



# Game-theoretic optimization of landslide susceptibility mapping: a comparative study between Bayesian-optimized basic neural network and new generation neural network models

Javed Mallick<sup>1</sup> · Meshel Alkahtani<sup>1</sup> · Hoang Thi Hang<sup>2</sup> · Chander Kumar Singh<sup>3</sup>

Received: 9 September 2023 / Accepted: 25 March 2024 / Published online: 9 April 2024  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

## Abstract

Landslide susceptibility mapping is essential for reducing the risk of landslides and ensuring the safety of people and infrastructure in landslide-prone areas. However, little research has been done on the development of well-optimized Elman neural networks (ENN), deep neural networks (DNN), and artificial neural networks (ANN) for robust landslide susceptibility mapping (LSM). Additionally, there is a research gap regarding the use of Bayesian optimization and the derivation of SHapley Additive exPlanations (SHAP) values from optimized models. Therefore, this study aims to optimize DNN, ENN, and ANN models using Bayesian optimization for landslide susceptibility mapping and derive SHAP values from these optimized models. The LSM models have been validated using the receiver operating characteristics curve, confusion matrix, and other twelve error matrices. The study used six machine learning-based feature selection techniques to identify the most important variables for predicting landslide susceptibility. The decision tree, random forest, and bagging feature selection models showed that slope, elevation, DFR, annual rainfall, LD, DD, RD, and LULC are influential variables, while geology and soil texture have less influence. The DNN model outperformed the other two models, covering 7839.54 km<sup>2</sup> under the very low landslide susceptibility zone and 3613.44 km<sup>2</sup> under the very high landslide susceptibility zone. The DNN model is better suited for generating landslide susceptibility maps, as it can classify areas with higher accuracy. The model identified several key factors that contribute to the initiation of landslides, including high elevation, built-up and agricultural land use, less vegetation, aspect (north and northwest), soil depth less than 140 cm, high rainfall, high lineament density, and a low distance from roads. The study's findings can help stakeholders make informed decisions to reduce the risk of landslides and ensure the safety of people and infrastructure in landslide-prone areas.

**Keywords** Landslide susceptibility · Neural network · Bayesian optimization · SHapley Additive exPlanations (SHAP) · Landslide risk mitigation

## Introduction

Landslides constitute a significant geomorphological hazard in mountainous regions around the world and pose considerable risks to human life, infrastructure, and economic stability. The annual economic impact of these events is

Responsible Editor: Philippe Garrigues

✉ Javed Mallick  
jmallick@kku.edu.sa  
Meshel Alkahtani  
mgalqhtani@kku.edu.sa  
Hoang Thi Hang  
hllhangstac@gmail.com  
Chander Kumar Singh  
chander.singh@terisas.ac.in

- <sup>1</sup> Department of Civil Engineering, College of Engineering, King Khalid University, P.O. Box: 394, Abha 61411, Kingdom of Saudi Arabia
- <sup>2</sup> Department of Geography, Faculty of Natural Science, Jamia Millia Islamia, New Delhi, India
- <sup>3</sup> Department of Energy and Environment, Analytical and Geochemistry Laboratory, TERI School of Advanced Studies, New Delhi, India

considerable as they affect roads, railways, and buildings (Arabameri et al. 2022; Achour and Pourghasemi 2020; Billaal et al. 2019). The occurrence and severity of landslides are influenced by a combination of natural phenomena such as seismic activity and extreme meteorological events, as well as anthropogenic factors such as deforestation and uncontrolled urban expansion (Haque et al. 2016; Sahin 2022; Kavzoglu et al. 2019; Althuwaynee et al. 2014). In India, especially in the Himalayan belt, landslides pose a formidable geological challenge with far-reaching socio-economic and environmental impacts. The state of Uttarakhand, nestled in the Indian Himalayas, is particularly prone to such events due to its unique geological and climatic conditions, which include frequent heavy rainfall and seismic activity. This susceptibility is exacerbated by human activities such as road construction and urbanization, making Uttarakhand a focus area for landslide research (Awasthi et al. 2022; Kaur et al. 2018; Martha et al. 2021; Kumar et al. 2017). To minimize landslide damage, landslide susceptibility mapping (LSM) must be used as a cornerstone of the hazard mitigation strategy. LSM provides important insights for preventive measures, ongoing monitoring, and informed land-use planning and thus plays a crucial role in reducing the damaging effects of these natural disasters (Razavi et al. 2021; Shi et al. 2017).

