



Machine Learning-Based Snow Avalanche Susceptibility Mapping Using Avalanche-Zone Training Scenarios in the Central Alborz Watershed, Iran

ARTICLE INFO

Article Type Original Research

Authors

Majid Kazemzadeh, Ph.D.^{1*}

How to cite this article

Kazemzadeh M. ZMachine Learning-Based Snow Avalanche Susceptibility Mapping Using Avalanche-Zone Training Scenarios in the Central Alborz Watershed, Iran. ECOPERSIA 2026;14(4): 359-381.

DOI:

10.48311/ecopersia.2026.122255.82937

¹ Ph.D., Assistant Professor, Department of Range and Watershed Management, Faculty of Natural Resources and Environment, Ferdowsi University of Mashhad, Mashhad, Khorasan Razavi, Iran.

* Correspondence

Address: Ph.D., Assistant Professor, Department of Range and Watershed Management, Faculty of Natural Resources and Environment, Ferdowsi University of Mashhad, Mashhad, Khorasan Razavi, Iran.
Phone: +989128635726
Email: Kazemzadeh@um.ac.ir
ORCID: 0000-0002-9150-353X

Article History

Received: August 3, 2026
Revised: September 21, 2026
Accepted: September 27, 2026
Published: October 3, 2026

ABSTRACT

Aims: This study used machine learning models, Random Forest (RF) and Boosted Regression Tree (BRT), to predict avalanche-susceptible areas in the Central Alborz watershed, Alborz Province, Iran.

Materials & Methods: We established a polygon-based framework based on three main avalanche zones: starting, track, and runout. Three training scenarios were considered for avalanche susceptibility modeling using RF and BRT: (1) the runout zone alone, (2) the starting and track zones, and (3) all three zones combined. We randomly split the avalanche inventory into 70% training and 30% validation datasets. We initially considered 21 topographic, geomorphological, and meteorological factors. We first conducted multicollinearity and Pearson correlation analyses to identify redundant predictors. We then applied Recursive Feature Elimination (RFE) with the Random Forest (RF) algorithm and out-of-bag (OOB) permutation importance to select the optimal predictor subset, retaining 13 factors for model development. We trained models under the three scenarios and evaluated them using the Receiver Operating Characteristic (ROC) curve, precision, sensitivity, F1 score, and Matthews Correlation Coefficient (MCC).

Findings: The BRT model trained with the starting and track zones showed the best overall performance across most evaluation metrics (AUC = 0.95, precision = 0.93, F1-score = 0.91, and accuracy = 0.91).

Overall, training both BRT and RF models using the starting and track zones substantially improved snow avalanche susceptibility prediction compared with the other scenarios. For the starting and track, runout, and all-zone scenarios, the RF model achieved AUC $\geq$ 0.90, while the corresponding BRT reached AUC $\geq$ 0.92 for all three scenarios.

Conclusion: The results showed that the BRT and RF models achieved their highest predictive performance under the starting and track zones scenario, with AUC values of 0.95 and 0.93, respectively. These findings indicate that incorporating the spatial characteristics of avalanche zones can improve susceptibility assessment and provide useful information for avalanche hazard management and disaster risk reduction in mountain watersheds.

Keywords: Boosted Regression Tree; Machine Learning; Natural Disasters; Random Forest; Snow Avalanche Susceptibility.

CITATION LINKS

[1] Viglietti D., Letey S., Mo... [2] Hao J., Zhang X., Cui P., ... [3] Ballesteros-Cánovas J.A., ... [4] Schweizer J., Bartelt P., ... [5] Durlević U., Valjarević A.... [6] Ahmad I., Mallick J., Hang... [7] Schweizer J., Jamieson J.B... [8] Yousefi S., Pourghasemi H.... [9] Rafique A., Dasti M.Y., Ul... [10] Wang H., Zhang L., Yin K.,... [11] Rahman M., Chen N., Islam ... [12] Bian R., Huang K., Liao X.... [13] Tang Y., Sun Y., Han Z., W... [14] Akay H. Spatial modeling o... [15] Viallon-Galinier L., Hagen... [16] Liu J., Zhang T., Hu C., W... [17] Bansal J.K., Goswami A., R... [18] Abhinav A., Sattar A. Snow... [19] Tian X., Yao X., Zhou Z., ... [20] Choubin B., Borji M., Mosa... [21] Rahmati O., Ghorbanzadeh O... [22] Yariyan P., Omidvar E., Ka... [23] Liu Y., Chen X., Yang J., ... [24] Kazemzadeh M., Noori Z. An... [25] McClung D., Schaerer P. Th... [26] Voiculescu M., Ardelean F... [27] Denissova N., Nurakynov S.... [28] Fazzini M., Cordeschi M., ... [29] Dormann C.F., Elith J., Ba... [30] Habumugisha J.M., Chen N.,... [31] Guns M., Vanacker V. Logis... [32] Lei X., Chen W., Avand M.,... [33] Chen W., Chen Y., Tsangara... [34] Darst B.F., Malecki K.C., ... [35] Liaw A., Wiener M. Classif... [36] Hastie T., Tibshirani R., ... [37] Bui D.T., Tuan T.A., Klem... [38] Breiman L. Random forests... [39] Robinsonov N. Advances in ... [40] Elith J., Leathwick J.R., ... [41] Youssef A.M., Pourghasemi ... [42] Schapire R.E. The boosting... [43] Fawcett T. An introduction... [44] Nachappa T.G., Ghorbanzade... [45] Ullah K., Wang Y., Fang Z.... [46] Sinčić M., Bernat Gazibara... [47] Frattini P., Crosta G., Ca... [48] Ziarh G.F., Asaduzzaman M.... [49] Rainio O., Teuho J., Klén ... [50] Matthews B.W. Comparison o... [51] Zheng H., Ding M., Huang T... [52] Li Y., Yang Z., Cheng Q., ... [53] Mosavi A., Shirzadi A., Ch... [54] Varol N. Avalanche suscept... [55] Cetinkaya S., Kocaman S. I... [56] McClung D.M. Effects of tr... [57] Yang J., He Q., Liu Y. Win... [58] Korzeniowska K., Bühler Y... [59] Gao M., Gu X., Liu Y., Zha...

Introduction

Snow avalanches are among the natural disasters that occur in mountain regions and, depending on their size and type, can affect mountain ecosystems, biodiversity, and human habitats ^[1, 2]. In recent decades, the risk of snow avalanches has increased due to climate change ^[2, 3]. Snow avalanches are a form of mass movement and come in different types (e.g., wet, dry, or powder) and sizes (small, medium, large, very large, and extreme) based on their destructiveness ^[4]. The main zones of a snow avalanche include starting (or release), track (or transition), and runout (or deposition) zones, which have different geomorphologic characteristics that hazard modeling must account for ^[5, 6]. The starting zone is the main area where a snow avalanche initiates and generally consists of terrain with slopes ranging from 30~45° ^[7]. Complex interactions among topography, snowpack properties, and meteorological conditions influence snow avalanche occurrence and behavior. Understanding these factors and incorporating them into predictive models are essential for improving avalanche risk assessment and mitigation strategies.

Snow avalanches have caused significant humanitarian and economic losses. However, predicting snow avalanche occurrence is very important for researchers. Similar to other regions worldwide, snow avalanches have caused the deaths of many people annually in Iran. For example, snow avalanches triggered by heavy snowfall in the Oshtorankuh Mountains of the Zagros Mountain Range and at the Dizin International Ski Resort in the Alborz Mountain Range have killed many people ^[8]. Although many mountain regions in Iran are prone to snow avalanches annually, spatial

and temporal data on avalanche hazards are lacking, and these hazards have not been mapped. Here, the first step is to create spatial distribution maps of avalanches and identify potential snow-avalanche-prone areas to reduce snow-avalanche damage ^[5, 9]. In recent decades, machine learning techniques have become a powerful tool in many environmental and natural hazards applications ^[10, 11, 12, 13]. Akay ^[14] used machine learning methods such as random trees (RT) and random forests (RF) to create snow avalanche susceptibility maps in Uzungol Basin, Turkey. Viallon-Galinier et al. ^[15] combined snow physical characteristics and avalanche occurrences within an RF approach to predict snow avalanche in a mountain area. Liu et al. ^[16] used novel tree-based machine-learning models with remote sensing images and observation data to predict snow avalanche susceptibility in the Tianshan Mountains, China. Bansal et al. ^[17] integrated climate data, terrain characteristics, and field validations to apply Object-Based Image Segmentation (OBIS) analysis for more precise hazard zonation in the Kullu region of Himachal Pradesh, India. Ahmad et al. ^[6] employed a Bayesian-optimized hybrid deep learning framework (CNN-LSTM-Attention) to map Himalayan snow avalanche susceptibility, concluding that land-cover, NDVI, and topography were the most influential factors via SHAP analysis. Abhinav and Sattar ^[18] assessed avalanche susceptibility in the Western Himalaya using RF, SVM, LR, and ANN models, finding Random Forest to be the most accurate (AUC = 0.95) and classifying 8% of the area as highly susceptible. Tian et al. ^[19] proposed a progressive avalanche hazard framework by integrating Bayesian-optimized XGBoost, SHAP, and RAMMS in the

Eastern Himalayas. The model achieved high predictive performance (AUC=0.932). NDVI, aspect, slope, and terrain roughness were identified as the most influential factors. Several studies have mapped snow avalanche susceptibility in Iran using machine learning techniques [8, 20, 21, 22]. Despite the increasing use of machine learning for avalanche susceptibility mapping, one methodological limitation of previous studies is how they represent avalanche occurrences spatially. Many previous studies represent avalanche events as single points, typically located in the runout (deposition) zone [8, 21, 22]. Although such point-based sampling is practical, it may not adequately represent the spatial extent of avalanche-prone areas. In particular, the starting, track, and runout zones constitute different parts of an avalanche path and are characterized by distinct geomorphological conditions. Consequently, relying on a single zone may introduce spatial sampling bias and limit susceptibility models' ability to characterize the full avalanche-affected terrain. This study addresses this limitation by incorporating polygon-based samples from all three avalanche zones into the training process and evaluating their effects on susceptibility-model reliability. In addition, conventional statistical approaches may struggle to capture the complex, nonlinear relationships among the factors influencing avalanche occurrence. Machine learning techniques such as BRT and RF can better accommodate such nonlinear relationships and interactions, providing a suitable framework for modeling avalanche susceptibility [23].

This study evaluates a polygon-based sampling framework for snow avalanche susceptibility mapping in the Central Alborz watershed, Alborz Province, Iran, using

machine learning. The framework uses polygon-based representations of the three snow avalanche zones (starting, track, and runout) to assess whether including them yields a more representative training dataset. Specifically, three distinct sampling scenarios were tested: (1) the runout zone only, (2) the starting and track zones combined, and (3) all three zones. We used these sampling schemes to develop corresponding training datasets, which we then used to train the Random Forest (RF) and Boosted Regression Tree (BRT) models. Because they capture complex, nonlinear relationships between avalanche occurrence and influencing factors, we used these models to evaluate how different spatial sampling strategies affect model performance. We initially considered 21 factors related to topography, geomorphology, and meteorological conditions that may influence avalanche occurrence. We identified the optimal combination of influencing factors using multicollinearity analysis and recursive feature elimination (RFE). Comparing these sampling scenarios provides an empirical basis for assessing how spatial sampling strategy affects susceptibility model performance and the delineation of avalanche-prone areas.

