IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 18, 2020, accepted August 2, 2020, date of publication August 6, 2020, date of current version August 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3014816

Towards an Ensemble Machine Learning Model of Random Subspace Based Functional Tree Classifier for Snow Avalanche Susceptibility Mapping

AMIRHOSEIN MOSAVI[1,2], ATAOLLAH SHIRZADI[3], BAHRAM CHOUBIN[4], FERESHTEH TAROMIDEH[5], FARZANEH SAJEDI HOSSEINI[6], MOSLEM BORJI[6], HIMAN SHAHABI[7], ARYAN SALVATI[6], AND ADRIENN A. DINEVA[8,9]

[1]Environmental Quality, Atmospheric Science and Climate Change Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam
[2]Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Vietnam
[3]Department of Rangeland and Watershed Management, College of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran
[4]West Azarbaijan Agricultural and Natural Resources Research and Education Center, Soil Conservation and Watershed Management Research Department, AREEO, Urmia 57169-63963, Iran
[5]Department of Irrigation, Sari Agricultural Sciences and Natural Resources University, Sari 48181 68984, Iran
[6]Reclamation of Arid and Mountainous Regions Department, Faculty of Natural Resources, University of Tehran, Karaj 31585-77871, Iran
[7]Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran
[8]Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
[9]Kalman Kando Faculty of Electrical Engineering, Obuda University, 1034 Budapest, Hungary

Corresponding authors: Bahram Choubin (b.choubin@areeo.ac.ir) and Adrienn A. Dineva (adriennndineva@duytan.edu.vn)

ABSTRACT Snow avalanche as a natural disaster severely affects socio-economic and geomorphic processes through damaging ecosystems, vegetation, landscape, infrastructures, transportation networks, and human life. Modeling the snow avalanche has been seen as an essential approach for understanding the mountainous landscape dynamics to assess hazard susceptibility leading to effective mitigation and resilience. Therefore, the main aim of this study is to introduce and implement an ensemble machine learning model of random subspace (RS) based on a classifier, functional tree (FT), named RSFT model for snow avalanche susceptibility mapping at Karaj Watershed, Iran. According to the best knowledge of literature, the proposed model, RSFT, has not earlier been introduced and implemented for snow avalanche modeling and mapping over the world. Four benchmark models, including logistic regression (LR), logistic model tree (LMT), alternating decision tree (ADT), and functional trees (FT) models were used to check the goodness-of-fit and prediction accuracy of the proposed model. To achieve this objective, the most important factors among many climatic, topographic, lithologic, and hydrologic factors, which affect the snow accumulation and snow avalanche occurrence, were determined by the information gain ratio (IGR) technique. The goodness-of-fit and prediction accuracy of the models were evaluated by some statistical-based indexes including, sensitivity, specificity, accuracy, kappa, and area under the ROC curve, Friedman and Wilcoxon sign rank tests. Results indicated that the ensemble proposed model (RSFT), had the highest performance (Sensitivity = 94.1%, Specificity = 92.4%, Accuracy = 93.3%, and Kappa = 0.782) rather than the other soft-computing benchmark models. The snow avalanche susceptibility maps indicated that the high and very high susceptibility avalanche areas are located in the north and northeast parts of the study area, which have a higher elevation with more precipitation and lower temperatures.

INDEX TERMS Snow avalanche, susceptibility mapping, ensemble approach, feature selection.

I. INTRODUCTION

Snow avalanche is referred to the sudden descending unstable snowpack often seen in cold and high altitude regions [1], [2]. Snow avalanche, as a natural disaster severely affects socio-economic and geomorphic processes through damaging ecosystems, vegetation, landscape, infrastructures, transportation networks, and human life [3], [4]. The recent frequency, irregularity, and uncertainty in the occurrence regime of avalanche had been strongly linked to climate

The associate editor coordinating the review of this manuscript and approving it for publication was Qichun Zhang.

change and global warming [5]–[8]. Therefore, an avalanche is fast becoming a significant risk in more nations [9]–[11]. Consequently, the advancement of the novel methods for hazard assessment and risk management of snow avalanche has increasingly become popular around the world for effective mitigation and resilience [12]–[14]. Modeling the snow motion has shown to be essential for understanding the mountainous landscape dynamics to assess the avalanche hazard susceptibility [15]–[17].

Modeling the snow avalanche is multidisciplinary in nature and literature demonstrates significant overlapping with atmosphere, hydrosphere, biosphere, lithosphere, and anthroposphere [18]–[23]. Among the principal interacting factors motivating the snow avalanche, the meteorology, terrain, and snowpack had been well studied in the literature [11], [24], [25]. On the other hand, the recent rising popularity of winter sports is increasing social exposure to snow avalanches hazard and a meaningful rise in fatal incidents [26]. Therefore, the advancement of novel techniques for susceptibility mapping has become widely popular to identify the hazardous regions [15], [16], [27], [28]. Susceptibility mapping techniques can complement early warning systems and temporal models to save lives [29]–[32]. Traditional techniques for susceptibility mapping limits to geospatial modeling using either simple statistical or decision analysis methods [15], [17], [27], [33]–[35]. Although the application of machine learning (ML) for susceptibility mapping of a wide range of natural hazards has been established, snow avalanche is yet to be explored [36].

ML methods have recently gained popularity for advancing high-performance prediction models in numerous application domains. ML models have shown promising results in a diverse range of applications, e.g., hydrological modeling [37], [38], earth system modeling [39], civil engineering and estimation [40], energy systems sciences [41], [42], and geosciences modeling [43]. ML models for susceptibility mapping have shown to outperform conventional models [44]. For instance, landslide and flood susceptibility mapping with machine learning presents promising results [45], [46]. However, applications of ML methods for modeling snow avalanche have been not adequate [47]. Recently Choubin *et al.* [48] used the support vector machine (SVM) considering several variables, e.g., terrain characteristics, avalanche locations, and meteorological factors. In another attempt, Rahmati *et al.* [49], in addition to SVM, used random forest and naïve Bayes to identify the susceptibility map and also to provide insight into the most important factors of snow avalanche distribution. Considering the limited number of literature on this realm, further research is essential to explore the application of more sophisticated machine learning methods to discover models with higher accuracy. Consequently, contribution of the present article is to explore the application and performance of an ensemble intelligent learning model to snow avalanche susceptibility mapping. According to the best of our knowledge, the proposed model, RSFT, has not earlier been introduced and implemented for snow avalanche modeling and mapping over the world. A comparative analysis of single models, several decision-trees, and an ensemble model is presented in this paper.

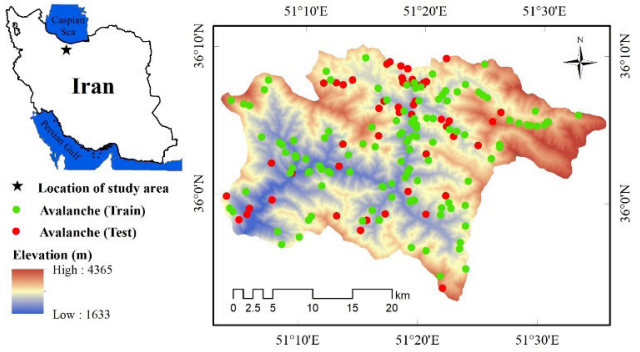

FIGURE 1. Location of the study area.

II. MATERIAL AND METHODS

A. STUDY AREA; DESCRIPTION AND LOCATION

The study area is Karaj watershed located between latitudes 35° 52' and 36° 10' N and longitudes 51° 03' and 51° 36' E in Iran (Fig. 1), with an area of is 845 km^2. Topographically, the elevation of the watershed changes from 4365 m to 1633 m a.s.l. respectively in the east and southwest. The climate of the watershed is mostly extra-cold-humid and cold-humid. The annual precipitation varies between 300 to 600 mm, which falls mostly as snow. This watershed is the main source of drinking water and irrigation water for Karaj and Tehran plain [50], [51].

The Karaj-Chalous way is located in this watershed. It usually has considerable traffics during the year, due to the main path between capital and the Caspian Sea, beautiful nature, and winter games. Therefore, heavy transit across stony and high-gradient paths makes this watershed to be important because of the natural hazards such as snow avalanches. According to the reports, the avalanches in this watershed have been caused to kill the people and fall vehicles into valleys [48].

B. SNOW AVALANCHE INVENTORY

Location of the occurred snow avalanches during the snow season (i.e., from December to March) from 2019 to 2020 were collected using field surveys. At first, according to the digital topographic and morphometric layers (such as valleys and slope) and Google Earth images, the sensitive hillsides to snow avalanches were determined. This helped to identify the probable location of the snow avalanches. Then, the estimated locations were visited and confirmed by the various field surveys, and finally, the final snow avalanche inventory map was prepared (Fig. 1). A total set of 171 snow avalanches were recorded, of which 70% (119 cases) of the data were randomly applied for training, and the rest (30% or 52) were considered for validating the machine learning models.

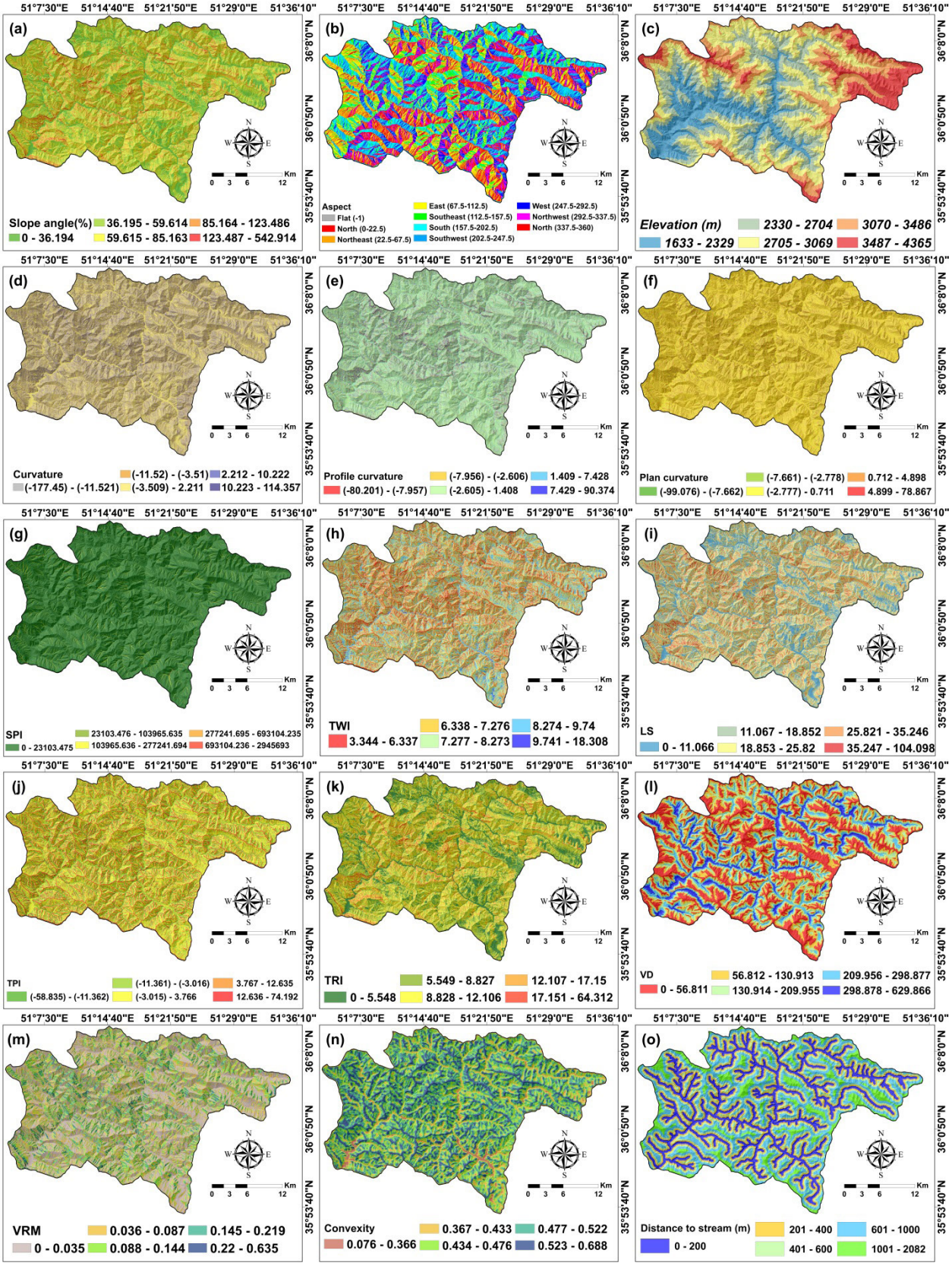

FIGURE 2. Snow avalanche conditioning factors: a) slope angle, b) aspect, c) elevation, d) curvature, e) profile curvature, f) plan curvature, g) stream power index (SPI), h) topographic wetness index (TWI), i) length of slope (LS), j) topographic position index (TPI), k) terrain roughness index (TRI), l) valley depth (VD), m) vector ruggedness measure (VRM), n) convexity, o) distance to river (DTR), p) drainage density, q) temperature, r) precipitation, s) distance to fault (DTF), t) lithology, and u) land use.

