Landslide susceptibility zoning north of Yenice (NW Turkey) by multivariate statistical techniques. *Nat. Hazards* **32**, 1–23 (2004)
2. Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P.: Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology* **31**, 181–216 (1999)
3. Pourghasemi, H.R., Pradhan, B., Gokceoglu, C.: Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Nat. Hazards* **63**, 965–996 (2012)
4. Oh, H.-J., Pradhan, B.: Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput. Geosci.* **37**, 1264–1276 (2011)

5. Ermini, L., Catani, F., Casagli, N.: Artificial neural networks applied to landslide susceptibility assessment. *Geomorphology* **66**, 327–343 (2005)
6. Pham, B.T., Tien Bui, D., Pourghasemi, H.R., Indra, P., Dholakia, M.B.: Landslide susceptibility assessment in the Uttarakhand area (India) using GIS: a comparison study of prediction capability of naïve Bayes, multilayer perceptron neural networks, and functional trees methods. *Theor. Appl. Climatol.* **122**, 1–19 (2015)
7. Tsangaratos, P., Ilia, I.: Landslide susceptibility mapping using a modified decision tree classifier in the Xanthi Perfection, Greece. *Landslides* **13**, 305–320 (2016)
8. Nefeslioglu, H., Sezer, E., Gokceoglu, C., Bozkir, A., Duman, T.: Assessment of landslide susceptibility by decision trees in the metropolitan area of Istanbul, Turkey. *Math. Probl. Eng.* **2010**, 0–15 (2010)
9. Kavzoglu, T., Sahin, E.K., Colkesen, I.: Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **11**, 425–439 (2014)
10. Pham, B.T., Bui, D.T., Prakash, I., Dholakia, M.: Evaluation of predictive ability of support vector machines and naïve Bayes trees methods for spatial prediction of landslides in Uttarakhand state (India) using GIS. *J. Geomat.* **10**, 71–79 (2016)
11. Pham, B.T., Bui, D.T., Prakash, I.: Landslide susceptibility assessment using bagging ensemble based alternating decision trees, logistic regression and J48 decision trees methods: a comparative study. *Geotech. Geol. Eng.* **31**, 1–15 (2017)
12. Ohlmacher, G.C., Davis, J.C.: Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Eng. Geol.* **69**, 331–343 (2003)
13. Tien Bui, D., Pham, B.T., Nguyen, Q.P., Hoang, N.-D.: Spatial prediction of rainfall-induced shallow landslides using hybrid integration approach of Least-Squares Support Vector Machines and differential evolution optimization: a case study in Central Vietnam. *Int. J. Digit. Earth* **9**, 1–21 (2016)
14. Pham, B.T., Tien Bui, D., Prakash, I., Dholakia, M.B.: Rotation forest fuzzy rule-based classifier ensemble for spatial prediction of landslides using GIS. *Nat. Hazards* **83**, 1–31 (2016)
15. Tien Bui, D., Tuan, T.A., Hoang, N.-D., Thanh, N.Q., Nguyen, D.B., Van Liem, N., Pradhan, B.: Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. *Landslides* **14**, 447–458 (2017)
16. Dehnavi, A., Aghdam, I.N., Pradhan, B., Varzandeh, M.H.M.: A new hybrid model using step-wise weight assessment ratio analysis (SWARA) technique and adaptive neuro-fuzzy inference system (ANFIS) for regional landslide hazard assessment in Iran. *CATENA* **135**, 122–148 (2015)
17. Rodriguez, J.J., Kuncheva, L.I., Alonso, C.J.: Rotation forest: a new classifier ensemble method. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**, 1619–1630 (2006)
18. Ozcift, A.: SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease. *J. Med. Syst.* **36**, 2141–2147 (2012)
19. Jolliffe, I.: *Principal component analysis*. Wiley Online Library (2002)
20. Kuncheva, L.I., Rodríguez, J.J.: An experimental study on rotation forest ensembles. In: *International Workshop on Multiple Classifier Systems*, pp. 459–468. Springer (2007)
21. Ozcift, A., Gulten, A.: A robust multi-class feature selection strategy based on rotation forest ensemble algorithm for diagnosis of Erythemato-Squamous diseases. *J. Med. Syst.* **36**, 941–949 (2012)
22. Ozcift, A., Gulten, A.: Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. *Comput. Methods Programs Biomed.* **104**, 443–451 (2011)