Materials & Methods

Study Area

The study area encompasses the Central Alborz watershed, specifically the Kasil-Nesa, Velayatrood, and Varangehrood sub-watersheds. This mountainous region in northern Alborz Province, Iran, covers approximately 272.55 km² and lies between 36°01'02"N–36°07'16"N latitude and 51°17'24"E–51°32'33"E longitude. The area receives a mean annual precipitation of 800

mm and has a mean annual temperature of 1.4 °C [24]. Rangeland constitutes over 90% of the land-use. The predominant geology comprises the Middle Tuff Member of the Cenozoic, consisting of thick-bedded green tuffs, minor lava flows, shale, pyroclastics, conglomerate, and limestone. The area is home to the Dizin International Ski Resort, attracting hundreds of tourists annually. Dizin's popularity as a skiing and mountaineering destination contributes to its high avalanche susceptibility, making it a critical area for avalanche hazard assessment and risk management. A significant avalanche event occurred on February 4, 2010, triggered by heavy snowfall, unstable snowpack (due to unstable snow layers), and adverse weather conditions, resulting in three consecutive avalanches. Thirty-one mountaineers and skiers were caught, tragically resulting in eight fatalities. Excessive snow accumulation, temperature fluctuations that weakened the snowpack, and human activity initiated the avalanches. This event represents only one of several avalanches that have occurred in the region historically. Figure 1 depicts the study area's geographic information.

Methods

This research consists of five main components: (1) collecting and processing inventory data from three main snow avalanche zones (Figure 2), (2) identifying and preparing a set of influencing factors for snow avalanche occurrence, (3) conducting multicollinearity analysis and selecting relevant features, (4) modeling snow avalanches using machine learning techniques (RF and BRT) under various scenarios and generating a snow avalanche susceptibility map, and (5) evaluating

the performance of the models. Figure 3 illustrates the study workflow.

Snow Avalanche Inventory

In this study, we compiled snow avalanche inventory data from local observations, historical avalanche events, and field surveys conducted during 2023–2024, identifying 50 polygon-based avalanche-prone areas (Figure 2). In addition, we randomly selected 70 non-avalanche polygons from areas identified as non-avalanche based on available evidence of past avalanche occurrence and the researcher's expert knowledge. We then created a spatial database that categorized avalanche zones into three distinct areas: starting, track, and runout zones (Figure 2a). Based on field observations, slope ranges, and satellite imagery, we classified avalanche-prone areas into three zones: starting zones (slopes >30°), tracks (15–30°), and runout zones (<15°) [25] (Figure 2b). Spatial analysis revealed that most avalanches in the study area are concentrated in the southern and western regions, particularly near roads and residential areas. For machine learning-based susceptibility mapping, we randomly divided the 120 polygons (50 avalanche and 70 non-avalanche polygons) at the polygon level into training (70%) and validation (30%) datasets before spatial extraction to reduce potential spatial leakage between the training and validation datasets. We maintained the same polygon-level partitioning across the three modeling scenarios, then used ArcGIS to extract point locations from the starting, track, and runout zone polygons in each scenario. Consequently, the number of extracted sample points varied across the three scenarios based on the spatial extent of the selected polygons. We then

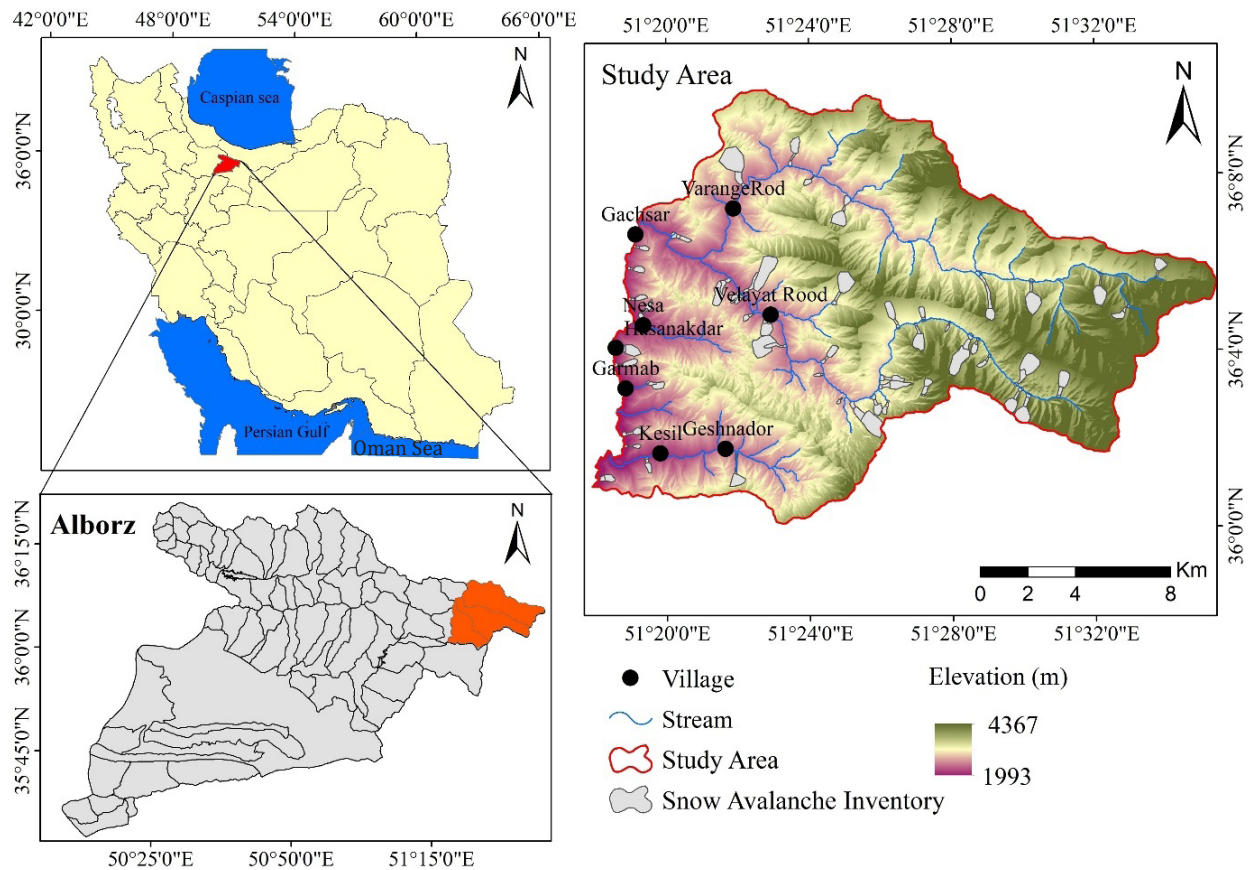


Figure 1) Location of the study area and spatial distribution of the snow avalanche inventory.

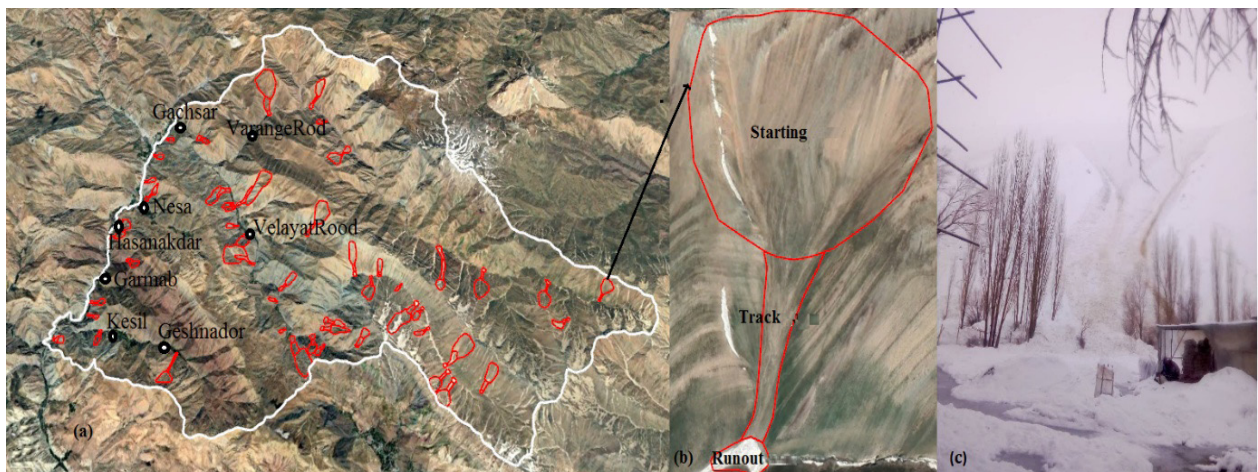


Figure 2) Snow avalanche occurrence areas in the study area: (a) an example of snow avalanche zones, (b) classification of avalanche occurrence areas into starting, track, and runout zones, and (c) a photograph of an avalanche occurrence in the study area.

used these extracted points to train and validate the machine learning models. To ensure balanced classification, the system randomly generated an equal number of

avalanche and non-avalanche samples. In the model, we assigned avalanche occurrences a value of "1," while non-avalanche locations received "0."

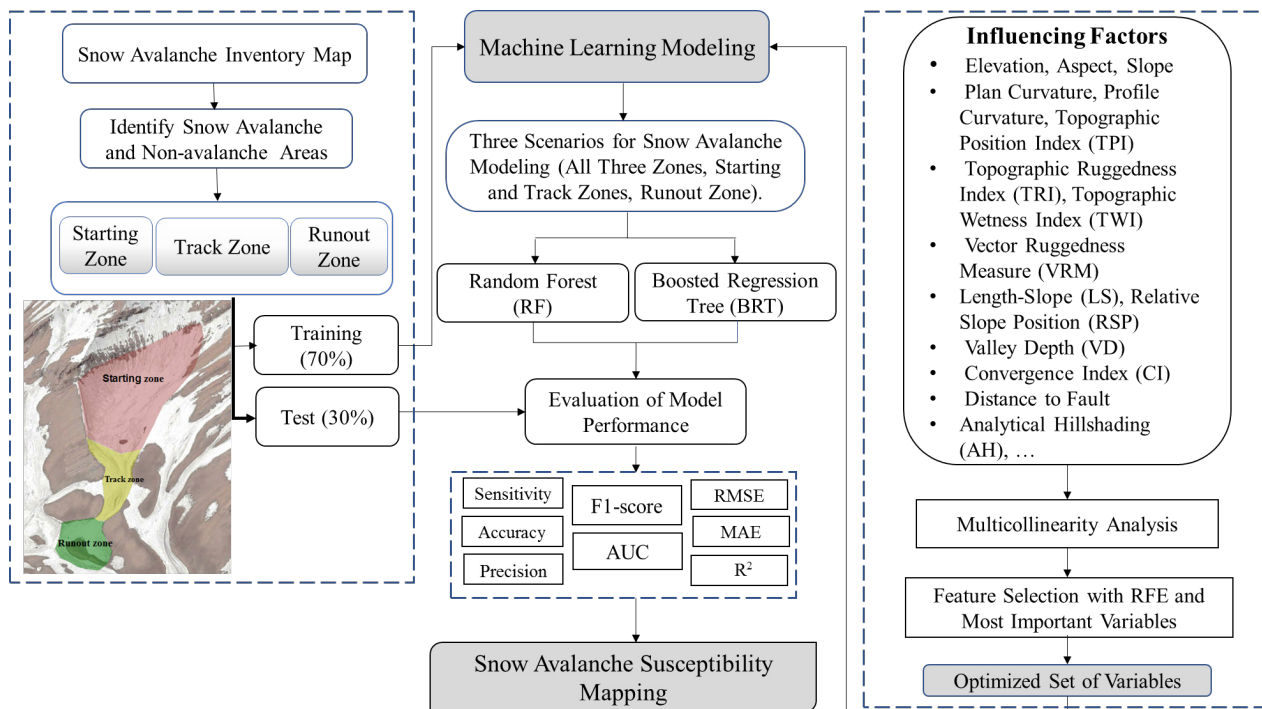


Figure 3) Flowchart of the snow avalanche susceptibility mapping framework used in this study.