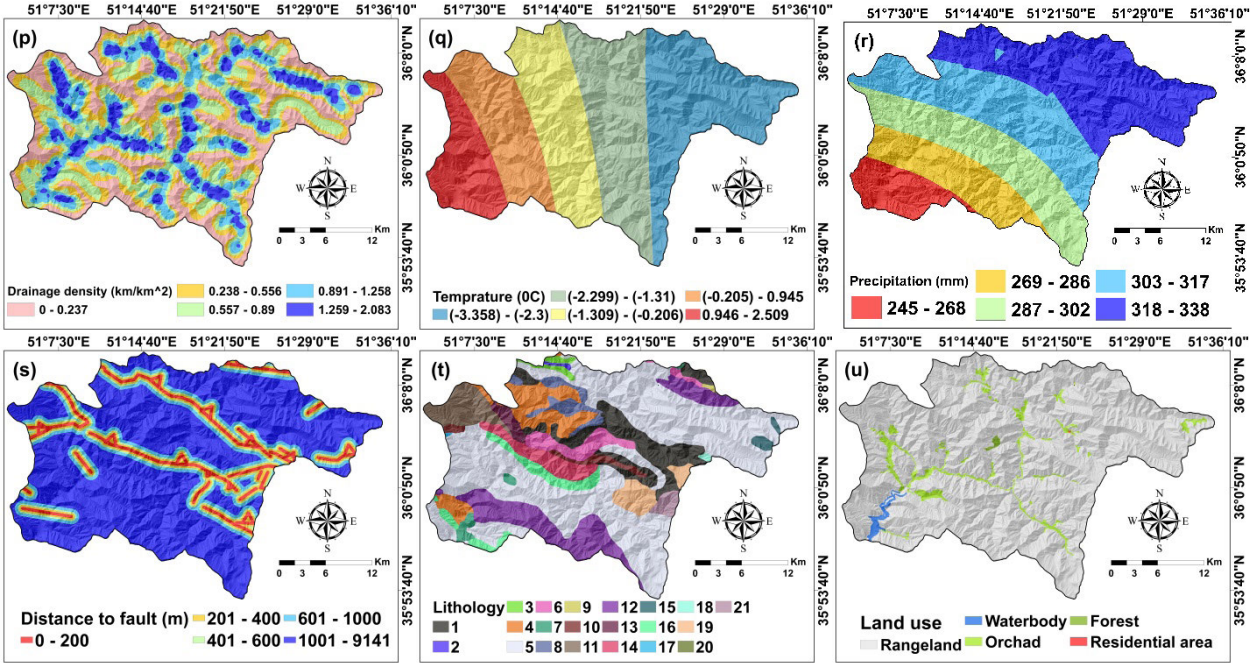

FIGURE 2. ***(Continued.)*** **Snow avalanche conditioning factors: a) slope angle, b) aspect, c) elevation, d) curvature, e) profile curvature, f) plan curvature, g) stream power index (SPI), h) topographic wetness index (TWI), i) length of slope (LS), j) topographic position index (TPI), k) terrain roughness index (TRI), l) valley depth (VD), m) vector ruggedness measure (VRM), n) convexity, o) distance to river (DTR), p) drainage density, q) temperature, r) precipitation, s) distance to fault (DTF), t) lithology, and u) land use.**

C. SNOW AVALANCHE CONDITIONING FACTORS

The construction and the flow regime of snow avalanches forcefully belong on the topography of the path. Its dynamics are generally controlled by the intrinsic features of snow, which are unstable and maybe to change during the flow itself [52]. However, various climatic, topographic, lithologic, and hydrologic variables affect the snow accumulation and snow avalanche occurrence. So, in this research, many important factors including elevation, slope, aspect, curvature, profile curvature, plan curvature, terrain roughness index (TRI), topographic position index (TPI), vector ruggedness measure (VRM), length of slope (LS), valley depth (VD), convexity, rainfall, temperature, lithology, distance to fault, topographic wetness index (TWI), stream power index (SPI), drainage density (DD), distance to river, and land uses were considered (Fig. 2).

An ASTER digital elevation map (DEM) with a cell size of 10×10 m was applied to calculate topographic variables. The elevation (Fig. 2c) of a beginning region is remarkable in that more snowfalls at greater elevations and there is more snow loading as a result of wind by the procedure of windward slopes being scoured and deposition happening over ridges onto slopes [53]. The mixed topographic effects can create situations that are desirable for avalanche formation [54]. The key characteristic for a beginning region is the attendance of a convenient slope (generally $25° - 45°$, Fig. 2a) that permits an avalanche to begin and move downhill [55]. The shape of slope impresses its capability to gather snow. The concavity or convexity (Fig. 2n) in the transverse profile (cross partial) orientation can affect deposition and depth of snow. Those paths that have a curved horizontal or cross-sectional profile are capable to ensnare breezing snow from many orientations related to the wind orientation [56]. Length of slope (LS) is the interval from the origination of the overland stream along its stream track to the position of either focused stream or deposition. The LS was evaluated by measuring perpendicular to the contour (Fig. 2i). Aspect (Fig. 2b) influences solar radiation earned at the snowpack and impresses the temperature gradient and snow level heating [57]. TPI (Fig. 2j) measures the comparative topographic situation of the focal point as the difference between the elevation at this point and the average elevation within predestinated vicinity. Employing TPI, landscapes can be categorized in slope situation classes [58], [59]. One exclusively appealing vector dispersion method is the vector ruggedness measure expanded by Sappington *et al.* [60], based on the vector approach offered by Hobson [61]. VRM (Fig. 2m) shows the roughness of the terrain that is most significant in the motion of snow avalanches. The measure is well suited to catch the procedure of ground smoothing of snow [62]. The TRI (Fig. 2k) represents the quantity of elevation diversity among contiguous cells of a digital elevation grid. It computes the diversity in elevation amounts from a center cell, and the eight surrounding cells [63]. The helpful geomorphological data can be gotten utilizing curvature analysis [64]. Curvature (Fig. 2d) is the initial derivation of the slope or aspect and parameterizes the concaveness and convexity of the region [65]. The process can be accomplished across

(plan curvature) or along the decline line of the slope (profile curvature). In this study, the profile curvature (Fig. 2e) is calculated as the alternation in the slope angle. Plan curvature (Fig. 2f) used for removing greatly concave or convex regions limiting the fraction diffusion of avalanche liberate and is evaluated as the alternation in aspect [24], [66], [67]. Valley depth (VD) (Fig. 2l) is computed as the diversity between the elevation and an interpolated ridge level. It was estimated by using the homonym algorithm on SAGA GIS.

Precipitation and temperature are the most momentous meteorological elements. Because rainfall and near-surface temperature variations can affect both the attributes of available slabs and the potential for the constitution of subsequent weak layers at the snowpack, therefore, they are the main consideration when estimating snow stability [53], [68]. In the current study, these factors are considered during the snow season, i.e., from December to March (Fig. 2q and Fig. 2r). The data recorded by the stations were achieved from the Iranian water Resource management company (IWRMC) also the Iran Meteorological Organization.

Lithology (Fig. 2t) and land use (Fig. 2u) affect snowpack, too. Land use, like the attendance of jungle, can bridle avalanching. In accordance with the land use map, rangelands extend most of this study region (Fig. 2u). Boulder structures and lithology are the principal agents accountable for the slope fracture in indigenous hazards, like landslides, snow avalanches, and other mass movements. Various lithological units may vary the capability of the study region into an avalanche area [15].

The proximity of a snowpack to faults or rivers can speed up the snowpack motion and construct a snow avalanche. The regions adjacent faults or rivers must be more sensitive to the avalanche occurrence. Spatial Analyst Tool (i.e., Euclidean distance) in the ArcGIS environment was used to calculate DTF and DTR (Fig. 2s and Fig. 2o). Another parameter is SPI (Fig. 2g), which evaluates the potential of streams to change the geomorphology of a region. SPI is the extent of the erosive power of flowing water by perceiving the relation between discharge and particular catchment zone [69], [70]. The drainage density (Fig. 2p) of an area is determined by the slope measure, the nature and propensity of the bedrock, and likewise by the structure and species of the geological configuration. They reflect the density of flows per unit ground and can be of momentous value for avalanche happening. TWI index plays a serious role in the spatial variation of hydrological conditions as soil moisture does. In this paper, TWI was prepared in SAGA GIS software (Fig. 2h).

D. MULTICOLLINEARITY DIAGNOSTIC STATISTICS TEST

As the results of the modeling process may be affected by the correlation between the factors, the multicollinearity of the factors should be analyzed. The tolerance ($TOL = 1-R^2$) and its reciprocal, named variance inflation factor ($VIF = 1/1-R^2$) measures are the two essential indicators to check the correlation among the independent variables. A TOL less than 0.2, and VIF exceeds 10 are the indicators for occurring the multicollinearity between independent variables [71], [72].

E. FACTOR SELECTION AND RANKING BY INFORMATION GAIN RATIO (IGR) TECHNIQUE

Feature selection with higher predictive ability is an essential step in hazard modeling [73]–[75]. Information gain ratio (IGR) is a ratio of information gain to the inherent information. This technique was suggested by Quinlan [76], to decrease a bias towards multi-valued properties by capturing the number and extent of branches into account when selecting a property. IGR biases the decision tree versus perceiving properties with a major number of distinguished amounts. The IGR is often applied to decide which of the parameters are the most relevant to use in the modeling process. In this study, a 10-folds cross-validation technique was used to obtain the final decision of the IGR factor selection method. The IGR weight of the slope angle in cluster C, is calculator as follows [76], [77]:

$$\begin{aligned}
\text{Gain ratio (Slope, C)} &= \frac{\text{Gain(Slope, C)}}{\text{Spliti info (Slope)}} \\
\text{Gain (Slope, C)} &= \text{Entropy (Slope, C)} \\
&\quad -\text{Entropy}_{\text{p}}\text{(Slope, C)}, \\
\text{Entropy (Slope, C)} &= -\text{p(Slope}\,|\text{C)}\log_2\text{p(Slope}\,|C) \\
&\quad (1-\text{p(Slope}|C))\log_2(1-p(Slope|C)), \\
\text{p(Slope}\,|\text{C)} &= \text{freq(Slope, C)}\big/|\text{C}|, \\
\text{Entropy}_{\text{p}}\text{(Slope, C)} &= \sum_{\text{i}} \frac{|\text{C}_\text{i}|}{|\text{C}|}\text{Entropy (Slope, C)}, \\
\text{Spliti info (Slope)} &= -\sum_{\text{i}} \frac{|\text{C}_\text{i}|}{|\text{C}|}\log \sum_{\text{i}} \frac{|\text{C}_\text{i}|}{|\text{C}|},
\end{aligned} \tag{1}$$

where the slope is slope angle as one of the snow avalanche conditioning factors, C is the class label (snow avalanche and non-snow avalanche), freq(Slope, C), C_i and $|C_i|$ are the frequency of the slope angle class in C, the *i*-th sub-cluster of C, and the number of slope angle class in C_i, respectively.

F. MODEL DESCRIPTION AND IMPLEMENTATION

After preparing the predictive variables and selecting the key ones using the IG, the k-fold (k = 10) cross-validation method was applied to calibrate the models by the 70% of the data. Then using the rest of the dataset (30%) the modeling process was evaluated. A summary description of the applied models is presented as follows:

Logistic regression: The major difficulty of implementing statistical models for avalanche susceptibility is concerned with the spatial variation of the morphological, metamorphism, snowpack dynamics factors. Furthermore, the dynamics of sun and wind and the angel of exposure would increase the model uncertainty and the overall snowpack quality. The efficacy of these parameters is modeled using the statistical technique of binary logistic regression (LR). LR is generally utilized to clarify event occurrence or

non-occurrence (dichotomous associate variable) from a set of parameters [78]. Such models have mostly been applied with prosperity in hazard management to illustrate the location of occurrences [79].