LSM is pivotal for managing landslide hazards, employing a wide array of analytical techniques from classical statistical methods to cutting-edge machine learning (ML) and deep learning (DL) models (Mallick et al. 2021). Traditional statistical methods like frequency ratio, principal component analysis (PCA), and weight of evidence (WOE) offer simplicity and interpretability but may not fully capture the intricate, non-linear dynamics of landslide data (Saha and Saha, 2020). While index-based methods like entropy index and logistic regression systematically assess susceptibility, they can overlook the subtle spatial interactions among landslide predictors (Alqadhi et al. 2022a, b, c). Recent advancements have seen ML algorithms such as random forest, support vector machine (SVM), and XGBoost gaining prominence due to their ability to manage complex datasets and variable interplays (Alqadhi et al., 2022a; Mallick et al., 2022). However, these methods are not without criticism, particularly for their tendencies toward overfitting and challenges in handling large data volumes, which can obscure their decision-making processes. To overcome these limitations, DL models with sophisticated architectures have been introduced, showing promise in deciphering complex data patterns and offering enhanced predictive accuracy, particularly with unstructured big data like images (Mahato et al., 2021; Zhang et al., 2019). Despite these advancements, the quest for the most effective LSM method remains unresolved, with a notable absence of consensus in regions with complex geological backgrounds like Uttarakhand, known

for its frequent landslides. While some studies (Dey et al., 2021; Gupta et al., 2016) have explored GIS and ML integration for LSM at local levels, comprehensive state-wide analyses remain scarce, underscoring the need for extensive research to inform disaster management strategies.

The success of LSM methods depends, to a considerable extent, on factors such as the choice and resolution of mapping units, algorithm selection and parameter optimization, sample selection, and identification of relevant influencing factors. The choice of mapping units is pivotal in LSM, as it directly influences the model's ability to accurately predict landslide susceptibility. In this study, we have opted for grid units with a resolution of 30 m, a decision driven by the balance between detail and practicality at the state level. High-resolution mapping, while offering finer detail, poses significant challenges due to the extensive computational resources required and the frequent unavailability of high-resolution data across large areas. A 30-m resolution strikes an optimal balance, ensuring sufficient detail for effective LSM while maintaining feasibility for state-wide modeling. This approach aligns with findings from previous studies which highlight the importance of selecting appropriate mapping units and resolutions to optimize LSM performance (Sun et al., 2023; Liao et al., 2022). In the quest for optimal hyperparameters within LSM models, various optimization techniques have been explored. Bayesian optimization has emerged as a particularly effective method, as evidenced by its successful implementation in previous studies (Sun et al., 2021a; Sun et al., 2020a, c). Unlike traditional optimization methods, Bayesian optimization excels in navigating complex optimization landscapes by efficiently balancing the exploration of new parameter spaces and the exploitation of known high-performing areas. This characteristic makes it especially suitable for optimizing deep learning models where the evaluation of model performance can be computationally expensive and time-consuming.

Building on the success of Bayesian optimization in previous research, this study aims to develop new DL models that leverage this optimization technique. By integrating Bayesian optimization, we aim to refine the selection of hyperparameters for artificial neural networks (ANN), deep neural networks (DNN), and Elman neural networks (ENN). This integration is anticipated to enhance the models' ability to capture the non-linear relationships between landslide conditional factors and landslide occurrence data, particularly within the chosen 30-m grid resolution framework. The rationale for this approach is supported by studies demonstrating the significant impact of mapping units on LSM performance and the proven effectiveness of Bayesian optimization in achieving optimal model configurations (Zhang et al., 2021; Sun et al., 2021b; Zhou et al., 2021). A novel aspect of this research is the integration of SHapley Additive