Influencing Factors in Snow Avalanche Occurrence

Various factors influence snow accumulation and snow avalanche occurrence, including topographic, geological, geomorphological, hydrological, and climatic conditions [16, 26, 27]. We derived the topographic and hydrological predictors from the 10m×10m resolution digital elevation model (DEM) using ArcGIS 10.3 and SAGA GIS tools. The influencing factors were classified into four groups based on their physical characteristics: topographic factors (elevation, slope, aspect, curvature, TPI, TRI, VD, RSP, PLA, PRO, DTR, and LS), climatic factors (Rainfall and Temperature), hydrological factors (TWI), and terrain-related factors (AH, VRM, CI, and DTF), lithology, and land-use, as illustrated in Figure 4. Topographic factors mainly control snow accumulation, redistribution, solar exposure, slope stability, and avalanche movement [28]. Meteorological factors, including rainfall, temperature, and

wind, directly affect snowpack stability and avalanche triggering [23, 27]. For modeling, we used monthly precipitation and temperature data from neighboring stations in the study area. In this study, precipitation and temperature served as spatial predictors representing regional climatic gradients across the study area rather than specific event-trigger conditions. Therefore, they characterized spatial variations in climatic conditions relevant to snow accumulation and stability, rather than predicting the timing of individual avalanche events. TWI represents the potential for water accumulation and moisture conditions, which may affect snowmelt and wet-snow instability. AH reflects local terrain morphology and relief, which can affect snow accumulation and redistribution. DTF represents the distance to faults and may indirectly influence avalanche occurrence by affecting terrain structure and slope stability.

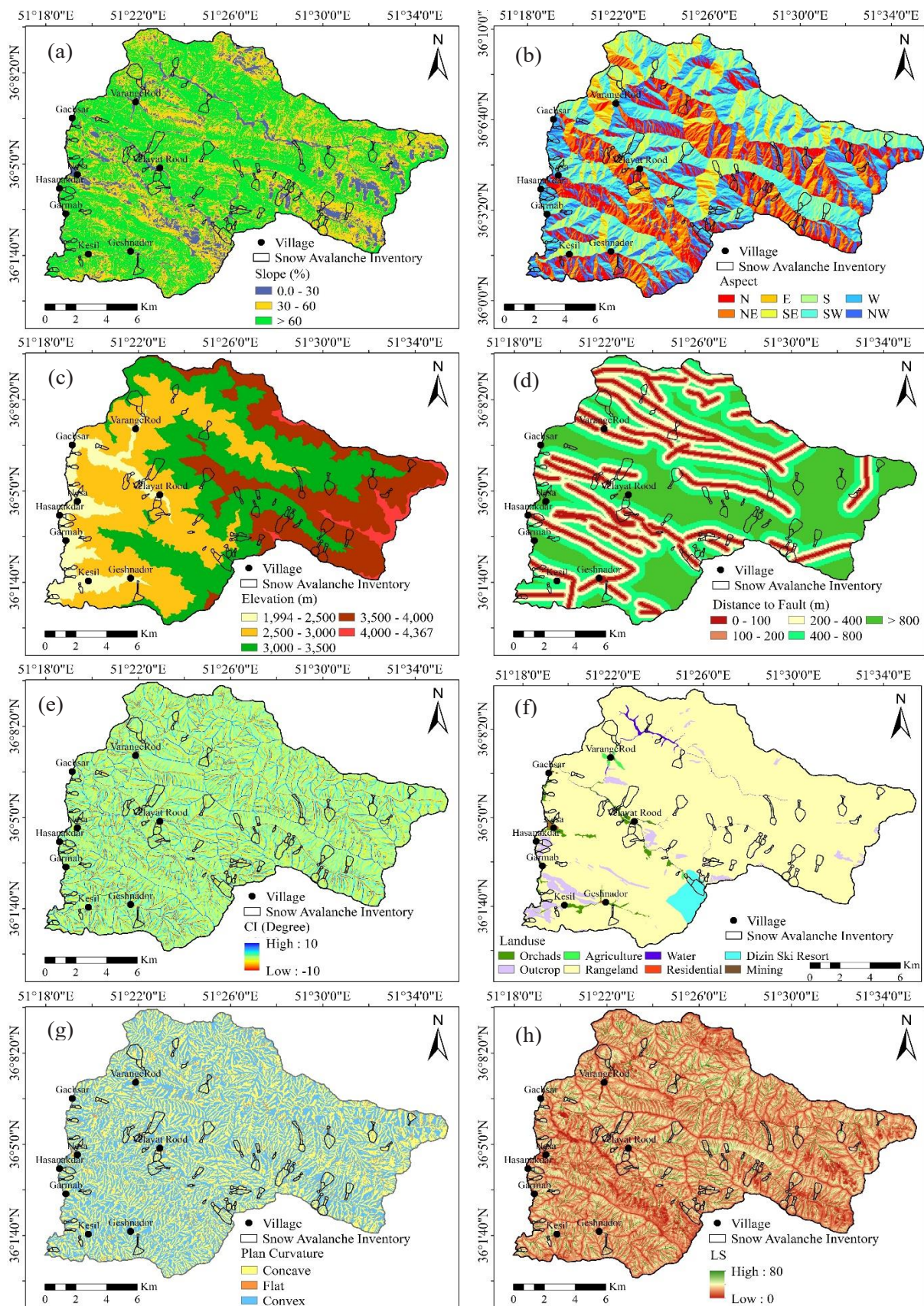


Figure 4) Snow avalanche influencing factors used in susceptibility modeling: (a) slope, (b) aspect, (c) elevation, (d) distance to fault, (e) CI, (f) land-use, (g) plan curvature, and (h) LS.

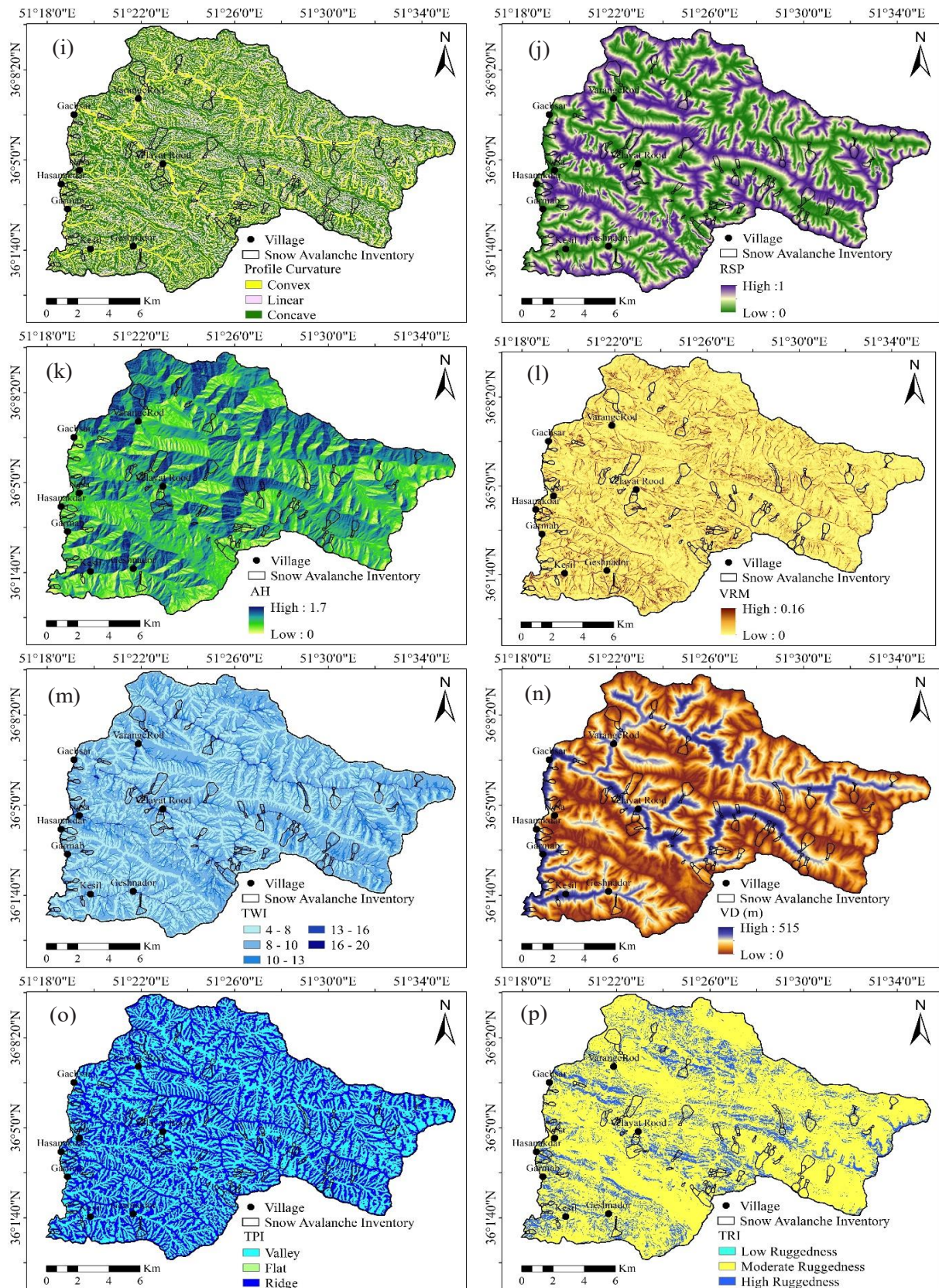


Figure 4 Continued) Snow avalanche influencing factors used in the susceptibility modeling: (i) profile curvature; (j) RSP; (k) AH; (l) VRM; (m) TWI; (n) VD; (o) TPI; (p) TRI.