Logistic Model Tree (LMT): The LMT basically consists of a model with a dependent controlled training algorithm that incorporates decision tree learning and LR. LMT is established according to the initial opinion of a model tree: a decision tree that has linear regression models at its leaves to supply a piecewise linear regression model also usual decision trees with permanents at their leaves would generate a piecewise permanent model [80]. In the logistic variant, the LogitBoost algorithm is utilized to generate a logistic regression model at each knot in the tree. Therefore, the knots are divided with the C4.5 pattern. Every LogitBoost invocation is warm-started from its outcomes in the parent knot. Eventually, the tree is pruned [81].

Alternating Decision Tree (ADT): The ADT is a generalization of weighted aggregates of decision trees (DT) for classification. ADT integrates the DT and boosting technique to produce decision legislations. ADT could create easy decision tree constructions where decision legislations are simple to explicate. Moreover, the training procedure needs fewer repetitions with the use of the AdaBoost algorithm [82]. An ADT includes of two sorts of knots formed in intermittent layers, prediction knots, and decision knots. Decision knots determine a declare situation whiles prediction knots include a unit number. ADT always has prediction knots as both root and leaves. A sample is categorized by an ADT by pursuing all ways for which all decision knots are correct and collecting any prediction knots that are wended [83], [84].

Functional Tree Base classifier: A functional tree (FT) is an algorithm that has represented hopeful consequences in other environmental fields such as landslide, floods, and forest fires; however, it has rarely been probed for snow avalanche modeling. The principal discrepancy between FT and traditional decision tree algorithms is that FT applies logistic regression functions for the dividing in the internal knots (named diagonal split) and forecast at the leaves, whiles these traditional algorithms apportion the entrance data at tree knots by comparing the quantity of several entrance features with a stable [83], [85]. FT has three variants of (i) FT internal applies regression models for only the internal knots, (ii) FT leaves applied regression models for only leaves, and (iii) the complete FT that applies regression models for the internal knots and the leaves, which used in this study.

Random Subspace ensemble classifier: Random Subspace (RS) classifier presented by Ho [86] presents an ensemble of diverse classifiers for particular data space. Classification consequences are on the basis of these specific classifiers' generation by most voting. As RS is a subspace of the principal data extent, the training objects would be smaller for the principal data, whereas it is bigger for the subspace data. The principal data extent is diminished, but the training object extent remains identical, which imputes more training sample extent, helps for better classification [87]. This algorithm is an appropriate selection where there are a large number of properties. RS is widely adopted by scholars for plenty of proofs such as the clarified model, plain explanation, and training times are lesser compared to others, increased extension with reduced overfitting. The most important parameters used in this ensemble model are the number of seed and the number of iterations that the optimal values of these factors are obtained base on trial and error technique. The RS method can increase generalization accuracies of decision tree-based classifiers without loss of accuracy on training data, which is one of the major problems when it comes to tree-based classifiers [88]. However, the proposed model, RS-FT, uses the FT as a base classifier and the RS to enhance the power prediction of the FT algorithm.

G. MODEL COMPARISON AND VALIDATION MEASURES

1) STATISTICAL-BASED EVALUATED MEASURES

An arbitrary test seldom ignores the stuff you are searching for (i.e., it is sensitive) and seldom mistakes it for something else (i.e., it is specific). Hence, when estimating diagnostic tests, it is significant to compute the sensitivity and specificity for that test to specify its benefit. The sensitivity of a test is likewise named the true positive rate (TPR) and is the ratio of samples that are really positive that impute a positive consequence using the test in question [89]. The specificity of a test also mentioned as the true negative rate (TNR), is the ratio of samples that test negative applying the test in question that are truly negative. Sensitivity and specificity of a diagnostic test are represented such as the possibility (as a percent). Accuracy demonstrates the deficit of forecasts that are true. Accuracy confines from 0 to 1 that 1 display the complete prediction. Kappa statistic evaluates the accuracy when compared with a random classifier (which exhibits a Kappa quantity of 0%). The greater the statistic is, the more accurate the consequence is [83]. AUC (Area Under Curve) of the ROC curve is one of the most significant estimate metrics for surveying any performance of the classification model. It shows the degree or measure of separability. It defines how much model is able to identify between the classes. An AUC quantity of 1 demonstrates a complete model. For this statistic 0 is a non-informative model [90], [91].

2) NON-PARAMETRIC STATISTICAL TESTS

The Freidman [92] and Wilcoxon tests [93] are non-parametric tests that were other evaluation metrics used to check the models' performance. Friedman's test shows generally the performance of the models [94], [95]. The null hypothesis indicates lack of a significant difference among the models, and if it is not true (p-value < 0.05), the null hypothesis will reject. In this situation, we can conclude that all models are significantly different in terms of performance. But it is not known between which models there are statistical differences. Therefore, to detect this challenge the Wilcoxon signed-rank test is often used to check a pairwise comparison

between two or more models. The null hypothesis in this test is judged based on p-value and z-value. If between two or more models p-value and z-value are < 0.05 and beyond the critical values of ± 1.96, respectively, then the null hypothesis is not true and rejected and there is a significant statistical difference between two or more models [96], [97].

III. RESULTS AND DISCUSSION

A. MULTICOLLINEARITY TEST

The result of the correlation among the independent variables is shown in Table 1. It indicating that there is no multicollinearity between the variables considered in this study [71], [72], [98]. Therefore, we considered all the factors for the modeling process by machine learning.

TABLE 1. The multicollinearity analysis for independent variables.

Factor	TOL	VIF
Slope	0.833	1.201
Aspect	0.832	1.202
Elevation	0.542	1.844
Curvature	0.813	1.230
Profile curvature	0.869	1.150
Plan curvature	0.605	1.654
Terrain roughness index (TRI)	0.796	1.256
Topographic position index (TPI)	0.456	2.193
Vector ruggedness measure (VRM)	0.857	1.167
Length of slope (LS)	0.473	2.113
Valley depth (VD)	0.392	2.549
Convexity	0.639	1.565
Rainfall	0.838	1.193
Temperature	0.826	1.210
Lithology	0.820	1.219
Distance to fault (DTF)	0.845	1.184
Stream power index (SPI)	0.487	2.052
Topographic wetness index (TWI)	0.630	1.588
Drainage density	0.361	2.773
Distance to river	0.445	2.247
Land use	0.809	1.236

B. THE MOST IMPORTANT FACTORS FOR SNOW AVALANCHE OCCURRENCE

The predictive capability of the snow avalanche conditioning factors using 10-folds cross-validation of the IGR technique is shown in Fig. 3. This figure has composed of red and blue rectangles that indicate ineffective and effective factors on snow avalanche occurrence in this study, respectively. According to this, the most important factor is TPI with AM equals to 0.306, follows by Slope (0.191), SPI (0.190), VD (0.157), DD (0.126), LS (0.113), distance to river (0.098), convexity (0.088), lithology (0.087), Aspect (0.045), plan curvature (0.044), TWI (0.034), and temperature (0.030). Additionally, it is concluding that elevation, VRM, curvature, profile curvature, distance to fault, TRI, rainfall, and land use factors due to having AM equals to 0 and non-having predictive capability were not taken into account for modeling process and they were removed from the snow avalanche modeling.

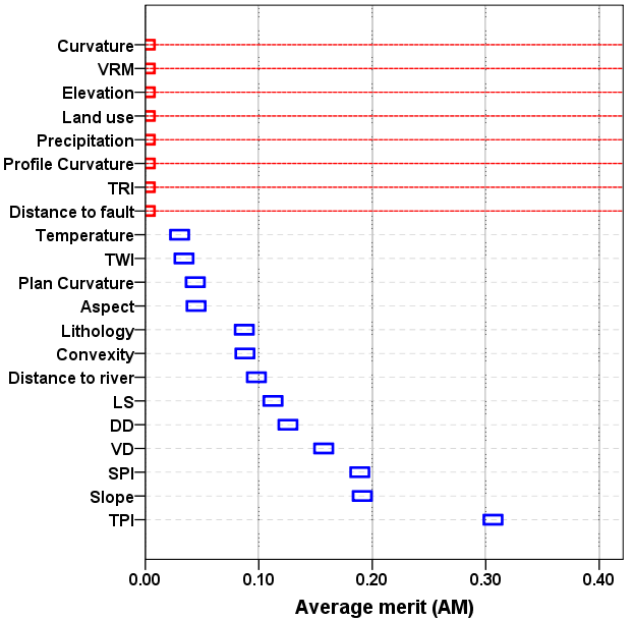

FIGURE 3. The most important snow avalanche conditioning factors in the study area.

Our results indicate that topographic factors (TPI, slope, LS, VD, convexity, aspect and plan curvature) along with climatic factors (temperature), geological factors (lithology) and hydrological factors (SPI, DD, distance to river and TWI) are responsible for snow avalanche occurring in the study area. Hydrologically, terrain factors such as TPI and slope angle (> 30 degrees) are the main factors that a?ects snowpack stability [48], [57], [99]; therefore, these factors can control or accelerate the avalanche occurrences [55], [100]. Few studies have been conducted on snow avalanche susceptibility mapping because of its high complexity [101]. For example, Choubin *et al.* [48] predicted the prone areas to snow avalanche using SVM and multivariate discriminant analysis (MDA). They pointed out that the TPI and slope angle was the most significant factor based on the sensitivity analysis of these two mentioned models. It indicates the obtained results in agreement with their results.

C. MODEL PERFORMANCE, VALIDATION, AND COMPARISON

We successfully trained the RS ensemble model based on selecting the optimal values of the number of seed and iteration as two most important parameters used in the RS ensemble model. Fig. 4a-b is shown the results of the modeling process by the new proposed model using AUC measures. Regarding these Figures, it is observed that the optimum values for number of seeds and iterations were 4 and 10, respectively. Indeed, based on these obtained values the RS was optimized and the result is shown in Tables 1 and 2 using training (performance/goodness-of-fit) and validation (prediction accuracy) datasets. According to Table 2, results of the training dataset showed that the most sensitivity was

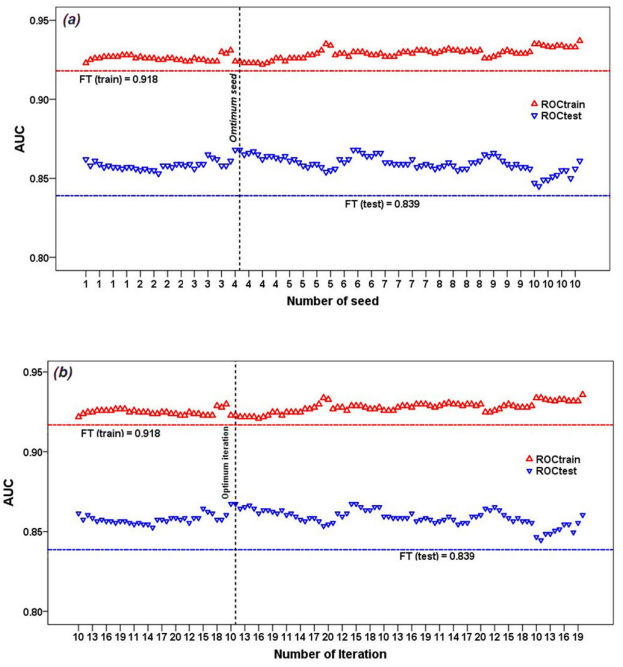

FIGURE 4. Number of the seeds and iterations of RS ensemble model; illustration of the optimal values.

obtained for the RS ensemble model (94.1%) indicated that 94.1% of the pixels determined as snow avalanche correctly classified as snow avalanche pixels. It is followed by LR and LMT (92.4%), FT (91.6%) and ADT (83.3%) models. Moreover, the highest value of specificity (92.4%) was acquired for the proposed ensemble model, RS. It is indicated that 92.4% of non-snow avalanche pixels correctly classified as non-snow avalanche pixels. The ADT and LR, LMT, and FT with values of 91.5% and 87.4% were ranked in the next positions. The accuracy is a standard and remarkable measure that can represent both sensitivity and specificity was the highest for the new proposed model (93.3%). However, the LR, LMT, and FT models with accuracy equal to 89.9% and ADT model with accuracy equal to 87% were other powerful compared models. Results also based on F1-score indicated that the RS ensemble model (93.3%) outclassed the LR and LMT (90.2%), the FT (89.7%) and the ADT (87.6%) models. The kappa similar to the above-mentioned measures also was the highest value for the RS ensemble model (0.812), followed by LR and LMT (0.798), FT (0.790) and ADT (0.740) models. Eventually, the AUC as one of the most important measures in the assessment of the model was the highest for the RS (0.935) model. It is followed by LR (0.930), LMT (0.927), ADT (0.929) and FT (0.918) models.