exPlanations (SHAP) values into the LSM models, which deviates from the traditional use of XGBoost models for SHAP calculation (Lundberg & Lee, 2017). By quantifying the contribution of each feature to the model's predictions, SHAP values provide interpretability, solving the challenges associated with the reliability and understanding of landslide risk assessments. This integrated methodology is suited to address the limitations highlighted in recent studies (Zhang et al., 2023; Liu et al., 2023; Rihan et al., 2023; Yu & Chen, 2024; Wang et al., 2024), which have primarily focused on model performance rather than interpretability and regional applicability.

Therefore, this research contributes to the LSM field by employing Bayesian optimization for hyperparameter tuning, an approach that has not been extensively explored in this context. By incorporating the SHAP value calculation into the modeling process, this study aims to provide a more accurate and reliable understanding of the determinants that influence landslide susceptibility. This study not only addresses the need for state-level LSM in Uttarakhand but also presents an innovative methodology that could shape future research and practical applications in landslide risk management and ultimately help stakeholders make informed decisions to strengthen the safety and resilience of communities and infrastructure in landslide-prone regions.

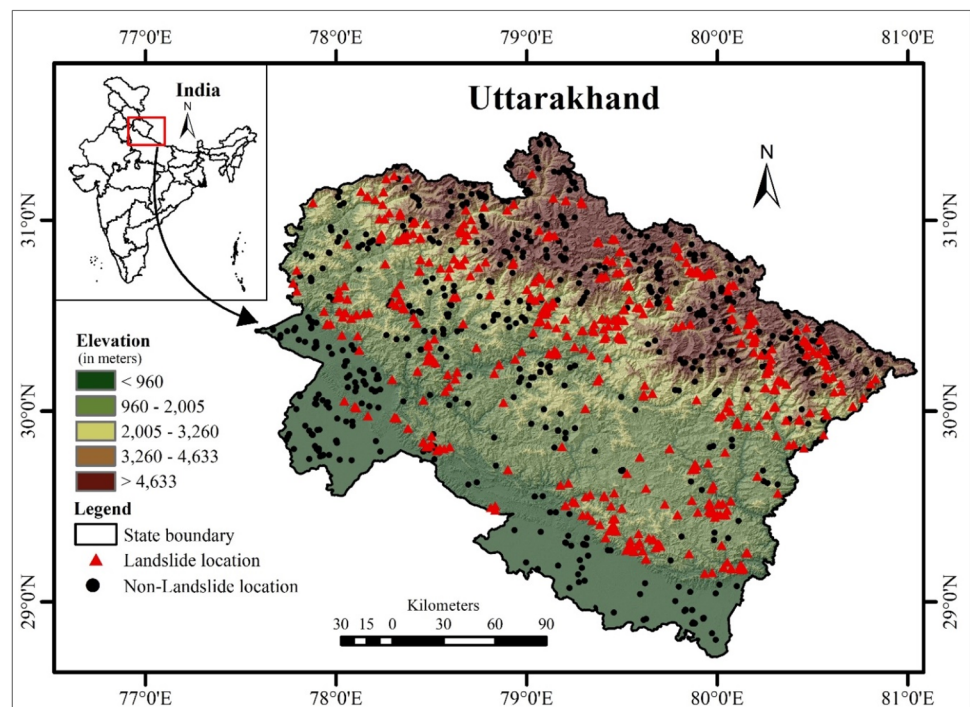
## Materials and methodology

### Study area

Uttarakhand is located in the Indian Himalayan range, between latitudes 28°43' N and 31°28' N and longitudes 77°34' E and 81°03' E. The state covers an area of approximately 53,483 km<sup>2</sup>, which is 1.63% of the country's total land area (refer to Fig. 1). As of the 2011 census, the population of Uttarakhand was approximately 10 million, and this population is evenly distributed across all 13 districts of the state. The two most populous districts in the state are Dehradun and Haridwar, with a combined population of over 3.59 million. The study area of this research is classified as a mountainous region with elevations ranging from 250 to 7817 m above mean sea level. The geology of Uttarakhand is complex and comprises of folded igneous, sedimentary, and metamorphic rocks. The state is geographically divided into six regions: the Great Himalaya, the Lesser Himalaya, the Shivalik Himalaya, the Upper Ganga Plain, the Terai region, and the Bhabar region (Anamika et al., 2020).