Data Preprocessing and Feature Selection

Data preprocessing is an important step in preparing appropriate data for creating a predictive model using machine learning. We conducted multicollinearity analysis to identify and select the most appropriate factors influencing snow avalanche occurrence. Multicollinearity quantifies correlations among predictor variables (influencing factors) in a regression model to eliminate redundant factors [29, 30]. We evaluated multicollinearity using Variance Inflation Factors (VIF) and the Tolerance Index (TOL). We considered predictor variables with a Pearson correlation coefficient (r) ≤ 0.50 appropriate for inclusion [30]. A $TOL < 0.2$ and $VIF > 10$ indicated multicollinearity [31, 32, 33]. We calculated VIF and TOL using Eqs. (1) and (2), respectively, where R_i^2 is the coefficient of determination obtained by regressing the i -th predictor on the remaining predictors, and i denotes the predictor variable.

$$VIF = \frac{1}{1 - R_i^2} \quad \text{Eq. (1)}$$

$$TOL = 1 - R_i^2 \quad \text{Eq. (2)}$$

After removing highly correlated variables, Recursive Feature Elimination (RFE) combined with the RF algorithm and out-of-bag (OOB) permutation importance was applied to select the most robust predictors. This method evaluates variable importance by measuring the drop in prediction accuracy when a specific predictor's OOB values are randomly permuted [34, 35, 36]. Consequently, key variables cause a sharper decline in model performance than redundant ones. Through this iterative ranking process, we retained an optimal subset of 13 features for the final model. In this study, we implemented

the RFE algorithm in Python using scikit-learn with ten-fold cross-validation.

Scenario-Based Snow Avalanche Modeling

After establishing the base dataset, we mapped each of the three snow avalanche zones separately and extracted corresponding data for each zone. We defined three distinct modeling scenarios based on combinations of sampling schemes for starting, track, and runout zones. We delineated the polygons for each zone independently to ensure distinct spatial extraction. We split polygon-based avalanche and non-avalanche inventories from the three main avalanche zones into two datasets, with training and validation polygons defined for each. We defined three scenarios for snow avalanche susceptibility analyses. Scenario (1): The runout zone, which used polygon-based locations of snow avalanches in the runout zone (50 polygons; 194 extracted points) for modeling. Scenario (2): The track and starting zones, consisting of polygon-based areas from the inventory data (100 polygons; 438 extracted points). Scenario (3): We considered all three main snow avalanche zones (150 polygons; 548 extracted points) for modeling. We then randomly partitioned these polygon datasets into training (70%) and testing (30%) subsets for model implementation. For each scenario, we extracted influencing factors based on the respective zone polygons, converted them into point features, and compiled them into distinct datasets.

As a result, we generated six modeling procedures for the three modeling scenarios. Next, we statistically modeled snow avalanche susceptibility using RF and BRT. These models are effective for avalanche susceptibility modeling because they capture complex, nonlinear relationships among environmental variables. RF excels

at handling large datasets and interactions between predictors using multiple decision trees [37]. BRT improves prediction accuracy by combining simple models through boosting, making both methods well-suited for predicting avalanche occurrences. The modeling was conducted using Python.

Random Forest (RF)

RF is an ensemble learning algorithm that generates predictions by combining multiple decision trees constructed from bootstrap samples of the training data [38]. In this study, we implemented the RF algorithm by randomly sampling data points in the forest (*ntree*) to build an ensemble of unpruned decision trees. At each node, we selected a random subset of controlling factors (*mtry*) to determine the optimal split, and we obtained the final susceptibility prediction by aggregating all trees via majority voting [38]. Breiman [38], Liaw and Wiener [35], and Hastie et al. [36] provide more detail on RF. Both *ntree* and *mtry* are hyperparameters that need tuning to obtain optimal predictions. The RF model used *ntree* = 200, *mtry* = 2, and minimum node size = 1.

Boosted Regression Tree (BRT)

BRT is a non-parametric model that fits complex nonlinear relationships between independent and dependent variables [39]. BRT combines many simple models through boosting to improve prediction accuracy. A more advanced application of the BRT method is modeling natural phenomena with nonlinear relationships [40, 41]. The BRT model uses regression trees. BRT improves prediction accuracy by combining and averaging multiple weak prediction rules, yielding more robust results than relying on a single highly precise prediction rule [42]. For the BRT model, the optimized values for the number of trees (*n.trees*), interaction depth

(*interaction.depth*), and minimum number of observations in terminal nodes (*n.minobs.innode*) were 300, 1, and 10, respectively.

Evaluation of Model Performance

Model performance is evaluated to assess the predictive performance of the machine learning models and their ability to predict snow avalanche susceptibility spatially. The performance of the models was assessed based on receiver operating characteristic (ROC) curves and the area under the curve (AUC) [43, 44, 45, 46], and common metrics for binary classification, such as precision, sensitivity (recall), accuracy, F1-score, and the Matthews correlation coefficient (MCC) [10, 11, 47, 48, 49, 50]. MCC is a robust, balanced metric that accounts for all confusion matrix categories and returns a value between -1 and +1, providing a reliable evaluation of classification quality. The formulas for these metrics are as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad \text{Eq. (3)}$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{Eq. (4)}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad \text{Eq. (5)}$$

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{Eq. (6)}$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad \text{Eq. (7)}$$

where Accuracy is the overall classification accuracy; Precision is the proportion of correctly predicted positive samples among all predicted positive samples; Sensitivity (Recall) is the proportion of correctly predicted positive samples among all actual positive samples; F1-score is the harmonic mean of Precision and Recall; TP, TN, FP, and

FN denote true positive, true negative, false positive, and false negative, respectively.

Findings

Selection of Snow Avalanche Influencing Factors

Table 1 presents the results of the multicollinearity analysis using the TOL and VIF methods. The lowest and highest VIF values are 1.27 and 6.11 for AH and lithology, respectively (Table 1). The findings indicate no multicollinearity among the selected influencing factors; therefore, no variable was excluded based on the VIF and TOL criteria. Linear correlations among influencing factors can reduce prediction accuracy. Figure 5 indicates the correlation between the independent variables (snow avalanche influencing factors) and the dependent variable (snow avalanche occurrence). We considered variables with a Pearson correlation coefficient ≤ 0.50 suitable for modeling. Most correlation coefficients ranged from 0.2 to 0.4, indicating weak correlations among the variables and making them appropriate for modeling. Consequently, we removed variables with correlation coefficients exceeding the predefined threshold.

The Most Important Factors for Snow Avalanche Modeling

Following Pearson correlation and multicollinearity analyses, we applied Recursive Feature Elimination (RFE) with the Random Forest (RF) algorithm and out-of-bag (OOB) permutation importance to identify the optimal subset of predictors. Through this iterative feature-selection procedure, we retained 13 factors for subsequent modeling. These factors included TPI, valley depth (VD), topographic wetness index (TWI), relative slope position (RSP), plan curvature (PLA), distance to river (DTR), slope length (LS), rainfall,

temperature, elevation, analytical hillshading (AH), aspect, and distance to fault (DTF). We then evaluated the relative importance of the 13 retained predictors using the RF mean decrease in accuracy measure (Figure 6). The results showed that TPI, RSP, VD, TWI, DTR, and LS were among the most influential predictors of snow avalanche occurrence in both model training and inference. Rainfall, elevation, temperature, aspect, AH, PLA, and DTF also contributed to model prediction, although their relative importance was lower than that of the aforementioned factors.

Model Performance, Validation and Comparison

We analyzed model performance using comprehensive evaluation metrics, including AUC, MCC, Precision, Sensitivity, Accuracy, and F1-score, on the training dataset (Table 2 and Figure 7). On the training dataset, the BRT model outperformed the RF model on most evaluation metrics. Furthermore, MCC values, a robust indicator of classification quality, ranged from 0.77 to 1.00 across models, with the BRT model showing higher values in the starting and track-zone scenarios (Table 2). The BRT model's accuracy values (0.96-1.00) were higher than the RF model's (0.89-0.98). Among the models trained with the three-zone scenario, performance was balanced. The BRT model trained with the starting and track zones scenario achieved the highest sensitivity and precision (1.00), followed by the model trained with the three-zone scenario (0.95, 0.99). The results showed that AUC values for the BRT model in the training datasets ranged from 0.96 to 0.99, with the highest value in the starting and track zones scenario (AUC=0.99). These results indicate that training with the track and starting zones scenario improved the models' avalanche prediction capabilities, followed by the three zones scenario,

Table 1) Multicollinearity assessment using variance inflation factors (VIFs)

Factors	VIF	TOL
Slope	1.55	0.65
Aspect	1.80	0.56
Elevation	1.55	0.65
Curvature	2.33	0.43
Profile Curvature (PRO)	2.86	0.34
Plan Curvature (PLA)	4.14	0.24
Analytical Hillshading (AH)	1.27	0.79
Topographic Position Index (TPI)	3.84	0.26
Terrain Ruggedness Index (TRI)	1.31	0.77
Vector Terrain Ruggedness (VRM)	2.45	0.40
Topographic Wetness Index (TWI)	1.87	0.53
Convergence Index (CI)	2.71	0.37
Relative Slope Position (RSP)	1.40	0.71
Length-Slope (LS)	1.32	0.76
Valley Depth (VD)	2.70	0.37
Rainfall	1.99	0.50
Temperature	2.18	0.46
Lithology	6.11	0.36
Land-Use	4.30	0.23
Distance to Fault (DTF)	2.70	0.37
Distance to River (DTR)	2.24	0.45

compared with models trained on the runout zone scenario.

Table 3 and Figure 8 show the overall performance of the BRT and RF models on the validation dataset. The BRT model showed higher values in some metrics, achieving the highest sensitivity (0.91), precision (0.93), and accuracy (0.91). The AUC values for the RF and BRT models trained with the three-zone, start-and-track, and runout-zone scenarios were 0.90 and 0.92, 0.93 and 0.92, and 0.95 and 0.95, respectively. Using the validation dataset, although the RF and BRT methods yielded reasonable predictive results, the BRT model (AUC: 0.92-0.95) achieved a higher area under the ROC curve than the RF model (AUC: 0.90-

0.93). Based on the evaluated performance metrics, models trained with the starting and track zones scenario performed better than those trained with the other scenarios.

Furthermore, MCC values ranged from 0.81 to 0.87 across the models, reflecting reliable classification performance (Table 3). AUC values for both models exceeded 0.90, indicating strong predictive capability. BRT and RF were both effective models in predicting snow avalanche susceptibility. Overall, model performance improved when using the starting and track zones scenario for snow avalanche modeling. The proposed methods enhanced the evaluation metrics of both the BRT and RF models.