From Table 3, based on the validation dataset, the results also indicated a higher ability for the RS ensemble model in terms of sensitivity (90.4%), specificity (92.3%), accuracy (91.3%), kappa (0.615), and AUC (0.868). Furthermore, the result indicated the superiority of the ADT model (75%) to LR, LMT, and FT models (73.1%) in terms of sensitivity.

TABLE 2. Model performance using the training dataset.

Criterion	RS	LR	LMT	FT	ADT
True positive (TP)	112	110	110	109	110
True negative (TN)	110	104	104	104	97
False positive (FP)	9	15	15	15	9
False negative (FN)	7	9	9	10	22
Sensitivity (%)	94.1	92.4	92.4	91.6	83.3
Specificity (%)	92.4	87.4	87.4	87.4	91.5
Accuracy (%)	93.3	89.9	89.9	89.5	87.0
F1-score	93.3	90.2	90.2	89.7	87.6
Kappa	0.782	0.798	0.798	0.790	0.740
AUC	0.935	0.930	0.927	0.918	0.929

However, the specificity based on the validation dataset was 86.5%, 84.6%, and 82.7%, respectively for FT, LMT, and LR and ADT commonly. After the proposed ensemble model, the FT had the highest accuracy (79.8%), followed by LMT and ADT (78.8%), and LR (77.9%). Similar to the earlier evaluation metrics, results concluded that the F1-score were the most values for the RS ensemble model (91.3%). It was followed by the FT, ADT, LMT, and LR models. According to the kappa and AUC, the powerful models after the RS model were known as LMT, FT, LR, and ADT. As a final result, it can be concluded that the new ensemble proposed model, RS, well trained and successfully had the highest performance and prediction accuracy base on the training and validation datasets, respectively, compared to other well-known soft computing benchmark models.

TABLE 3. Model performance using the validation dataset.

Criterion	RS	LR	LMT	FT	ADT
True positive (TP)	47	38	38	38	39
True negative (TN)	48	43	44	45	43
False positive (FP)	4	9	8	7	9
False negative (FN)	5	14	14	14	13
Sensitivity (%)	90.4	73.1	73.1	73.1	75.0
Specificity (%)	92.3	82.7	84.6	86.5	82.7
Accuracy (%)	91.3	77.9	78.8	79.8	78.8
F1-score	91.3	76.8	77.6	78.4	78
Kappa	0.615	0.558	0.577	0.596	0.577
AUC	0.868	0.828	0.839	0.839	0.800

The proposed model, RS, had a high goodness-of-fit and prediction accuracy for the study area, but the comparison of this model with previous studies in this field is not possible, due to lack of similar study. According to the literature, it can be declared that this work is a pioneer study on snow avalanche susceptibility by a machine learning ensemble model. Although many studies have been focused on the RS as an ensemble model on natural hazard events such as landslide susceptibility mapping [94], [102]–[104], and groundwater potential mapping [105], there is no research

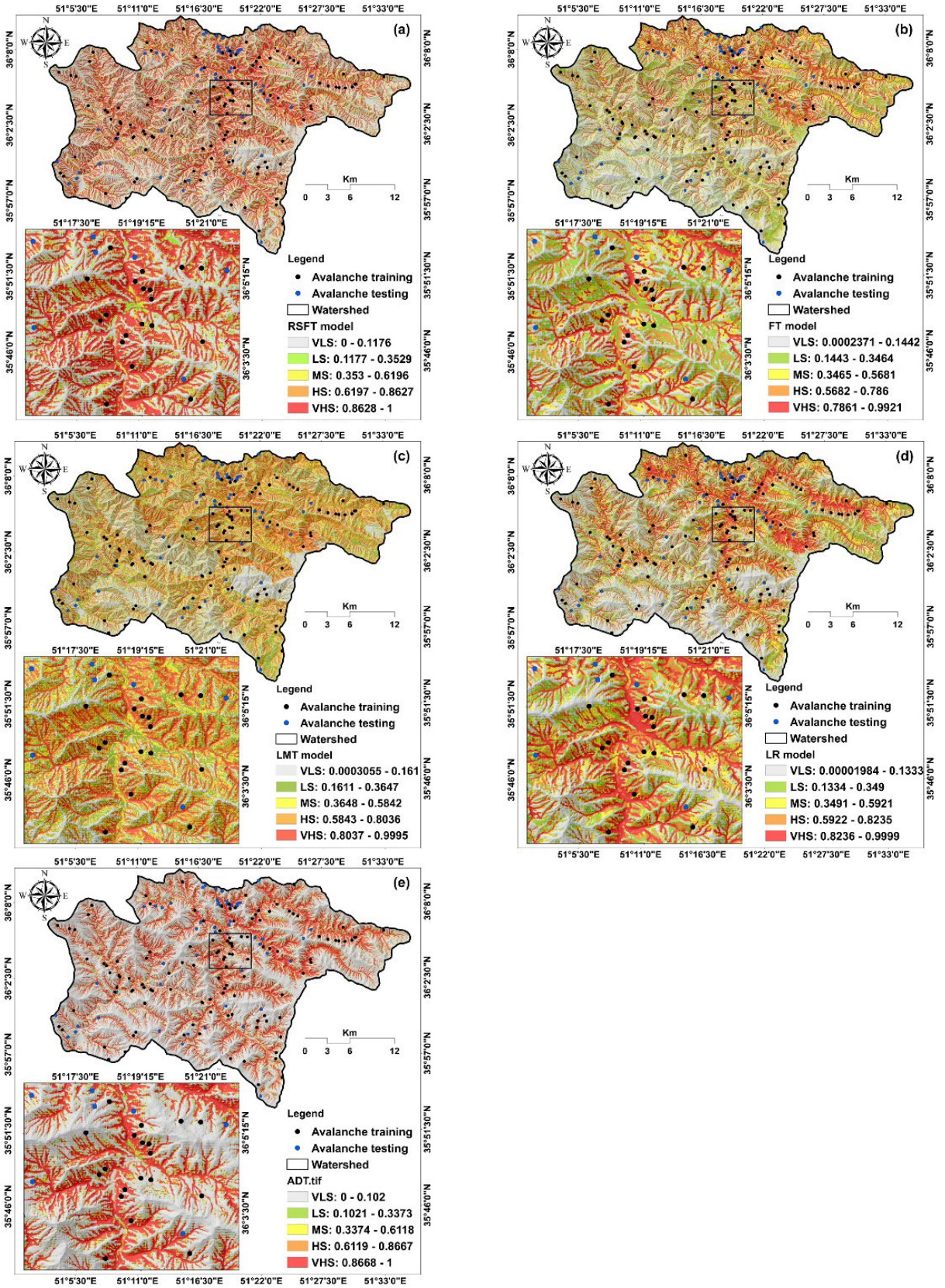

FIGURE 5. Snow avalanche susceptibility maps prepared by the models: (a) RS ensemble model, (b) FT, (c) LMT, (d) LR, and (e) ADT.

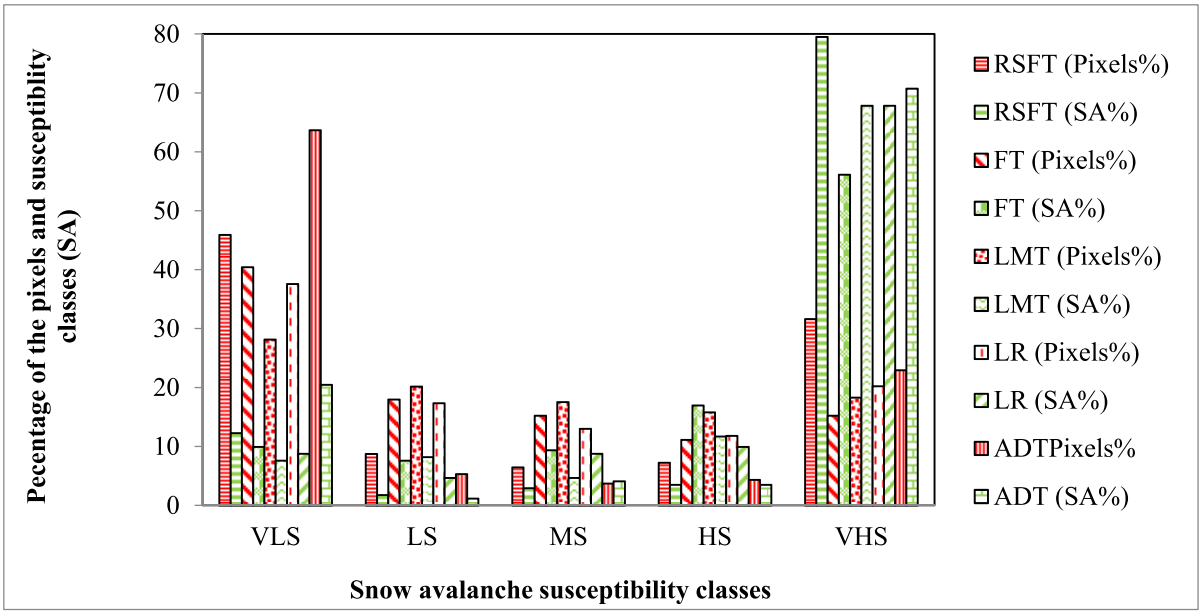

FIGURE 6. Histogram of snow avalanche susceptibility classes of all machine learning models.

on the RS based on the FT model on snow avalanche susceptibility mapping worldwide. In other words, the ability of RS for modeling of natural hazards has been earlier confirmed the abovementioned researchers. According to our results, this model can be used for other mountain areas prone to snow avalanche. However, the parameters used in the proposed model should be optimized based on the input dataset in each study area.

Compared to the previous study in this watershed, the RS model used in this study has a higher accuracy (91.3%) than the SVM and MDA models conducted by Khosravi *et al.* [39], respectively with 83% and 85% accuracy. Therefore, RS ensemble technique could well enhance the performance and prediction accuracy in snow avalanche mapping. The obtained result was for the ability of RS including: (i) in the smaller subspaces training the base classifier is easy, (ii) the substituting a single classifier with an ensemble is not adversely affected the accuracy of classification, (iii) in the random subspaces the base classifier is well enhanced than those in the original feature space [104].

D. SNOW AVALANCHE SUSCEPTIBILITY MAPPING

The RS ensemble model and all four soft computing benchmark models well trained based on the training dataset and snow avalanche susceptibility indexes (SAIs) were computed and then it was obtained and assigned for each pixel of the study area. Consequently, snow avalanche susceptibility maps were generated for each model. Although there are some classification methods to classify the SAIs in the ArcGIS such as equal interval, manual, quintile, natural breaks, geometrical interval, and standard deviation, we tested all these methods and finally we selected natural breaks as the classification method to classify the probability of snow avalanche occurrence in the study area due to more agreement and concordance it to the snow avalanche location occurred in the study area. Basically, we classified all maps into five classes that are shown in Fig. 5a-e.

With a detailed and expert look at the snow avalanches susceptibility maps (Fig. 5a-e) and matching with the conditioning factors considered in this study, it can be concluded that the HS and VHS avalanche classes are located in the north and northeast parts of the study area. These areas are mostly at higher altitudes with more rainfall and lower temperatures, which are more likely to occur. Choubin *et al.* [48] with SVM and MDA models prepares snow avalanche susceptibility maps for the used study area and they found that the HS class is mostly near the streams and matches with hillsides around the water pathways that their results confirmed also the obtained results of this study.