23. Xia, J., Du, P., He, X., Chanussot, J.: Hyperspectral remote sensing image classification based on rotation forest. *IEEE Geosci. Remote Sens. Lett.* **11**, 239–243 (2014)
24. Kavzoglu, T., Colkesen, I.: An assessment of the effectiveness of a rotation forest ensemble for land-use and land-cover mapping. *Int. J. Remote Sens.* **34**, 4224–4241 (2013)
25. Pham, B.T., Bui, D.T., Dholakia, M.B., Prakash, I., Pham, H.V., Mehmood, K., Le, H.Q.: A novel ensemble classifier of rotation forest and Naïve Bayer for landslide susceptibility assessment at the Luc Yen district, Yen Bai Province (Viet Nam) using GIS. *Geomat. Nat. Hazards Risk*, 1–23 (2016)
26. Pham, B.T., Tien Bui, D., Prakash, I., Dholakia, M.B.: Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *CATENA* **149**(Part 1), 52–63 (2017)
27. Gama, J.: Functional trees for classification. In: *Proceedings IEEE International Conference on Data Mining, 2001, ICDM 2001*, pp. 147–154. IEEE (2001)
28. Gama, J.: Functional trees. *Mach. Learn.* **55**, 219–250 (2004)
29. Shirzadi, A., Bui, D.T., Pham, B.T., Solaimani, K., Chapi, K., Kavian, A., Shahabi, H., Revhaug, I.: Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ. Earth Sci.* **76**, 60 (2017)
30. Tien Bui, D., Ho, T.-C., Pradhan, B., Pham, B.-T., Nhu, V.-H., Revhaug, I.: GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with AdaBoost, Bagging, and MultiBoost ensemble frameworks. *Environ. Earth Sci.* **75**, 1–22 (2016)
31. Pham, B.T., Tien Bui, D., Dholakia, M.B., Prakash, I., Pham, H.V.: A comparative study of least square support vector machines and multiclass alternating decision trees for spatial prediction of rainfall-induced landslides in a tropical cyclones area. *Geotech. Geol. Eng.* **34**, 1–18 (2016)
32. Pham, B.T., Tien Bui, D., Pham, H.V., Le, H.Q., Prakash, I., Dholakia, M.B.: Landslide hazard assessment using random subspace fuzzy rules based classifier ensemble and probability analysis of rainfall data: a case study at Mu Cang Chai District, Yen Bai Province (Viet Nam). *J. Indian Soc. Remote Sens.* **45**, 1–11 (2016)
33. Dou, J., Bui, D.T., Yunus, A.P., Jia, K., Song, X., Revhaug, I., Xia, H., Zhu, Z.: Optimization of causative factors for landslide susceptibility evaluation using remote sensing and GIS data in parts of Niigata, Japan. *PLoS ONE* **10**, e0133262 (2015)
34. Bennett, N.D., Croke, B.F., Guariso, G., Guillaume, J.H., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T., Norton, J.P., Perrin, C.: Characterising performance of environmental models. *Environ. Model Softw.* **40**, 1–20 (2013)
35. Pham, B.T., Pradhan, B., Tien Bui, D., Prakash, I., Dholakia, M.B.: A comparative study of different machine learning methods for landslide susceptibility assessment: a case study of Uttarakhand area (India). *Environ. Model Softw.* **84**, 240–250 (2016)
36. Thanh, T.T.M., Vung, V.V., Miyake, H., Irikura, K.: Simulated ground motion of the earthquake on October 22nd, 2012, M4. 6 at Song Tranh 2 dam area. *J. Earth Sci.* **37**, 241–251 (2016)
37. Thanh, T.T.M., Minh, N.L., Vung, V.V., Irikura, K.: Values for peak ground acceleration and peak ground velocity using in seismic hazard assessment for Song Tranh 2 hydropower region. *Vietnam J. Earth Sci.* **36**, 462–469 (2014)
38. Nam, N.T.: Probabilistic seismic hazard assessment for the Tranh river hydropower plant No2 site, Quang Nam Province. *Vietnam J. Earth Sci.* **38**, 188–201 (2016)
39. Toan, D.V., Phong, L.H., Vu, T.A., Quang, N.T.H.: Study of the Earth's crustal structure in the Area of Song Tranh and it's adjacents. *Vietnam J. Earth Sci.* **37**, 127–138 (2015)



40. Duan, B.V., Giang, H.T., Duong, N.A., Nguyen, P.D.: About factors related to the occurrence of earthquakes in the Song Tranh 2 hydropower area in period 2011–2014. *Vietnam J. Earth Sci.* **37**, 228–240 (2016)
41. Pham, B.T., Khosravi, K., Prakash, I.: Application and comparison of decision tree-based machine learning methods in landside susceptibility assessment at Pauri Garhwal Area, Uttarakhand, India. *Environ. Processes* **4**, 1–20 (2017)
42. Pham, B.T., Tien Bui, D., Pham, H.V.: Spatial prediction of rainfall induced landslides using Bayesian network at Luc Yen District, Yen Bai Province (Viet Nam). In: *International Conference on Environmental Issues in Mining and Natural Resources Development (EMNR 2016)*, pp. 1–10 (2016)
43. Liu, K.-H., Huang, D.-S.: Cancer classification using rotation forest. *Comput. Biol. Med.* **38**, 601–610 (2008)
44. Ayalew, L., Yamagishi, H., Ugawa, N.: Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. *Landslides* **1**, 73–81 (2004)