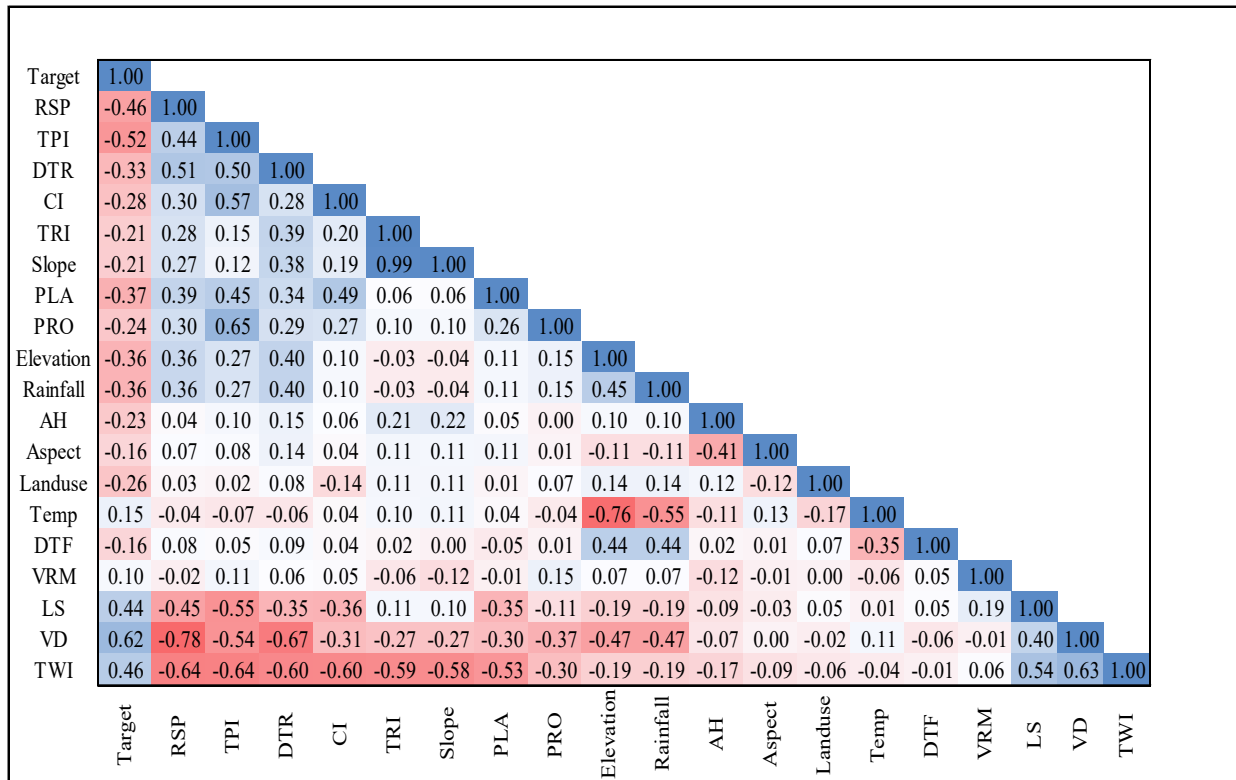


Figure 5) Correlation analysis of snow avalanche influencing factors.

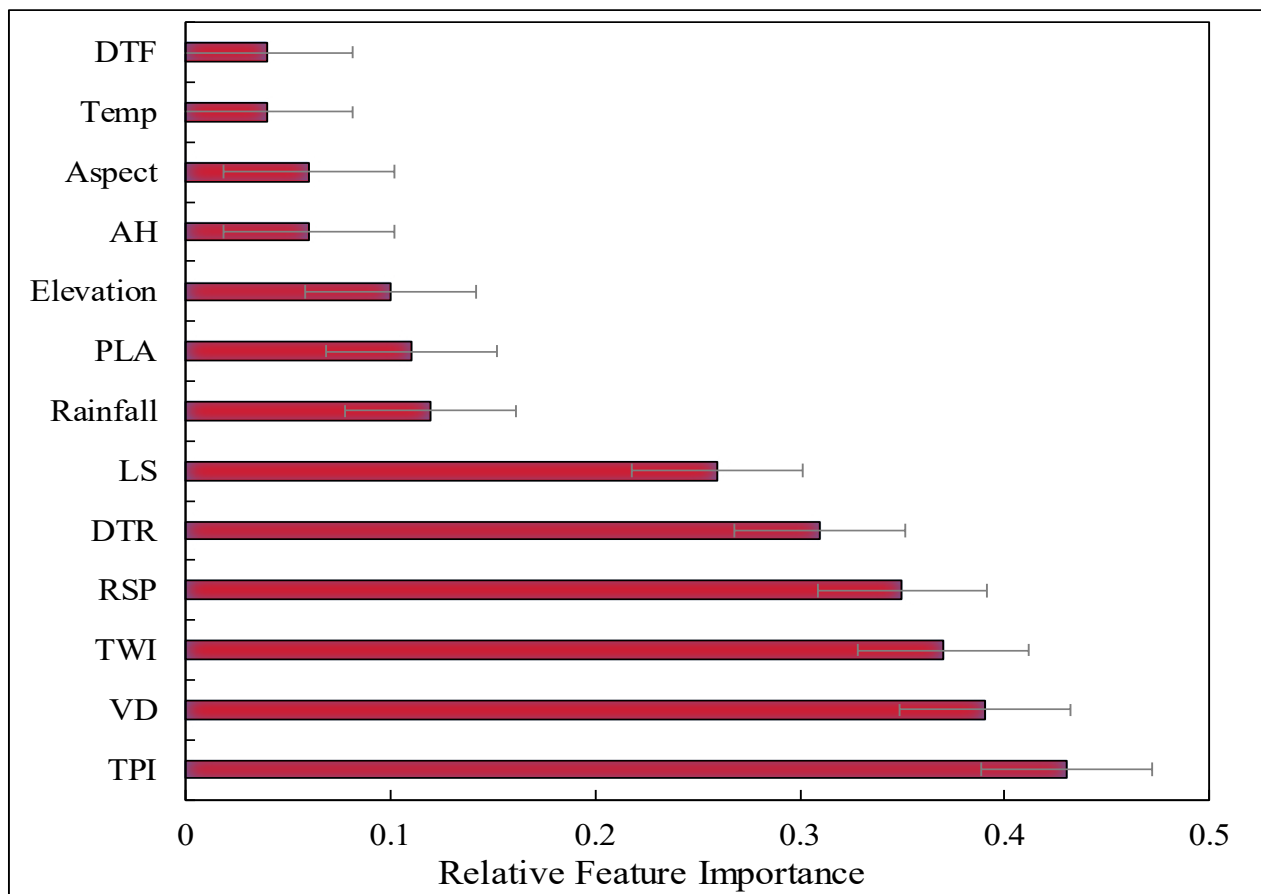


Figure 6) Importance of influencing factors in RF modeling.

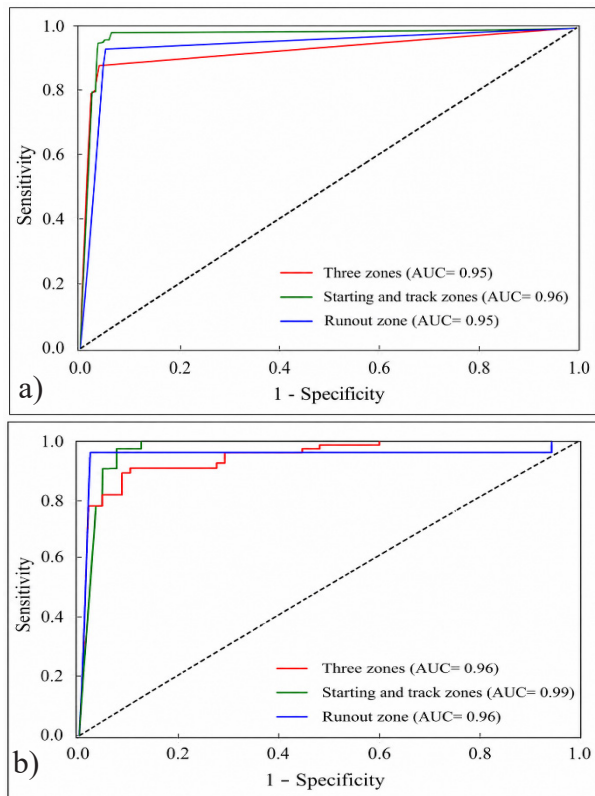


Figure 7) ROC curves and AUC values for the RF (a) and BRT (b) models using the training dataset.

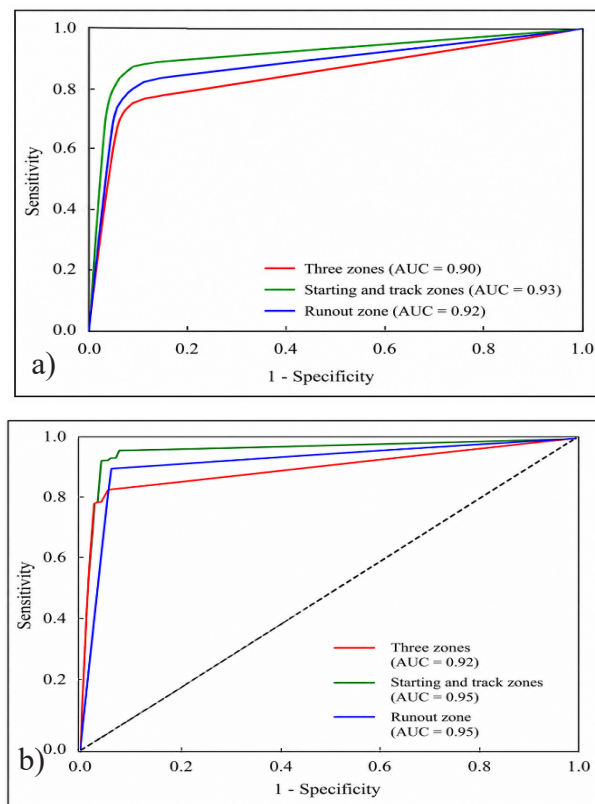


Figure 8) ROC curves and AUC values for the RF (a) and BRT (b) models using the validation dataset.

Snow Avalanche Susceptibility Mapping

The RF and BRT models were employed to model snow avalanche susceptibility maps. The study area susceptibility was classified into five main classes: very low (0-0.2), low (0.2-0.4), medium (0.4-0.6), high (0.6-0.8), and very high (0.8-1.0) using the natural breaks classification scheme. Figure 9 illustrates the snow avalanche susceptibility zones for the study area using the BRT model. The BRT model performed best for snow avalanche modeling when trained with the starting and track zones scenario (Figure 9b). This was followed by a model trained with the three-zone scenario (Figure 9c). However, training the model using only the runout zone substantially reduced its prediction capability (Figure 9a). As shown in Figure 9a, most of the study area was classified as very low susceptibility (79%).

In contrast, the BRT model trained using the three-zone scenario assigned a larger proportion of the study area to the high and very high susceptibility classes, which collectively accounted for 34.8% of the study area, a higher proportion than under the other scenarios. Similarly, Figure 10a illustrates that the RF model trained with the runout zone displayed the largest area share in the low and very low susceptibility categories (79%). In comparison, its classification of medium- and high-susceptibility areas remained below 20%. Conversely, the RF model trained with the three-zone scenario exhibited the highest area proportion in the very high susceptibility class (41.1%) (Figure 10c). On average, the BRT model classified less than 30% of the study area as very high susceptibility.