Fig. 6 shows the histogram of snow avalanche susceptibility classes and its coverage by pixels of the study area. According to this Figure, it can be concluded that in the RS ensemble model 45.927%, 8.737%, 6.454%, 7.238% and 31.644% of VLS, LS, MS, HS, and VHS susceptibility classes respectively are corresponded to snow avalanche of 12.281%, 1.754%, 2.924%, 3.509%, and 79.532%. In the FT model, the susceptibility classes have been occupied by the percentage of pixels of 40.442%, 17.969%, 15.222%, 11.121% and 15.247% and percentage of snow avalanche of 9.942%, 7.602%, 9.357%, 16.959%, and 56.140%. In the LMT model, VLS, LS, MS, HS, and VHS susceptibility classes have pixels of 28.172%, 20.177%, 17.543%, 15.813%, and 18.296%, respectively. However, these classes have been covered by snow avalanche of 7.602%, 8.187%, 4.678%, 11.696%, and 67.836%, respectively. In the LR model most of the area has been covered

by VLS class (37.595%), follows by VHS (20.240%), MS (13.004%), LS (17.339%), and HS (11.821%) classes. However, VHS class has covered most of snow avalanche locations (67.836%), followed by HS (9.942%), MS and VLS (7.772%), and LS (4.678%). In the ADT model 63.589%, 5.318%, 3.715%, 4.344% and 22.934% of VLS, LS, MS, HS, and VHS susceptibility classes respectively are corresponded to snow avalanche of 20.648%, 1.170%, 4.094%, 3.509% and 70.760%.

Overall results based on Fig. 5a-e and Fig. 6 show that although in all models from VLS to VHS the number of snow avalanche locations has been increased, this process is more considerable in the RS ensemble mode that 79.532% of snow avalanches are located in the VHS class.

E. VALIDATION OF SNOW AVALANCHE SUSCEPTIBILITY MAPS

We designed the ROC curve and AUC to check the performance of the RS ensemble model and four soft computing benchmark models based on training (goodness-of-fit/performance) and validation (prediction accuracy) datasets (Fig. 7a, b). The result explored that the RS reached a high performance (AUC = 0.889) and prediction accuracy (AUC = 0.867). To check its applicability, performance, and perdition accuracy, the result was compared with four machine learning benchmark models. It can be observed that the AUC based on the training dataset for the RS, LR, FT, ADT, and LMT respectively, were 0.889, 0.879, 0.878, 0.872, and 0.864 (Fig. 7a). However, it has been reported by some researchers that the results based on the validation dataset are showed the priority of the models in terms of usage [106]–[108]. Based on the validation dataset, the result concluded that the RS had the highest value of AUC (0.867), followed by LR and LMT (0.850), ADT (0.842), and FT (0.830) models. In other words, the RS ensemble model not only could well enhance the performance of the FT base classifier but also had the highest prediction accuracy compared to other machine learning benchmark models (Fig. 7b).

In addition to the ROC curve and AUC metric, we checked the applicability and statistical treatment of the RS ensemble proposed model with other benchmark models. The results based on the Friedman test showed that there was a statistical difference between all models because of significant equaled to zero at a 95% confidence level (Table 4), but it could not able to specify between which one of the models. To detect this challenge and know the statistical treatment between the models Wilcoxon signed-rank test was conducted (Table 5). Also, the result indicated a difference (rejecting the null hypothesis) between the ensemble proposed model and each compared model at a 95% confidence level because of a significance level less than 0.05 and z value exceeding the critical values ranging from −1.96 to +1.96.

Overall, it is observed that the ensemble proposed model had a higher performance and well prediction accuracy than all comparison models in the study area. One important issue for both classification or regression of machine learning methods particular for susceptibility mapping is uncertainty. The sources of the uncertainty can be related to input factors, model structure, model parameters, and etc [109],

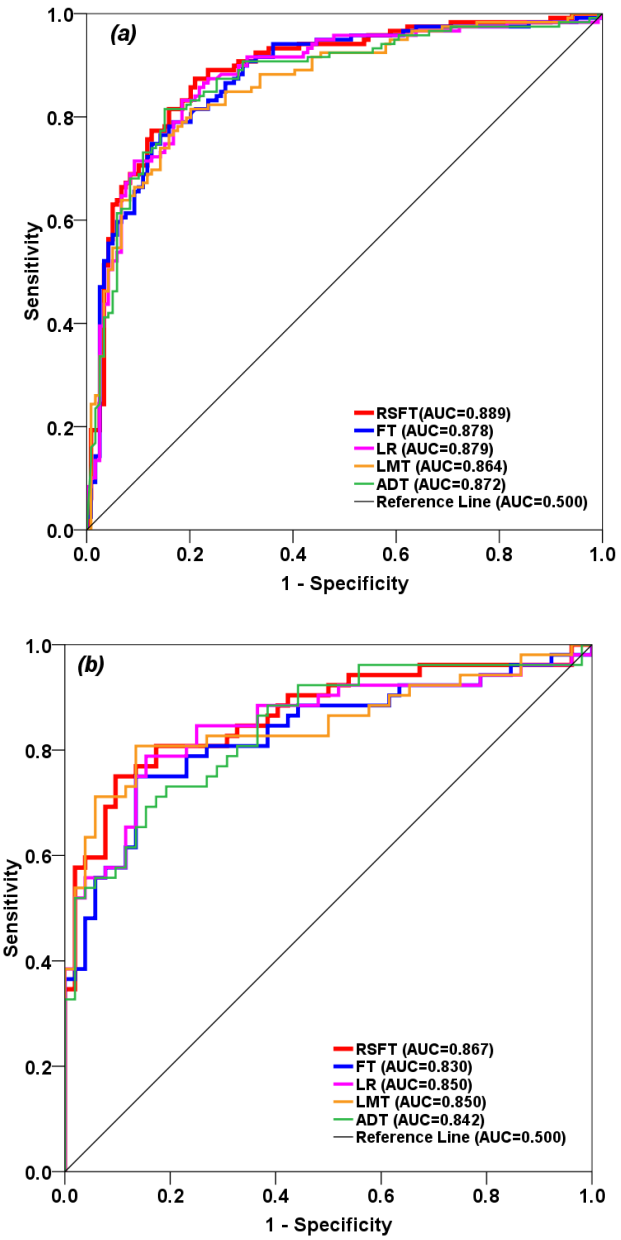

FIGURE 7. ROC curve and AUC: (a) Training (goodness-of-fit/performance), and (b) Validation (prediction accuracy) datasets.

TABLE 4. Performance evaluation of the snow avalanche susceptibility models using Friedman's test.

No	Snow avalanche models	Mean ranks	χ^2	Significance
1	RS	3.40	67.135	0.000
2	FT	2.47		
3	LR	3.12		
4	LMT	3.35		
5	ADT	2.65		

TABLE 5. Performance evaluation of the snow avalanche susceptibility models using Wilcoxon signed-rank test (two-tailed).

NO	Pairwise comparison	NPR	NNR	*Z*-value	*P*-value	Significance
1	RS vs. FT	102	136	-5.436	0.000	Yes
2	RS vs. LR	105	133	-2.135	0.031	Yes
3	RS vs. LMT	719	637	-17.758	0.000	Yes
4	RS vs. ADR	106	132	-2.993	0.032	Yes

NPR: number of positive ranks; NNR: number of negative ranks

[110]. Although we tried to reduce the uncertainties of the machine learning models through (i) best feature selection, and (ii) optimizing the parameters [95], there are some approaches such as Gaussian Process Regression (GPR) can address this by calculating the uncertainties of predicted values [111], [112]; which is recommended for future studies.

IV. CONCLUSION

The current paper presented an intelligent ensemble model (i.e., RSFT) for snow avalanche susceptibility mapping in the Karaj watershed, Iran. 21 conditioning factors tested by the IGR technique, and the ineffective factors (no predictive ability) were removed. Modeling with the most important factors (including TPI, slope, SPI, VD, DD, LS, distance to river, convexity, lithology, aspect, plan curvature, TWI, and temperature), demonstrated that the proposed RSFT model had a higher prediction accuracy in the spatial prediction of snow avalanches. Therefore, the results highlighted that the combination of the RS ensemble intelligent model and the functional tree as a base classifier provides a high quality of snow avalanche susceptibility map which is outperformed the LR, LMT, FT, and ADT models.

With a detailed and expert look at the snow avalanches susceptibility maps and matching with the conditioning factors considered in this study, it can be concluded that the high and very high susceptibility avalanche areas are located in the north and northeast parts of the study area. The study area is mostly at higher elevations with more precipitation and lower temperatures, which are more likely to occur.

The predictive capability of the RS model is depended on the inputs, and also parameters used such as the number of seeds and number of iterations not only for snow avalanche but also for other natural hazards. Optimizing the best number of inputs and the best values for parameters was the main difficulty of the research. Since these parameters reflect the uncertainty and can affect the results of the modeling process, other optimization methods should be checked.

Snow avalanche forecasting and simulation is a difficult task because of the variability of spatiotemporal of snowpack properties and the complex interactions of the snow layers. On the other hand, snow avalanches are serious threats for people, structures (buildings), and infrastructures. Therefore, the methodology developed in this study as a promising technique can help risk-based decision making in reducing the loss of lives, decreasing the damage to infrastructure caused by avalanches, and providing critical information for hazard managers.

REFERENCES

[1] A. I. Mears, "A fragment–flow model of dry–snow avalanches," *J. Glaciology*, vol. 26, no. 94, pp. 153–163, 1980, doi: 10.3189/s0022143000010698.

[2] R. Perla, T. T. Cheng, and D. M. McClung, "A two–parameter model of snow–avalanche motion," *J. Glaciology*, vol. 26, no. 94, pp. 197–207, 1980, doi: 10.3189/s002214300001073x.

[3] V. I. Osipov, N. A. Rumyantseva, and O. N. Eremina, "Living with risk of natural disasters," *Russian J. Earth Sci.*, vol. 19, no. 6, pp. 1–10, Dec. 2019, doi: 10.2205/2019ES000673.

[4] I. A. Torgoev, Y. G. Alioshin, and I. T. Aitmatov, "Danger and risk of natural and man-caused disasters in the mountains of Kyrgyzstan," *Freiberg Online Geosci.*, vol. 33, pp. 105–129, Sep. 2013. Accessed: Jun. 10, 2020. [Online]. Available: http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=14347512&asa=Y&AN=90537862&h=BKTuuWZF6kXQy16XblOAx0x3UcIhJhhnvc3NyvsU2cYgcOal-Skck6M1GbukAokhsUQ8b%2BTXvG1FwrFJuPxbYug%3D%3D&crl=c

[5] G. de Bouchard d'Aubeterre, A. Favillier, R. Mainieri, J. Lopez Saez, N. Eckert, M. Saulnier, J.-L. Peiry, M. Stoffel, and C. Corona, "Tree-ring reconstruction of snow avalanche activity: Does avalanche path selection matter?" *Sci. Total Environ.*, vol. 684, pp. 496–508, Sep. 2019, doi: 10.1016/j.scitotenv.2019.05.194.

[6] B. Gądek, R. J. Kaczka, Z. Rączkowska, E. Rojan, A. Casteller, and P. Bebi, "Snow avalanche activity in Źleb Żandarmerii in a time of climate change (Tatra Mts., Poland)," *Catena*, vol. 158, pp. 201–212, Nov. 2017, doi: 10.1016/j.catena.2017.07.005.

[7] K. Laute and A. A. Beylich, "Potential effects of climate change on future snow avalanche activity in western norway deduced from meteorological data," *Geografiska Annaler, A, Phys. Geography*, vol. 100, no. 2, pp. 163–184, Apr. 2018, doi: 10.1080/04353676.2018.1425622.

[8] M. Maggioni, D. Godone, B. Frigo, and M. Freppaz, "Snow gliding and glide-snow avalanches: Recent outcomes from two experimental test sites in Aosta Valley (Northwestern Italian Alps)," *Natural Hazards Earth Syst. Sci.*, vol. 19, no. 11, pp. 2667–2676, Nov. 2019, doi: 10.5194/nhess-19-2667-2019.

[9] J. Blahut, J. Klimeš, J. Balek, P. Hájek, L. Červená, and J. Lysák, "Snow avalanche hazard of the krkonoše national park, czech republic," *J. Maps*, vol. 13, no. 2, pp. 86–90, Nov. 2017, doi: 10.1080/17445647.2016.1262794.

[10] A. Rakesh and M. C. Sharma, "Snow avalanche hazard vulnerability analysis of Himachal Pradesh," *Disaster Adv.*, vol. 12, no. 10, pp. 34–42, Oct. 2019.

[11] D. K. Singh, V. D. Mishra, H. S. Gusain, N. Gupta, and A. K. Singh, "Geo-spatial modeling for automated demarcation of snow avalanche hazard areas using Landsat-8 satellite images and *in situ* data," *J. Indian Soc. Remote Sens.*, vol. 47, no. 3, pp. 513–526, Mar. 2019, doi: 10.1007/s12524-018-00936-w.