Highly susceptible areas strongly correlated with drainage and steep slopes, whereas low-susceptibility regions were predominantly on gentle slopes, where avalanche formation conditions are unfavorable. The findings further indicate that high-susceptibility areas identified by the BRT model trained with starting and track zones overlapped more with historical avalanche

events, consistent with the performance evaluation metrics in Table 3. Overall, models trained with avalanche occurrence data concentrated in starting and track zones exhibited superior predictive performance in snow avalanche susceptibility mapping.

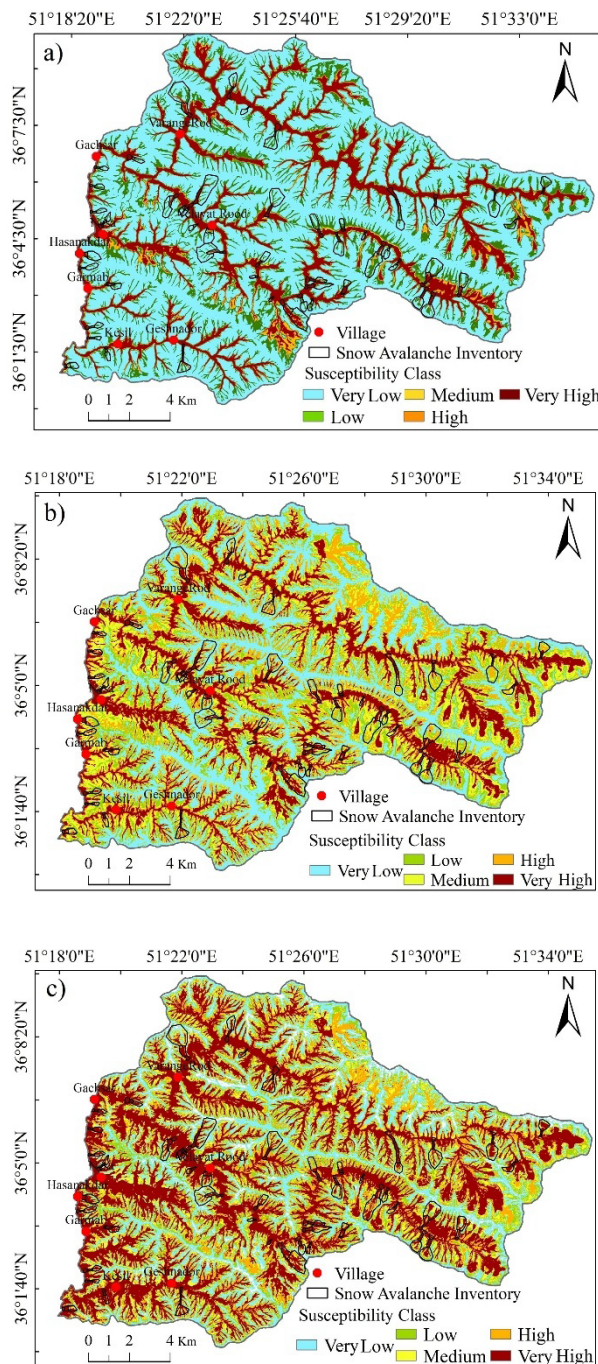


Figure 9) Snow avalanche susceptibility maps generated using the BRT model for the runout-zone scenario (a), starting- and track-zone scenario (b), and three-zone scenario (c).

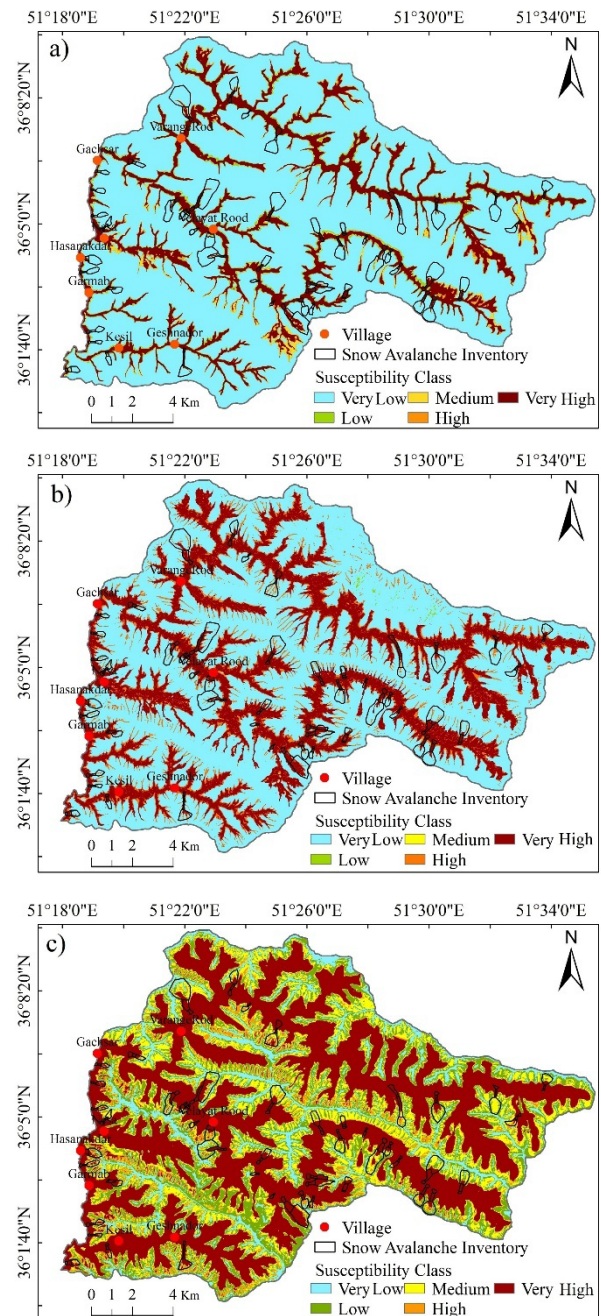


Figure 10) Snow avalanche susceptibility maps generated using the RF model for the runout-zone scenario (a), starting- and track-zone scenario (b), and three-zone scenario (c).

Figure 11 presents the spatial distribution of snow avalanche susceptibility classes predicted by the BRT and RF models, along with their respective area percentages. In the BRT model trained using the runout zone scenario, 79.2%, 0.6%, 2.1%, 1.6%, and 16.6% of the study area were classified as very low,

Table 2) Model performance using the training dataset.

Criterion	Three Zones		Starting and Track Zones		Runout Zone	
	RF	BRT	RF	BRT	RF	BRT
AUC	0.95	0.96	0.96	0.99	0.95	0.96
MCC	0.86	0.84	0.90	1.00	0.90	0.90
Precision	0.90	0.99	0.95	1.00	0.95	0.96
Sensitivity	0.87	0.95	1.00	1.00	0.95	0.97
Accuracy	0.89	0.96	0.98	1.00	0.98	0.99
F1-score	0.88	0.97	0.97	1.00	0.97	0.98

Table 3) Model performance using the validation dataset

Criterion	Three Zones		Starting and Track Zones		Runout Zone	
	RF	BRT	RF	BRT	RF	BRT
AUC	0.90	0.92	0.93	0.95	0.92	0.95
MCC	0.87	0.82	0.83	0.85	0.81	0.81
Precision	0.88	0.89	0.89	0.93	0.89	0.89
Sensitivity	0.83	0.91	0.90	0.90	0.90	0.89
Accuracy	0.89	0.91	0.91	0.91	0.90	0.90
F1-score	0.85	0.89	0.90	0.91	0.89	0.89

low, medium, high, and very high susceptibility classes, respectively. In contrast, when the BRT model was trained using the starting and track zones scenario, these classes covered 25.5% (very low), 16.6% (low), 10.4% (medium), 14.6% (high), and 33.2% (very high) of the study area. Moreover, in the BRT model trained with the three-zone scenario, the very high susceptibility class covered the largest proportion of the study area (34.8%), followed by very low (21.3%), high (16.5%), medium (13.7%), and low (13.6%) classes. Similarly, for the RF model trained using the runout zone scenario, the area proportions for very low, low, medium, high, and very high susceptibility zones were 78.5%, 0.4%, 0.1%, 4.2%, and 16.8%, respectively. When the RF model was trained with the starting-

and-track zones scenario, most of the study area was classified as either very low (63.3%) or very high (35.7%) susceptibility zones. Additionally, under the three-zone training scenario, the RF model predicted the highest area share in the very high susceptibility class (41.1%). Overall, the results indicate that both the BRT and RF models classify the majority of the study area within the very low susceptibility class when trained using only the runout zone scenario. Conversely, when trained with the three zones scenario, both models predict the highest proportion of the study area within the very high susceptibility class, highlighting the importance of incorporating data from starting, track, and runout zones for improved avalanche susceptibility modeling.

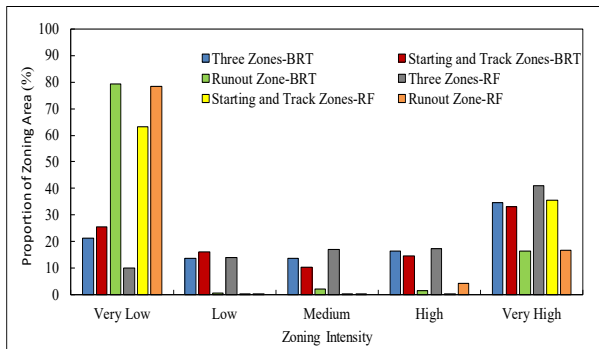


Figure 11) Proportion of the study area within each snow avalanche susceptibility class.

Discussion

Snow avalanches in the study area are the most common event causing annual financial and human losses. Therefore, accurately predicting the spatial distribution of this phenomenon and improving predictive capabilities are crucial for decision-making and disaster reduction policies. In this study, we used two machine learning models, BRT and RF, to generate snow avalanche susceptibility maps. We compared the models' performance in delineating spatial susceptibility using statistical indices, including the F1 score, sensitivity, precision, accuracy, and the Receiver Operating Characteristic (ROC) curve. The evaluation showed that the BRT model performed better than the RF model in identifying the spatial susceptibility of areas prone to snow avalanches. Various studies show that the relationship between environmental and topographic factors and snow avalanche occurrence is inherently nonlinear, involving complex interactions and threshold effects across mountainous landscapes [51, 52]. Although both models are nonlinear, BRT frequently surpasses RF because its sequential boosting iteratively targets residuals, capturing subtle thresholds and local topographic variations more effectively than RF's parallel bagging.