[12] A. Casteller, T. Häfelfinger, E. Cortés Donoso, K. Podvin, D. Kulakowski, and P. Bebi, "Assessing the interaction between mountain forests and snow avalanches at nevados de Chillán, chile and its implications for ecosystem-based disaster risk reduction," *Natural Hazards Earth Syst. Sci.*, vol. 18, no. 4, pp. 1173–1186, Apr. 2018, doi: 10.5194/nhess-18-1173-2018.

[13] C. García-Hernández, J. Ruiz-Fernández, and B. González-Díaz, "Inherited memory, social learning, and resilience: Lessons from Spain's great blizzard of 1888," *Geographical Res.*, vol. 57, no. 2, pp. 189–203, May 2019, doi: 10.1111/1745-5871.12322.

[14] D. Germain, "Snow avalanche hazard assessment and risk management in northern quebec, eastern canada," *Natural Hazards*, vol. 80, no. 2, pp. 1303–1321, Jan. 2016, doi: 10.1007/s11069-015-2024-z.

[15] S. Beato Bergua, M. Á. P. Piedrabuena, and J. L. M. Alfonso, "Snow avalanche susceptibility in the eastern hillside of the aramo range (Asturian central massif, cantabrian mountains, NW Spain)," *J. Maps*, vol. 14, no. 2, pp. 373–381, Nov. 2018, doi: 10.1080/17445647.2018.1480974.

[16] T. Sardar, A. Raziq, A. Rashid, and G. Saddiq. (2019). *Snow Avalanche Based Susceptibility Assessment of Selected Districts in Northern Zone of Pakistan Applying MCDA Approach in GIS*. Accessed: Jun. 10, 2020. [Online]. Available: http://nceg.uop.edu.pk/GeologicalBulletin/Vol-52(2)-2019/Vol-52(2)-2019-Paper6.pdf

[17] R. L. Soteres, J. Pedraza, and R. M. Carrasco, "Snow avalanche susceptibility of the circo de gredos (Iberian central system, Spain)," *J. Maps*, vol. 16, no. 2, pp. 155–165, Dec. 2020, doi: 10.1080/17445647.2020.1717655.

[18] A. Chourot and J.-P. Martin, "Comparison of logistic regressions and snowfall intensity–duration threshold as forecasting tools for direct-action snow avalanches in the presidential range, new hampshire, USA," *Natural Hazards*, vol. 93, no. 3, pp. 1649–1656, Sep. 2018, doi: 10.1007/s11069-018-3361-5.

[19] L. Dreier, S. Harvey, A. van Herwijnen, and C. Mitterer, "Relating meteorological parameters to glide-snow avalanche activity," *Cold Regions Sci. Technol.*, vol. 128, pp. 57–68, Aug. 2016, doi: 10.1016/j.coldregions.2016.05.003.

[20] H. Hancock, A. Prokop, M. Eckerstorfer, and J. Hendrikx, "Combining high spatial resolution snow mapping and meteorological analyses to improve forecasting of destructive avalanches in Longyearbyen, Svalbard," *Cold Regions Sci. Technol.*, vol. 154, pp. 120–132, Oct. 2018, doi: 10.1016/j.coldregions.2018.05.011.

[21] M. Heck, C. Hammer, A. van Herwijnen, J. Schweizer, and D. Fäh, "Automatic detection of snow avalanches in continuous seismic data using hidden Markov models," *Natural Hazards Earth Syst. Sci.*, vol. 18, no. 1, pp. 383–396, Jan. 2018, doi: 10.5194/nhess-18-383-2018.

[22] A. Van Herwijnen, J. Gaume, E. H. Bair, B. Reuter, K. W. Birkeland, and J. Schweizer, "Estimating the effective elastic modulus and specific fracture energy of snowpack layers from field experiments," *J. Glaciology*, vol. 62, no. 236, pp. 997–1007, Dec. 2016, doi: 10.1017/jog.2016.90.

[23] N. Wever, C. Vera Valero, and F. Techel, "Coupled snow cover and avalanche dynamics simulations to evaluate wet snow avalanche activity," *J. Geophys. Res., Earth Surf.*, vol. 123, no. 8, pp. 1772–1796, Aug. 2018, doi: 10.1029/2017JF004515.

[24] Y. Bühler, D. von Rickenbach, A. Stoffel, S. Margreth, L. Stoffel, and M. Christen, "Automated snow avalanche release area delineation–validation of existing algorithms and proposition of a new object-based approach for large-scale hazard indication mapping," *Natural Hazards Earth Syst. Sci.*, vol. 18, no. 12, pp. 3235–3251, Dec. 2018, doi: 10.5194/nhess-18-3235-2018.

[25] S. Kumar, P. K. Srivastava, A. Snehmani, and S. Bhatiya, "Geospatial probabilistic modelling for release area mapping of snow avalanches," *Cold Regions Sci. Technol.*, vol. 165, Sep. 2019, Art. no. 102813, doi: 10.1016/j.coldregions.2019.102813.

[26] F. Brigo, G. Strapazzon, W. M. Otte, S. C. Igwe, and H. Brugger, "Web search behavior for snow avalanches: An italian study," *Natural Hazards*, vol. 80, no. 1, pp. 141–152, Jan. 2016, doi: 10.1007/s11069-015-1961-x.

[27] S. Kumar, P. K. Srivastava, and A. Snehmani, "Geospatial modelling and mapping of snow avalanche susceptibility," *J. Indian Soc. Remote Sens.*, vol. 46, no. 1, pp. 109–119, Jan. 2018, doi: 10.1007/s12524-017-0672-z.

[28] M. P. Piedrabuena, S. B. Bergua, and J. L. M. Alfonso, "The snow avalanches risk in the upper san isidro valley (Asturian central massif): The vulnerability of the as-253 road," *Cuaternario y Geomorfol.*, vol. 33, nos. 3–4, pp. 79–104, Dec. 2019, doi: 10.17735/cyg.v33i3-4.72057.

[29] T. Arumugam, S. K. Dewali, A. Snehmani, and A. Ganju, "Integration of terrain and AVHRR-derived multi-temporal snow cover data for statistical assessment of avalanches: Case study of a part of NW himalaya," *Arabian J. Geosci.*, vol. 12, no. 17, Sep. 2019, doi: 10.1007/s12517-019-4691-7.

[30] M. Eckerstorfer, H. Vickers, E. Malnes, and J. Grahn, "Near-real time automatic snow avalanche activity monitoring system using Sentinel-1 SAR data in norway," *Remote Sens.*, vol. 11, no. 23, p. 2863, Dec. 2019, doi: 10.3390/rs11232863.

[31] H. S. Gusain, V. D. Mishra, and D. K. Singh, "Study of a snow avalanche accident along Chowkibal–Tangdhar road in Kupwara district, Jammu and Kashmir, India," *Current Sci.*, vol. 115, no. 5, pp. 969–972, 2018, doi: 10.18520/cs/v115/i5/969-972.

[32] M. Voiculescu and A. Onaca, "Spatio-temporal reconstruction of snow avalanche activity using dendrogeomorphological approach in Bucegi mountains romanian carpathians," *Cold Regions Sci. Technol.*, vols. 104–105, pp. 63–75, Aug. 2014, doi: 10.1016/j.coldregions.2014.04.005.

[33] A. Aydin and R. Eker, "GIS-based snow avalanche hazard mapping: Bayburt-Asagi dere catchment case," *J. Environ. Biol.*, vol. 38, no. 5, pp. 937–943, 2017, doi: 10.22438/jeb/38/5(SI)/GM-10.

[34] M. Boltižiar, M. Biskupič, and I. Barka, "Spatial modelling of avalanches by application of GIS on selected slopes of the western tatra Mts. And belianske tatra Mts., slovakia," *Geographia Polonica*, vol. 89, no. 1, pp. 79–90, 2016, doi: 10.7163/GPol.0047.

[35] E. H. Peitzsch, J. Hendrikx, and D. B. Fagre, "Terrain parameters of glide snow avalanches and a simple spatial glide snow avalanche model," *Cold Regions Sci. Technol.*, vol. 120, pp. 237–250, Dec. 2015, doi: 10.1016/j.coldregions.2015.08.002.

[36] S. Salcedo-Sanz, P. Ghamisi, M. Piles, M. Werner, L. Cuadra, A. Moreno-Martínez, E. Izquierdo-Verdiguier, J. Muñoz-Marí, A. Mosavi, and G. Camps-Valls, "Machine learning information fusion in Earth observation: A comprehensive review of methods, applications and data sources," *Inf. Fusion*, vol. 63, pp. 256–272, Nov. 2020. Accessed: Jul. 18, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253520303171.

[37] Z. M. Yaseen, W. H. M. W. Mohtar, A. M. S. Ameen, I. Ebtehaj, S. F. M. Razali, H. Bonakdari, S. Q. Salih, N. Al-Ansari, and S. Shahid, "Implementation of univariate paradigm for streamflow simulation using hybrid data-driven model: Case study in tropical region," *IEEE Access*, vol. 7, pp. 74471–74481, 2019, doi: 10.1109/ACCESS.2019.2920916.

[38] Z. Yaseen, I. Ebtehaj, S. Kim, H. Sanikhani, H. Asadi, M. Ghareb, H. Bonakdari, W. Wan Mohtar, N. Al-Ansari, and S. Shahid, "Novel hybrid data-intelligence model for forecasting monthly rainfall with uncertainty analysis," *Water*, vol. 11, no. 3, p. 502, Mar. 2019, doi: 10.3390/w11030502.

[39] K. Khosravi, P. Daggupati, M. T. Alami, S. M. Awadh, M. I. Ghareb, M. Panahi, B. T. Pham, F. Rezaie, C. Qi, and Z. M. Yaseen, "Meteorological data mining and hybrid data-intelligence models for reference evaporation simulation: A case study in Iraq," *Comput. Electron. Agricult.*, vol. 167, Dec. 2019, Art. no. 105041, doi: 10.1016/j.compag.2019.105041.

[40] S. Q. Salih, A. Sharafati, K. Khosravi, H. Faris, O. Kisi, H. Tao, M. Ali, and Z. M. Yaseen, "River suspended sediment load prediction based on river discharge information: Application of newly developed data mining models," *Hydrological Sci. J.*, vol. 65, no. 4, pp. 624–637, Mar. 2020, doi: 10.1080/02626667.2019.1703186.

[41] X. Tang, K. Liu, X. Wang, F. Gao, J. Macro, and W. D. Widanage, "Model migration neural network for predicting battery aging trajectories," *IEEE Trans. Transport. Electrific.*, vol. 6, no. 2, pp. 363–374, Jun. 2020, doi: 10.1109/TTE.2020.2979547.

[42] K. Liu, Y. Shang, Q. Ouyang, and W. D. Widanage, "A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery," *IEEE Trans. Ind. Electron.*, early access, Mar. 18, 2020, doi: 10.1109/tie.2020.2973876.

[43] S. Maroufpoor, E. Maroufpoor, O. Bozorg-Haddad, J. Shiri, and Z. M. Yaseen, "Soil moisture simulation using hybrid artificial intelligent model: Hybridization of adaptive neuro fuzzy inference system with grey wolf optimizer algorithm," *J. Hydrol.*, vol. 575, pp. 544–556, Aug. 2019, doi: 10.1016/j.jhydrol.2019.05.045.

[44] M. Emadi, R. Taghizadeh-Mehrjardi, A. Cherati, M. Danesh, A. Mosavi, and T. Scholten, "Predicting and mapping of soil organic carbon using machine learning algorithms in Northern Iran," *Remote Sens.*, vol. 12, no. 14, p. 2234, Jul. 2020, doi: 10.3390/rs12142234.

[45] H. R. Pourghasemi and M. Rossi, "Landslide susceptibility modeling in a landslide prone area in Mazandarn Province, North of Iran: A comparison between GLM, GAM, MARS, and M-AHP methods," *Theor. Appl. Climatol.*, vol. 130, nos. 1–2, pp. 609–633, Oct. 2017, doi: 10.1007/s00704-016-1919-2.

[46] E. Dodangeh, B. Choubin, A. N. Eigdir, N. Nabipour, M. Panahi, S. Shamshirband, and A. Mosavi, "Integrated machine learning methods with resampling algorithms for flood susceptibility prediction," *Sci. Total Environ.*, vol. 705, Feb. 2020, Art. no. 135983, doi: 10.1016/j.scitotenv.2019.135983.