Appropriate factor selection is essential for improving model performance and prediction accuracy. Of the 21 potential influencing factors initially considered, we retained 13 for snow avalanche modeling: TPI, VD, TWI, RSP, PLA, DTR, LS, AH, DTF, elevation, aspect, rainfall, and temperature. Notably, adding factors does not necessarily improve prediction accuracy, as redundant variables may introduce noise and reduce model performance. VRM, CI, curvature, and PRO characterize terrain roughness and surface geometry, which influence snow accumulation and redistribution but may provide information similar to TPI, RSP, VD, and PLA. Thus, their exclusion does not indicate a lack of physical relevance to avalanche occurrence, but rather limited additional predictive value in the modeling. In contrast, selecting relevant factors can improve model predictive performance. Previous studies have generally grouped the factors influencing snow avalanche occurrence into three main categories: terrain characteristics, meteorological conditions, and snow conditions [4, 12, 53]. Among these factors, elevation influences local climatic conditions, aspect affects the amount of solar radiation slopes receive, and TPI reflects local topographic position and may influence snow accumulation and redistribution. These factors can therefore play important roles in controlling snow avalanche occurrence [9, 54].

To evaluate the models' ability to predict areas prone to snow avalanche occurrence, we considered three polygon-based sampling scenarios for model training: the runout zone, the starting and track zones, and all three zones. As shown in Table 3, the models trained using the starting and track zone scenario outperformed those

trained under the other scenarios across all performance metrics. In contrast, the models trained using the runout zone scenario showed lower performance based on the evaluated statistical and accuracy metrics. In the runout zone scenario, most of the study area was classified as having very low susceptibility (BRT = 79.2%; RF = 78.5%). Because a large proportion of the study area consists of mountainous terrain, this classification suggests that the runout-based sampling strategy had limited ability to identify areas prone to snow avalanches. In contrast, polygon-based sampling from the starting and track zones improved the model performance and enabled more effective identification of areas with high avalanche susceptibility.

The higher performance in the starting-and-track-zone scenario can be attributed to the distinct roles of the three avalanche zones. The starting zone is directly associated with snowpack instability and avalanche release, while the track zone represents subsequent avalanche movement; thus, their environmental and topographic characteristics are more closely related to avalanche initiation and propagation. In contrast, the runout zone mainly represents deceleration and deposition and may reflect the consequences of avalanche movement rather than the conditions of avalanche initiation. This interpretation is consistent with Çetinkaya and Kocaman^[55], who investigated the effect of different learning sets (all zones and starting and track zones) on snow avalanche susceptibility mapping and reported better performance when they used the starting and track zones without the runout zone. Similarly, Akay^[14] described the starting, track, and runout zones as successive stages of avalanche movement,

from release and flow to final deposition. To demonstrate qualitative differences among the three scenarios, we selected the BRT model, which showed superior predictive performance (Figure 12)—based on a close-up of the points of interest shown in Figure 12A and B, polygon-based sampling from the starting and track zones improved the model's ability to identify areas of high and very high susceptibility, such as the Dizin ski resort. Conversely, models trained using the runout-zone scenario showed less ability to correctly identify areas affected by snow avalanches, as many mapped avalanche polygons fell outside the high and very high susceptibility classes. The findings indicated that slope, TWI, TPI, and elevation were among the most influential factors affecting snow avalanche occurrence. This may be attributed to the strong influence of terrain characteristics, particularly slope and elevation, on avalanche release and movement. Several studies corroborate this, indicating that steep slopes (28°–60°) are more prone to snow avalanches, while gentler slopes do not provide favorable conditions for snow avalanche occurrence [25, 56, 57].

By comparing the model predictions with the historical avalanche inventory, we found that both machine learning models (BRT and RF) performed better when trained using the starting-and-track-zone scenario, followed by the three-zone scenario, compared with the runout-zone scenario. Overall, the results demonstrate the usefulness of machine learning and remote sensing approaches for spatial snow avalanche susceptibility mapping, particularly in mountainous regions where historical avalanche inventories are limited [5, 58, 59].

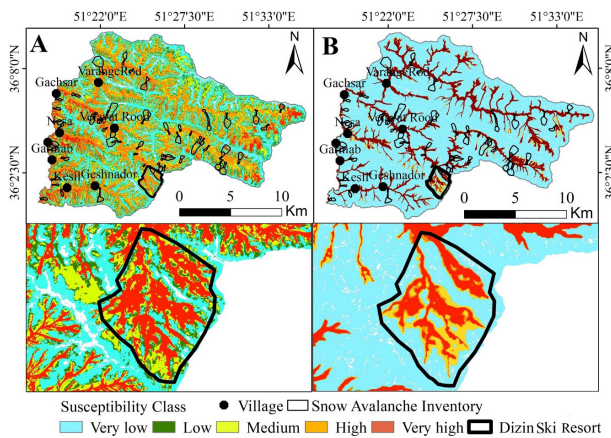


Figure 12) Snow avalanche susceptibility maps generated using the BRT model for scenario 3 (A) and scenario 1 (B).

Avalanche susceptibility mapping has traditionally relied on point-based data [8, 20, 21, 22], representing avalanches as single locations rather than spatially distributed processes. However, avalanches are not point phenomena; they initiate from a starting zone, travel through a track zone, and finally accumulate in a runout (deposition) zone, each covering a distinct area. Using single-point data without identifying these three stages oversimplifies avalanche dynamics and reduces the accuracy of susceptibility models. In this study, we addressed this limitation by utilizing polygon-based historical data, allowing for a more precise representation of avalanche-prone areas. Unlike point-based approaches, our method captures the full spatial extent of past avalanche occurrences, enhancing the accuracy and reliability of susceptibility predictions. The main advantage of this framework lies in the design and spatial representation of the training data, not in developing a new machine learning algorithm. Moreover, while many previous studies have not explicitly considered different avalanche zones, our framework incorporates the starting, track, and runout

areas, allowing systematic evaluation of their effects on susceptibility assessment. This approach provides more spatially informed susceptibility maps and can support avalanche hazard management and risk reduction. Nevertheless, further evaluation in other mountain environments and under different avalanche regimes is required to assess the transferability and generalizability of this framework. Future studies should apply bootstrapping to address uncertainty in the modeling results further.

An important limitation of this study is the lack of a systematic snow and avalanche monitoring network in the study area. Therefore, detailed and continuous snowpack observations, such as snow depth, SWE, and snowpack characteristics, were not available. Nevertheless, the models incorporated meteorological variables, including precipitation and temperature, because these variables influence snow accumulation, melt processes, and snowpack stability, thereby contributing to the environmental conditions associated with avalanche occurrence. Uncertainties also arise from the random splitting of training and validation datasets and the feature selection process, which can affect model stability and performance. Although we extracted many points from the polygons, increasing the number and spatial diversity of independent avalanche-prone polygons could provide more representative training data, reduce uncertainty, and potentially improve model robustness and generalizability. The high AUC values should nevertheless be interpreted with caution, as they may partly reflect the quality and spatial structure of the historical avalanche inventory. Although polygon-level splitting

reduced spatial leakage, residual spatial autocorrelation may still affect model performance. Independent validation using multi-year inventories or data from other watersheds would further strengthen the reliability and generalizability of the results.

Conclusion

This study evaluates avalanche susceptibility using machine learning models in the Central Alborz watershed, Alborz Province, Iran. The primary objective was to develop an effective training scheme for snow avalanche susceptibility mapping using machine learning techniques, specifically BRT and RF. A polygon-based framework was established that incorporated polygon-based representations of the three primary avalanche zones, including starting, track, and runout zones, across three distinct scenarios. We selected the influencing factors based on relevant literature, data availability, and the study area's specific characteristics. The models were trained using three scenarios: (1) using only the runout zone, (2) utilizing the starting and track zones, and (3) incorporating all three zones (starting, track, and runout). We then used statistical and accuracy metrics, including the Receiver Operating Characteristic (ROC) curve, precision, sensitivity, F1 score, and Matthews Correlation Coefficient (MCC), to evaluate model performance. The main conclusions of this study are as follows: (1) among the models evaluated, BRT generally demonstrated better predictive performance than RF across the considered evaluation metrics; (2) the application of polygon sampling for training enhanced the models' ability to delineate areas susceptible to snow avalanches; (3) models trained with polygon sampling from the starting and track zones

demonstrated a greater ability to delineate high-susceptibility areas compared with those trained using data from all three zones; and (4) overall, the results contribute to improving the effectiveness of spatial snow avalanche susceptibility mapping and provide valuable insights for avalanche hazard management.

Acknowledgments

The authors gratefully acknowledge the funding support of Ferdowsi University of Mashhad, Mashhad, Iran.

Author's Contributions: M KAZEMZADEH; Conceptualization, Sampling, Methodology, Analysis, Software, Writing, Review, Editing.

Ethical Permission: This study did not require ethical approval.

Funding/ Support: This work was supported by funding from Ferdowsi University of Mashhad, Iran [grant number 64099].

Conflict of Interest: The authors declare no conflicts of interest.

AI Use Declaration: The authors used ChatGPT-4 (OpenAI) in August 2026 solely to improve and polish the manuscript's language and readability. The authors reviewed, edited, and validated the content and take full responsibility for the final manuscript.

References

1. Viglietti D., Letey S., Motta R., Maggioni M., Freppaz M. Snow avalanche release in forest ecosystems: A case study in the Aosta Valley Region (NW-Italy). *Cold Reg. Sci. Technol.* 2010; 64(2):167-173.
2. Hao J., Zhang X., Cui P., Li L., Wang Y., Zhang G., Li C. Impacts of climate change on snow avalanche activity along a transportation corridor in the Tianshan Mountains. *Int. J. Disaster Risk Sci.* 2023; 14(4):510-522.
3. Ballesteros-Cánovas J.A., Trappmann D., Madrigal-González J., Eckert N., Stoffel M. Climate warming enhances snow avalanche risk in the