[47] T. Thüring, M. Schoch, A. van Herwijnen, and J. Schweizer, "Robust snow avalanche detection using supervised machine learning with infrasonic sensor arrays," *Cold Regions Sci. Technol.*, vol. 111, pp. 60–66, Mar. 2015, doi: 10.1016/j.coldregions.2014.12.014.

[48] B. Choubin, M. Borji, A. Mosavi, F. Sajedi-Hosseini, V. P. Singh, and S. Shamshirband, "Snow avalanche hazard prediction using machine learning methods," *J. Hydrol.*, vol. 577, Oct. 2019, Art. no. 123929, doi: 10.1016/j.jhydrol.2019.123929.

[49] O. Rahmati, O. Ghorbanzadeh, T. Teimurian, F. Mohammadi, J. P. Tiefenbacher, F. Falah, S. Pirasteh, P.-T.-T. Ngo, and D. T. Bui, "Spatial modeling of snow avalanche using machine learning models and geo-environmental factors: Comparison of effectiveness in two mountain regions," *Remote Sens.*, vol. 11, no. 24, p. 2995, Dec. 2019, doi: 10.3390/rs11242995.

[50] M. Sakizadeh, "Assessment the performance of classification methods in water quality studies, a case study in Karaj River," *Environ. Monitor. Assessment*, vol. 187, no. 9, pp. 1–12, Sep. 2015, doi: 10.1007/s10661-015-4761-6.

[51] S. F. Tavassol and G. Gs. (2016). *GIS- Based Morphometric Analysis of Major Watersheds of Tehran-Karaj, Central of Iran*. Accessed: Jun. 10, 2020. [Online]. Available: http://www.ijlsci.in/

[52] M. Naaim, F. Naaim-Bouvet, T. Faug, and A. Bouchet, "Dense snow avalanche modeling: Flow, erosion, deposition and obstacle effects," *Cold Regions Sci. Technol.*, vol. 39, nos. 2–3, pp. 193–204, Oct. 2004, doi: 10.1016/j.coldregions.2004.07.001.

[53] D. M. Delparte, "Avalanche terrain modeling in Glacier National Park, Canada," Univ. Calgary, Calgary, AB, Canada, Tech. Rep., 2008.

[54] M. J. Smith and D. M. Mcclung, "Avalanche frequency and terrain characteristics at Rogers' Pass, British Columbia, Canada," *J. Glaciol.*, vol. 43, no. 143, pp. 165–171, 1997, doi: 10.1017/S0022143000002926.

[55] J. Schweizer, "Snow avalanche formation," *Rev. Geophys.*, vol. 41, no. 4, pp. 122–143, 2003, doi: 10.1029/2002RG000123.

[56] C. Jaccard, "B. Armstrong, and K. Williams 1986. The avalanche book. Golden, CO, fulcrum Inc.," *J. Glaciology*, vol. 33, no. 113, p. 129, 1987, doi: 10.3189/s002214300000544x.

[57] L. Bakermans, "Near-surface snow temperature changes over terrain," Montana State Univ., Bozeman, MT, USA, Tech. Rep., 2006.

[58] J. P. Gallant and J. C. Wilson, "Primary topographic attributes," in *Terrain Analysis: Principles and Applications*, J. P. Wilson and J. C. Gallant, Eds. Amsterdam, The Netherlands: Elsevier, 2000, pp. 51–85.

[59] J. De Reu, J. Bourgeois, M. Bats, A. Zwertvaegher, V. Gelorini, P. De Smedt, W. Chu, M. Antrop, P. De Maeyer, P. Finke, M. Van Meirvenne, J. Verniers, and P. Crombéa, "Application of the topographic position index to heterogeneous landscapes," *Geomorphology*, vol. 186, pp. 39–49, Mar. 2013, doi: 10.1016/j.geomorph.2012.12.015.

[60] J. M. Sappington, K. M. Longshore, and D. B. Thompson, "Quantifying landscape ruggedness for animal habitat analysis: A case study using bighorn sheep in the mojave desert," *J. Wildlife Manage.*, vol. 71, no. 5, pp. 1419–1426, Jul. 2007, doi: 10.2193/2005-723.

[61] R. D. Hobson, "Surface roughness in topography: Quantitative approach," *Spat. Anal. Geomorphol.*, vol. 6, no. 3, pp. 221–245, 2019, doi: 10.4324/9780429273346-8.

[62] J. Veitinger, B. Sovilla, and R. S. Purves, "Influence of snow depth distribution on surface roughness in alpine terrain: A multi-scale approach," *Cryosphere*, vol. 8, no. 2, pp. 547–569, Apr. 2014, doi: 10.5194/tc-8-547-2014.

[63] S. Riley, "Index that quantifies topographic heterogeneity," *Intermountain J. Sci.*, vol. 5, nos. 1–4, pp. 23–27, 1999. Accessed: Jun. 10, 2020. [Online]. Available: http://download.osgeo.org/qgis/doc/reference-docs/Terrain_Ruggedness_Index.pdf

[64] D. D. Moghaddam, M. Rezaei, H. R. Pourghasemi, Z. S. Pourtaghie, and B. Pradhan, "Groundwater spring potential mapping using bivariate statistical model and GIS in the Taleghan Watershed, Iran," *Arabian J. Geosci.*, vol. 8, no. 2, pp. 913–929, Feb. 2015, doi: 10.1007/s12517-013-1161-5.

[65] S. M. Manson, P. A. Burrough, and R. A. McDonnell, "Principles of geographical information systems: Spatial information systems and geostatistics," *Econ. Geography*, vol. 75, no. 4, p. 422, Oct. 1999, doi: 10.2307/144481.

[66] J.-T. Fischer, J. Kowalski, and S. P. Pudasaini, "Topographic curvature effects in applied avalanche modeling," *Cold Regions Sci. Technol.*, vols. 74–75, pp. 21–30, May 2012, doi: 10.1016/j.coldregions.2012.01.005.

[67] A. van Herwijnen, M. Heck, and J. Schweizer, "Forecasting snow avalanches using avalanche activity data obtained through seismic monitoring," *Cold Regions Sci. Technol.*, vol. 132, pp. 68–80, Dec. 2016, doi: 10.1016/j.coldregions.2016.09.014.

[68] K. W. Birkeland, "Terminology and predominant processes associated with the formation of weak layers of near-surface faceted crystals in the mountain snowpack," *Arctic Alpine Res.*, vol. 30, no. 2, pp. 193–199, 1998, doi: 10.2307/1552134.

[69] C.-Y. Chen and F.-C. Yu, "Morphometric analysis of debris flows and their source areas using GIS," *Geomorphology*, vol. 129, nos. 3–4, pp. 387–397, Jun. 2011, doi: 10.1016/j.geomorph.2011.03.002.

[70] H. Pourghasemi, B. Pradhan, C. Gokceoglu, and K. D. Moezzi, "A comparative assessment of prediction capabilities of Dempster–Shafer and weights-of-evidence models in landslide susceptibility mapping using GIS," *Geomatics, Natural Hazards Risk*, vol. 4, no. 2, pp. 93–118, Jun. 2013, doi: 10.1080/19475705.2012.662915.

[71] H. C. Jessen and S. Menard, "Applied logistic regression Analysis.," *Statistician*, vol. 45, no. 4, p. 534, 1996, doi: 10.2307/2988559.

[72] S. Menard, "Applied logistic regression analysis: Sage University series on quantitative applications in the social sciences," Sage Publ. Inc, Sage Univ., Tech. Rep., 2002.

[73] B. Choubin, A. Mosavi, E. H. Alamdarloo, F. S. Hosseini, S. Shamshirband, K. Dashtekian, and P. Ghamisi, "Earth fissure hazard prediction using machine learning models," *Environ. Res.*, vol. 179, Dec. 2019, Art. no. 108770, doi: 10.1016/j.envres.2019.108770.

[74] F. S. Hosseini, B. Choubin, A. Mosavi, N. Nabipour, S. Shamshirband, H. Darabi, and A. T. Haghighi, "Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: Application of the simulated annealing feature selection method," *Sci. Total Environ.*, vol. 711, Apr. 2020, Art. no. 135161, doi: 10.1016/j.scitotenv.2019.135161.

[75] B. Choubin, M. Abdolshahnejad, E. Moradi, X. Querol, A. Mosavi, S. Shamshirband, and P. Ghamisi, "Spatial hazard assessment of the PM10 using machine learning models in Barcelona, Spain," *Sci. Total Environ.*, vol. 701, Jan. 2020, Art. no. 134474, doi: 10.1016/j.scitotenv.2019.134474.

[76] J. R. Quinlan, "Induction of Decision Trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986, doi: 10.1023/A:1022643204877.

[77] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith. (2007). *Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations*. Accessed: Jul. 7, 2020. [Online]. Available: https://elibrary.asabe.org/abstract.asp?aid=23153

[78] F. Gauthier, D. Germain, and B. Hétu, "Logistic models as a forecasting tool for snow avalanches in a cold maritime climate: Northern Gaspésie, Québec, Canada," *Natural Hazards*, vol. 89, no. 1, pp. 201–232, Oct. 2017, doi: 10.1007/s11069-017-2959-3.

[79] V. Jomelli, C. Delval, D. Grancher, S. Escande, D. Brunstein, B. Hetu, L. Filion, and P. Pechc, "Probabilistic analysis of recent snow avalanche activity and weather in the French Alps," *Cold Reg. Sci. Technol.*, vol. 47, nos. 1–2, pp. 180–192, 2007, doi: 10.1016/j.coldregions.2006.08.003.

[80] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Mach. Learn.*, vol. 59, nos. 1–2, pp. 161–205, May 2005, doi: 10.1007/s10994-005-0466-3.

[81] M. Sumner, E. Frank, and M. Hall, "Speeding up logistic model tree induction," in *Knowledge Discovery in Databases: PKDD* (Lecture Notes in Computer Science), vol. 3721. Springer, 2005, pp. 675–683, doi: 10.1007/11564126_72.

[82] Y. Freund and L. Mason, "The alternating decision tree learning algorithm," in *Proc. ICML*, vol. 99, 1999, pp. 124–133, doi: 10.1093/jxb/ern164.

[83] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*. Hoboken, NJ, USA: Wiley, 2016.

[84] H. Hong, B. Pradhan, C. Xu, and D. Tien Bui, "Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines," *Catena*, vol. 133, pp. 266–281, Oct. 2015, doi: 10.1016/j.catena.2015.05.019.

[85] J. Gama, "Functional trees," *Mach. Learn.*, vol. 55, no. 3, pp. 219–250, Jun. 2004, doi: 10.1023/B:MACH.0000027782.67192.13.

[86] T. Kam Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, 1998, doi: 10.1109/34.709601.

[87] M. Skurichina and R. P. W. Duin, "Bagging, boosting and the random subspace method for linear classifiers," *Pattern Anal. Appl.*, vol. 5, no. 2, pp. 121–135, Jun. 2002, doi: 10.1007/s100440200011.

[88] T. Lasota, T. Łuczak, M. Niemczyk, M. Olszewski, and B. Trawiński, "Investigation of property valuation models based on decision tree ensembles built over noised data," in *Computational Collective Intelligence. Technologies and Applications* (Lecture Notes in Computer Science). Springer, 2013, doi: 10.1007/978-3-642-40495-5_42.

[89] Y. Bühler, M. Christen, J. Kowalski, and P. Bartelt, "Sensitivity of snow avalanche simulations to digital elevation model quality and resolution," *Ann. Glaciology*, vol. 52, no. 58, pp. 72–80, 2011, doi: 10.3189/172756411797252121.

[90] A. Ozdemir, "Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the sultan mountains (Aksehir, Turkey)," *J. Hydrol.*, vol. 405, nos. 1–2, pp. 123–136, Jul. 2011, doi: 10.1016/j.jhydrol.2011.05.015.

[91] Y. Razandi, H. R. Pourghasemi, N. S. Neisani, and O. Rahmati, "Application of analytical hierarchy process, frequency ratio, and certainty factor models for groundwater potential mapping using GIS," *Earth Sci. Informat.*, vol. 8, no. 4, pp. 867–883, Dec. 2015, doi: 10.1007/s12145-015-0220-8.

[92] M. Friedman, "A comparison of alternative tests of significance for the problem of *m* rankings," *Ann. Math. Statist.*, vol. 11, no. 1, pp. 86–92, Mar. 1940, doi: 10.1214/aoms/1177731944.

[93] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, vol. 1, no. 6, p. 80, Dec. 1945, doi: 10.2307/3001968.

[94] A. Shirzadi, K. Soliamani, M. Habibnejhad, A. Kavian, K. Chapi, H. Shahabi, W. Chen, K. Khosravi, B. Thai Pham, B. Pradhan, A. Ahmad, B. Bin Ahmad, and D. Tien Bui, "Novel GIS based machine learning algorithms for shallow landslide susceptibility mapping," *Sensors*, vol. 18, no. 11, p. 3777, Nov. 2018, doi: 10.3390/s18113777.

[95] A. Shirzadi, K. Solaimani, M. H. Roshan, A. Kavian, K. Chapi, H. Shahabi, S. Keesstra, B. B. Ahmad, and D. T. Bui, "Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution," *Catena*, vol. 178, pp. 172–188, Jul. 2019, doi: 10.1016/j.catena.2019.03.017.

[96] V.-H. Nhu, A. Shirzadi, H. Shahabi, W. Chen, J. J. Clague, M. Geertsema, A. Jaafari, M. Avand, S. Miraki, D. T. Asl, B. T. Pham, B. B. Ahmad, and S. Lee, "Shallow landslide susceptibility mapping by random forest base classifier and its ensembles in a semi-arid region of Iran," *Forests*, vol. 11, no. 4, p. 421, Apr. 2020, doi: 10.3390/F11040421.

[97] V.-H. Nhu, A. Shirzadi, H. Shahabi, S. K. Singh, N. Al-Ansari, J. J. Clague, A. Jaafari, W. Chen, S. Miraki, J. Dou, C. Luu, K. Górski, B. Thai Pham, H. D. Nguyen, and B. B. Ahmad, "Shallow landslide susceptibility mapping: A comparison between logistic model tree, logistic regression, Naïve Bayes tree, artificial neural network, and support vector machine algorithms," *Int. J. Environ. Res. Public Health*, vol. 17, no. 8, p. 2749, Apr. 2020, doi: 10.3390/ijerph17082749.

[98] A. Azareh, E. Rafiei Sardooi, B. Choubin, S. Barkhori, A. Shahdadi, J. Adamowski, and S. Shamshirband, "Incorporating multi-criteria decision-making and fuzzy-value functions for flood susceptibility assessment," *Geocarto Int.*, vol. 11, pp. 1–21, Nov. 2019, doi: 10.1080/10106049.2019.1695958.

[99] M. Dekanova, F. Duchon, M. Dekan, F. Kyzek, and M. Biskupic, "Avalanche forecasting using neural network," in *Proc. Elektro*, May 2018, pp. 1–5, doi: 10.1109/ELEKTRO.2018.8398359.

[100] D. McClung and P. A. Schaerer, *The Avalanche Handbook*, vol. 31, no. 7. Hoboken, NJ, USA: Wiley, 1994.

[101] A. Singh and A. Ganju. (2008). *Artificial Neural Networks for Snow Avalanche Forecasting in Indian Himalaya*. Accessed: Jun. 10, 2020. [Online]. Available: https://www.researchgate.net/publication/263651736

[102] A. Shirzadi, D. T. Bui, B. T. Pham, K. Solaimani, K. Chapi, A. Kavian, H. Shahabi, and I. Revhaug, "Shallow landslide susceptibility assessment using a novel hybrid intelligence approach," *Environ. Earth Sci.*, vol. 76, no. 2, Jan. 2017, Art. no. 60, doi: 10.1007/s12665-016-6374-y.

[103] M. Abedini, B. Ghasemian, A. Shirzadi, H. Shahabi, K. Chapi, B. T. Pham, B. B. Ahmad, and D. T. Bui, "A novel hybrid approach of Bayesian logistic regression and its ensembles for landslide susceptibility assessment," *Geocarto Int.*, vol. 34, no. 13, pp. 1427–1457, Nov. 2019, doi: 10.1080/10106049.2018.1499820.

[104] B. T. Pham, I. Prakash, S. K. Singh, A. Shirzadi, H. Shahabi, T.-T.-T. Tran, and D. T. Bui, "Landslide susceptibility modeling using reduced error pruning trees and different ensemble techniques: Hybrid machine learning approaches," *Catena*, vol. 175, pp. 203–218, Apr. 2019, doi: 10.1016/j.catena.2018.12.018.

[105] S. Miraki, S. H. Zanganeh, K. Chapi, V. P. Singh, A. Shirzadi, H. Shahabi, and B. T. Pham, "Mapping groundwater potential using a novel hybrid intelligence approach," *Water Resour. Manage.*, vol. 33, no. 1, pp. 281–302, Jan. 2019, doi: 10.1007/s11269-018-2102-6.

[106] Y. Wang, H. Hong, W. Chen, S. Li, M. Panahi, K. Khosravi, A. Shirzadi, H. Shahabi, S. Panahi, and R. Costache, "Flood susceptibility mapping in dingnan county (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm," *J. Environ. Manage.*, vol. 247, pp. 712–729, Oct. 2019, doi: 10.1016/j.jenvman.2019.06.102.

[107] M. Ahmadlou, M. Karimi, S. Alizadeh, A. Shirzadi, D. Parvinnejhad, H. Shahabi, and M. Panahi, "Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA)," *Geocarto Int.*, vol. 34, no. 11, pp. 1252–1272, Sep. 2019, doi: 10.1080/10106049.2018.1474276.

[108] K. Khosravi, H. Shahabi, B. T. Pham, J. Adamowski, A. Shirzadi, B. Pradhan, J. Dou, H.-B. Ly, G. Gróf, H. L. Ho, H. Hong, K. Chapi, and I. Prakash, "A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods," *J. Hydrol.*, vol. 573, pp. 311–323, Jun. 2019, doi: 10.1016/j.jhydrol.2019.03.073.

[109] D. P. Solomatine and D. L. Shrestha, "A novel method to estimate model uncertainty using machine learning techniques," *Water Resour. Res.*, vol. 45, no. 12, Dec. 2009, doi: 10.1029/2008WR006839.

[110] O. Rahmati, B. Choubin, A. Fathabadi, F. Coulon, E. Soltani, H. Shahabi, E. Mollaefar, J. Tiefenbacher, S. Cipullo, B. B. Ahmad, and D. Tien Bui, "Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods," *Sci. Total Environ.*, vol. 688, pp. 855–866, Oct. 2019, doi: 10.1016/j.scitotenv.2019.06.320.

[111] K. Liu, X. Hu, Z. Wei, Y. Li, and Y. Jiang, "Modified Gaussian process regression models for cyclic capacity prediction of lithium-ion batteries," *IEEE Trans. Transport. Electrific.*, vol. 5, no. 4, pp. 1225–1236, Dec. 2019, doi: 10.1109/TTE.2019.2944802.

[112] K. Liu, Y. Li, X. Hu, M. Lucu, and W. D. Widanage, "Gaussian process regression with automatic relevance determination kernel for calendar aging prediction of lithium-ion batteries," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 3767–3777, Jun. 2020, doi: 10.1109/TII.2019.2941747.

AMIRHOSEIN MOSAVI graduated from Kingston University London, U.K. He received the Ph.D. degree in applied informatics. He is currently a Data Scientist in climate change, sustainability, and hazard prediction. He is a Senior Research Fellow with Oxford Brookes University. He is also an Alexander von Humboldt Research Fellow in big data, the IoT, and machine learning. He was a recipient of the Green-Talent Award, the UNESCO Young Scientist Award, the ERCIM Alain Bensoussan Fellowship Award, the Campus France Fellowship Award, the Campus Hungary Fellowship Award, and the Endeavour-Australia Leadership.

ATAOLLAH SHIRZADI received the M.Sc. and Ph.D. degrees in watershed management engineering and sciences from Sari Agricultural Sciences and Natural Resources University, Mazandaran, Iran, in 2007 and 2017, respectively. He is currently a Research Assistant. He is also with the University of Kurdistan, Kurdistan, Iran. He is the author or coauthor of 60 articles published in international previewed (JCR-indexed) and refereed journals indexed, such as SCI, SCIE, and SCOPUS. He is a Scientific Reviewer of more than 15 ISI journals. His current research interests include development of artificial intelligence approaches, such as machine learning, deep learning, uncertainty concept in modeling process, and optimization algorithms on natural hazards phenomena, such as landslides, flash flood, soil erosion, and other environmental challenges.

BAHRAM CHOUBIN received the B.S. degree in rangeland and watershed management from Shiraz University, Shiraz, Iran, in 2011, and the M.S. degree in watershed management and engineering from the University of Tehran, Tehran, Iran, in 2013, and the Ph.D. degree in watershed management from Sari Agricultural Sciences and Natural Resources University, Sari, Iran, in 2017. From 2018 to 2019, he was a Postdoctoral Researcher with the University of Tehran. He was also an Assistant Professor with the Soil Conservation and Watershed Management Research Department, AREEO, Urmia, Iran, in 2019. He is a Scientific Reviewer of more than 50 ISI journals. He has published more than 30 articles in JCR-indexed journals. His research interests include hydrometeorology, natural hazard, predictions in ungauged basins (PUB), remote sensing, and hazard modeling. He was a recipient of the Academic Award from the National Elites Foundation, Iran, in 2016.

FERESHTEH TAROMIDEH received the B.S. degree in water engineering from the University of Guilan, Guilan, Iran, in 2013, and the M.S. degree in water engineering from Sari Agricultural Sciences and Natural Resources University, Sari, Iran, in 2016. Her research interests include water resources management, hydrology, river hydraulics, sediment transport, modeling and measurements of flows, and water quality.

FARZANEH SAJEDI HOSSEINI received the B.S. degree in rangeland and watershed management from Shiraz University, Shiraz, Iran, in 2011, and the M.S. degree in watershed management and engineering from Sari Agricultural Sciences and Natural Resources University, Sari, Iran, in 2015. She is currently pursuing the Ph.D. degree in watershed management with the University of Tehran, Tehran, Iran. Her research interests include hydrology, spatial, and natural hazard modeling.

MOSLEM BORJI received the B.S. degree in rangeland and watershed management from the Shahid Bahonar University of Kerman, Kerman, Iran, in 2013, and the M.S. degree in watershed management and engineering from the University of Tehran, Iran, in 2015, where he is currently pursuing the Ph.D. degree. He is also working on implementation projects to control land degradation and sediment fingerprinting. He has published more than five papers in JCR-indexed journals. His research interests include dynamic systems, sediment fingerprinting, land hazards modeling, data mining, and statistic. He was a recipient of the Academic Award from the National Elites Foundation, Iran, in 2017.

HIMAN SHAHABI received the Ph.D. degree in remote sensing-GIS from the Department of Remote sensing, Faculty of Geoinformation and Real Estate and the Ph.D. degree in remote sensing-GIS from the Institute of Geoscience and Digital Earth Centre (Geo-DEC), Research Institute for Sustainability and Environment (RISE), Universiti Teknologi Malaysia (UTM), Malaysia. He is currently an Assistant Professor in remote sensing with the Department of Geomorphology (Remote Sensing/GIS), Faculty of Natural Resources, University of Kurdistan, Sanandaj, Iran. His research has been published in 135 journal articles, two books, two book chapters, and more than 40 papers presented in international and national conferences. His research interests include image processing, GIS application, soft computing techniques in natural hazards, and geosciences. He served as an Associate Editor and a Reviewer for the Board of GIS Modeling, *Remote Sensing*, *Natural Hazards*, and *Machine Learning Journals*.

ARYAN SALVATI received the B.Sc. degree in rangeland and watershed management engineering from the University of Kurdistan, Kurdistan, Sanandaj, Iran, in 2017, and the M.Sc. degree in watershed management engineering and sciences from the College of Natural Resources, University of Tehran, Tehran, Karaj, in 2020. He is interested in the development of technology in collecting and analyzing environmental data. He has several national inventions. His current research interests include GIS-based hydrology (hydro-informatic), developing machine learning models, and optimization algorithms in hydrology fields by Python, modeling, measurement of flow, and water quality.

ADRIENN A. DINEVA received the B.Sc. and M.Sc. degrees in mechatronics engineering from Obuda University, in 2013, and the Ph.D. degrees in applied informatics from the Universita degli Studi di Milano and Obuda University, in 2017. She is currently an Assistant Professor with Obuda University. Her research interests include the area of adaptive control of nonlinear systems and applied engineering, including machine learning and soft computing techniques, and also the area of adaptive signal processing. She has authored more than 20 ISI articles in these fields.