- Western Himalayas. *Proc. Natl. Acad. Sci. U.S.A.* 2018; 115(13):3410-3415
4. Schweizer J., Bartelt P., Van Herwijnen A. Snow avalanches. In: *Snow and Ice-Related Hazards, Risks, and Disasters*. Amsterdam: Elsevier. 2015; 395-436p.
 5. Durlević U., Valjarević A., Novković I., Vujović F., Josifov N., Krušić J., Ivanović M. Universal snow avalanche modeling index based on SAFI-Flow-R approach in poorly-gauged regions. *ISPRS Int. J. Geo-Inf.* 2024; 13(9):315.
 6. Ahmad I., Mallick J., Hang H.T., Rihan M., Bakshi A. An explainable Bayesian-optimized hybrid deep learning framework with attention mechanisms for modeling snow avalanche susceptibility. *J. Indian Soc. Remote Sens.* 2026; 54:1-25.
 7. Schweizer J., Jamieson J.B. Snow cover properties for skier triggering of avalanches. *Cold Reg. Sci. Technol.* 2001; 33(2-3):207-221.
 8. Yousefi S., Pourghasemi H.R., Emami S.N., Pouyan S., Eskandari S., Tiefenbacher J.P. A machine learning framework for multi-hazard modeling and mapping in a mountainous area. *Sci. Rep.* 2020;10(1):12144.
 9. Rafique A., Dasti M.Y., Ullah B., Awwad F.A., Ismail E.A., Saqib Z.A. Snow avalanche hazard mapping using a GIS-based AHP approach: A case of glaciers in Northern Pakistan from 2012 to 2022. *Remote Sens.* 2023; 15(22):5375.
 10. Wang H., Zhang L., Yin K., Luo H., Li J. Landslide identification using machine learning. *Geosci. Front.* 2021; 12(1):351-364.
 11. Rahman M., Chen N., Islam M.M., Mahmud G.I., Pourghasemi H.R., Alam M., Dewan A. Development of flood hazard map and emergency relief operation system using hydrodynamic modeling and machine learning algorithm. *J. Clean. Prod.* 2021; 311:127594.
 12. Bian R., Huang K., Liao X., Ling S., Wen H., Wu X. Snow avalanche susceptibility assessment based on ensemble machine learning model in the central Shaluli Mountain. *Front. Earth Sci.* 2022; 10: 880711.
 13. Tang Y., Sun Y., Han Z., Wu Q., Tan B., Hu C. Flood forecasting based on machine learning pattern recognition and dynamic migration of parameters. *J. Hydrol. Reg. Stud.* 2023; 47:101406.
 14. Akay H. Spatial modeling of snow avalanche susceptibility using hybrid and ensemble machine learning techniques. *Catena.* 2021; 206:105524.
 15. Viallon-Galinier L., Hagenmuller P., Eckert N. Combining snow physics and machine learning to predict avalanche activity: Does it help? *Cryosphere.* 2022;1-23.
 16. Liu J., Zhang T., Hu C., Wang B., Yang Z., Sun X., Yao S. A study on avalanche-triggering factors and activity characteristics in Aerxiangou, West Tianshan Mountains, China. *Atmosphere.* 2023; 14(9):1439.
 17. Bansal J.K., Goswami A., Roy A. Snow avalanche in the Indian Himalayas: Hazard zonation and climate change trends in Kullu region of Himachal Pradesh, India. *Phys. Chem. Earth.* 2025; 138: 103882.
 18. Abhinav A., Sattar A. Snow avalanche susceptibility, hazard, and exposure assessment in the Western Himalaya using machine learning and numerical modelling. *Sci. Rep.* 2025; 15: 38093.
 19. Tian X., Yao X., Zhou Z., Chen X., Song G. A progressive snow avalanche hazard assessment framework coupling interpretable machine learning with dynamic simulation. *Prog. Disaster Sci.* 2026; 31:100660.
 20. Choubin B., Borji M., Mosavi A., Sajedi-Hosseini F., Singh V.P., Shamshirband S. Snow avalanche hazard prediction using machine learning methods. *J. Hydrol.* 2019; 577:123929.
 21. Rahmati O., Ghorbanzadeh O., Teimurian T., Mohammadi F., Tiefenbacher J.P., Falah F., Bui D. Spatial modeling of snow avalanche using machine learning models and geo-environmental factors: Comparison of effectiveness in two mountain regions. *Remote Sens.* 2019; 11(24):2995.
 22. Yariyan P., Omidvar E., Karami M., Cerdà A., Pham Q.B., Tiefenbacher J.P. Evaluating novel hybrid models based on GIS for snow avalanche susceptibility mapping: A comparative study. *Cold Reg. Sci. Technol.* 2022; 194:103453.
 23. Liu Y., Chen X., Yang J., Li L., Wang T. Snow avalanche susceptibility mapping from tree-based machine learning approaches in ungauged or poorly-gauged regions. *Catena.* 2023; 224:106997.
 24. Kazemzadeh M., Noori Z. Analyzing the Factors Influencing Avalanche Occurrence Using Multivariate Statistical Methods (Case Study: Karaj Dam Watershed). *J. Geogr. Environ. Hazard.* 2026; 15(2): 118-140.
 25. McClung D., Schaerer P. *The Avalanche Handbook*. 4rd ed. Seattle: The Mountaineers Books; 2006; 368p.
 26. Voiculescu M., Ardelean F., Török-Oance M., Milian N. Topographical factors, meteorological variables and human factors in the control of the main snow avalanche events in the Făgăraș Massif (Southern Carpathians, Romanian Carpathians): Case studies. *Geogr. Pol.* 2016; 89(1):47-64.
 27. Voiculescu M., Ardelean F., Török-Oance M., Milian N. Topographical factors, meteorological variables and human factors in the control of the main snow avalanche events in the Făgăraș Massif (Southern Carpathians, Romanian Carpathians):

- Case studies. *Geogr. Pol.* 2016; 89(1):47-64.
28. Denissova N., Nurakynov S., Petrova O., Chepashev D., Daumova G., Yelisseyeva A. Remote sensing techniques for assessing snow avalanche formation factors and building hazard monitoring systems. *Atmosphere.* 2024; 15(11):1343.
 29. Fazzini M., Cordeschi M., Carabella C., Paglia G., Esposito G., Miccadei E. Snow avalanche assessment in mass movement-prone areas: Results from climate extremization in relationship with environmental risk reduction in the Prati di Tivo area (Gran Sasso Massif, Central Italy). *Land.* 2021;10(11):1176.
 30. Dormann C.F., Elith J., Bacher S., Buchmann C., Carl G., Carré G., Lautenbach S. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography.* 2013; 36(1):27-46.
 31. Habumugisha J.M., Chen N., Rahman M., Islam M.M., Ahmad H., Elbeltagi A., Dewan A. Landslide susceptibility mapping with deep learning algorithms. *Sustainability.* 2022; 14(3):1734.
 32. Guns M., Vanacker V. Logistic regression applied to natural hazards: rare event logistic regression with replications. *Nat. Hazard. Earth Syst. Sci.* 2012; 12(6):1937-1947.
 33. Lei X., Chen W., Avand M., Janizadeh S., Kariminejad N., Shahabi H., Mosavi A. GIS-based machine learning algorithms for gully erosion susceptibility mapping in a semi-arid region of Iran. *Remote Sens.* 2020; 12(15):2478.
 34. Chen W., Chen Y., Tsangaratos P., Ilia I., Wang X. Combining evolutionary algorithms and machine learning models in landslide susceptibility assessments. *Remote Sens.* 2020; 12(23):3854.
 35. Darst B.F., Malecki K.C., Engelman C.D. Using recursive feature elimination in random forest to account for correlated variables in high-dimensional data. *BMC Genet.* 2018; 19 (Suppl 1):65.
 36. Liaw A., Wiener M. Classification and regression by random Forest. *R New.* 2002; 2(3):18-22.
 37. Hastie T., Tibshirani R., Friedman J. The elements of statistical learning: Data mining, inference, and prediction. 2nd ed. New York: Springer. 2009; 764.
 38. Bui D.T., Tuan T.A., Klempe H., Pradhan B., Revhaug I. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides.* 2016; 13(2): 361-378.
 39. Breiman L. Random forests. *Mach. Learn.* 2001; 45(1):5-32.
 40. Robinzonov N. Advances in boosting of temporal and spatial models. PhD thesis. Ludwig-Maximilians-Universität München. 2013;135p
 41. Elith J., Leathwick J.R., Hastie T. A working guide to boosted regression trees. *J. Anim. Ecol.* 2008; 77(4):802–813.
 42. Youssef A.M., Pourghasemi H.R., Pourtaghi Z.S., Al-Katheeri M.M. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides.* 2016; 13(5):839-856.
 43. Schapire R.E. The boosting approach to machine learning: An overview. In: *Nonlinear Estimation and Classification.* New York: Springer. 2003; 149-171p.
 44. Fawcett T. An introduction to ROC analysis. *Pattern Recognit. Lett.* 2006; 27(8):861-874.
 45. Nachappa T.G., Ghorbanzadeh O., Gholamnia K., Blaschke T. Multi-hazard exposure mapping using machine learning for the State of Salzburg, Austria. *Remote Sens.* 2020;12(17):2757.
 46. Ullah K., Wang Y., Fang Z., Wang L., Rahman M. Multi-hazard susceptibility mapping based on convolutional neural networks. *Geosci. Front.* 2022;13(5):101425.
 47. Sinčić M., Bernat Gazibara S., Rossi M., Krkač M., Mihalčić Arbanas S. A comprehensive comparison of stable and unstable area sampling strategies in large-scale landslide susceptibility models using machine learning methods. *Remote Sens.* 2024; 16(16):2923.
 48. Frattini P., Crosta G., Carrara A. Techniques for evaluating the performance of landslide susceptibility models. *Eng. Geol.* 2010; 111(1-4):62-72.
 49. Ziarh G.F., Asaduzzaman M., Dewan A., Nashwan M.S., Shahid S. Integration of catastrophe and entropy theories for flood risk mapping in Peninsular Malaysia. *J. Flood Risk Manag.* 2021; 14(1):e12686.
 50. Rainio O., Teuho J., Klén R. Evaluation metrics and statistical tests for machine learning. *Sci. Rep.* 2024; 14(1):6086.
 51. Matthews B.W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta Protein Struct.* 1975; 405(2): 442-451.
 52. Zheng H., Ding M., Huang T., He Y., Gao Z., Duan Y. Integrating the frequency ratio and index of entropy with an extreme learning machine to map post-earthquake landslide susceptibility: A case study in Xingwen County, China. *Stoch. Environ. Res. Risk Assess.* 2024; 39(10): 4749-4771.
 53. Li Y., Yang Z., Cheng Q., Qiang X., Liu J. Avalanche Susceptibility Mapping with Explainable Machine Learning: A Case Study of the Kanas Scenic

- Transportation Corridor in the Altay Mountains, China. *Appl. Sci.* 2025; 15(21): 11631.
54. Mosavi A., Shirzadi A., Choubin B., Taromideh F., Hosseini F.S., Borji M., Dineva, A.A. Towards an ensemble machine learning model of random subspace-based functional tree classifier for snow avalanche susceptibility mapping. *IEEE Access*.2020;8(1):145968-145983.
 55. Varol N. Avalanche susceptibility mapping with the use of frequency ratio, fuzzy, and classical analytical hierarchy process for Uzungol area, Turkey. *Cold Reg. Sci. Technol.* 2022; 194:103439.
 56. Cetinkaya S., Kocaman S. Impact of learning set and sampling for snow avalanche susceptibility mapping with random forest. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* 2023; 48(1):57–64.
 57. McClung D.M. Effects of triggering mechanism on snow avalanche slope angles and slab depths from field data. *Nat. Hazard.* 2013; 69(3):1721-1731.
 58. Yang J., He Q., Liu Y. Winter-spring prediction of snow avalanche susceptibility using optimisation multi-source heterogeneous factors in the Western Tianshan Mountains, China. *Remote Sens.* 2022;14(6):1340.
 59. Korzeniowska K., Bühler Y., Marty M., Korup O. Regional snow-avalanche detection using object-based image analysis of near-infrared aerial imagery. *Nat. Hazards Earth Syst. Sci.* 2017; 17(10):1823-1836.
 60. Gao M., Gu X., Liu Y., Zhan Y., Wei X., Yu H., Ding Y. An improved spatiotemporal data fusion method for snow-covered mountain areas using snow index and elevation information. *Sensor.* 2022; 22(21):8524.