Uttarakhand has a subtropical monsoon climate, with an average annual rainfall ranging between 920 and 2370 mm. The monsoon season usually occurs from July to August and brings persistent and heavy rainfall. The maximum temperature typically reaches about 40 °C in May and June, while

**Fig. 1** Location of the study area and some landslide spots



the minimum temperature drops to  $-7^{\circ}\text{C}$  in December and January (Dobriyal and Bijalwan, 2017). The vegetation and climate in this region vary significantly with altitude, from the Ganga plains to the high-altitude glaciers.

## Materials

This study area experiences recurrent landslides annually, and to obtain the historical landslide inventory for this study, data was gathered from the Geological Survey of India (<https://bhukosh.gsi.gov.in/Bhukosh/Public>) (Table 1). Satellite images of Landsat-9 (OLI-2/TIRS-2) were downloaded from the United States Geological Survey (USGS) website (<https://earthexplorer.usgs.gov/>) to generate a land use and land cover (LULC) map and normalized differential vegetation index (NDVI). The topographic and hydrological features of the study area were extracted from the SRTM map with a spatial resolution of 30 m. Soil depth and soil texture were downloaded from the Bhuvan Portal, India (<https://bhuvan.nrsc.gov.in/home/index.php>). The annual rainfall data for the study area was collected from the India Meteorological Department (IMD).

## Creation of landslide inventory

The landslide inventory, essential for developing the landslide susceptibility map through Bayesian-optimized deep learning techniques, was compiled from historic landslide data sourced from the Bhukosh website, a portal managed by the Geological Survey of India. This portal provides spatial data on landslide occurrences across India in polygon format. For this study, 1600 random landslide incidents were selected from the Bhukosh dataset and converted from polygons to point representations to serve as the target variable for model training and validation. It is noteworthy to mention that the Bhukosh portal, while a rich source of spatial data on landslides, does not provide specific temporal information for each landslide event. Consequently, the dataset used in this study encompasses landslide locations without explicit temporal details. These landslide points were further

validated through field investigations, GPS surveys, and the review of various landslide reports and articles, ensuring their accuracy and relevance for the susceptibility modeling. Alongside the 1600 landslide points, an equivalent number of non-landslide locations, which have not experienced any recorded landslides for a substantial period, were identified to create a balanced binary classification model from Google Earth Imageries and field survey. These non-landslide samples were selected based on the absence of landslide events in historical records and were included to improve the robustness of the predictive models. For the modeling process, the dataset of 3200 points was divided into training and testing sets, with 70% of the points used for model training and the remaining 30% allocated for model testing. This segmentation facilitated the effective training and evaluation of the Bayesian-optimized deep learning models, which were then applied to develop comprehensive landslide susceptibility maps. The lack of temporal data in the landslide inventory is a limitation acknowledged in this study, and future research could benefit from incorporating time-specific landslide occurrences to enhance the predictive capabilities of susceptibility models.

## Preparation of landslide conditioning parameters

The accuracy of landslide prediction depends largely on the quality of the prediction parameters (Azarafza et al., 2021). In this study, sixteen landscape characteristic parameters (LCPs) were employed to predict landslide susceptibility and map the study area. These LCPs include elevation, slope, curvature, aspect, annual rainfall, distance from river (DFR), drainage density (DD), lineament density (LD), distance from built-up (DFB), topographic wetness index (TWI), road density (RD), land use/land cover (LULC), normalized difference vegetation index (NDVI), soil depth, soil texture, and geology, each of which will be detailed below.

Firstly, elevation is a critical component of topography and a crucial indicator for detecting changes in relative relief as it reveals the highest and lowest points of elevation (Nawazelibe et al., 2022; Liu et al., 2021). The study

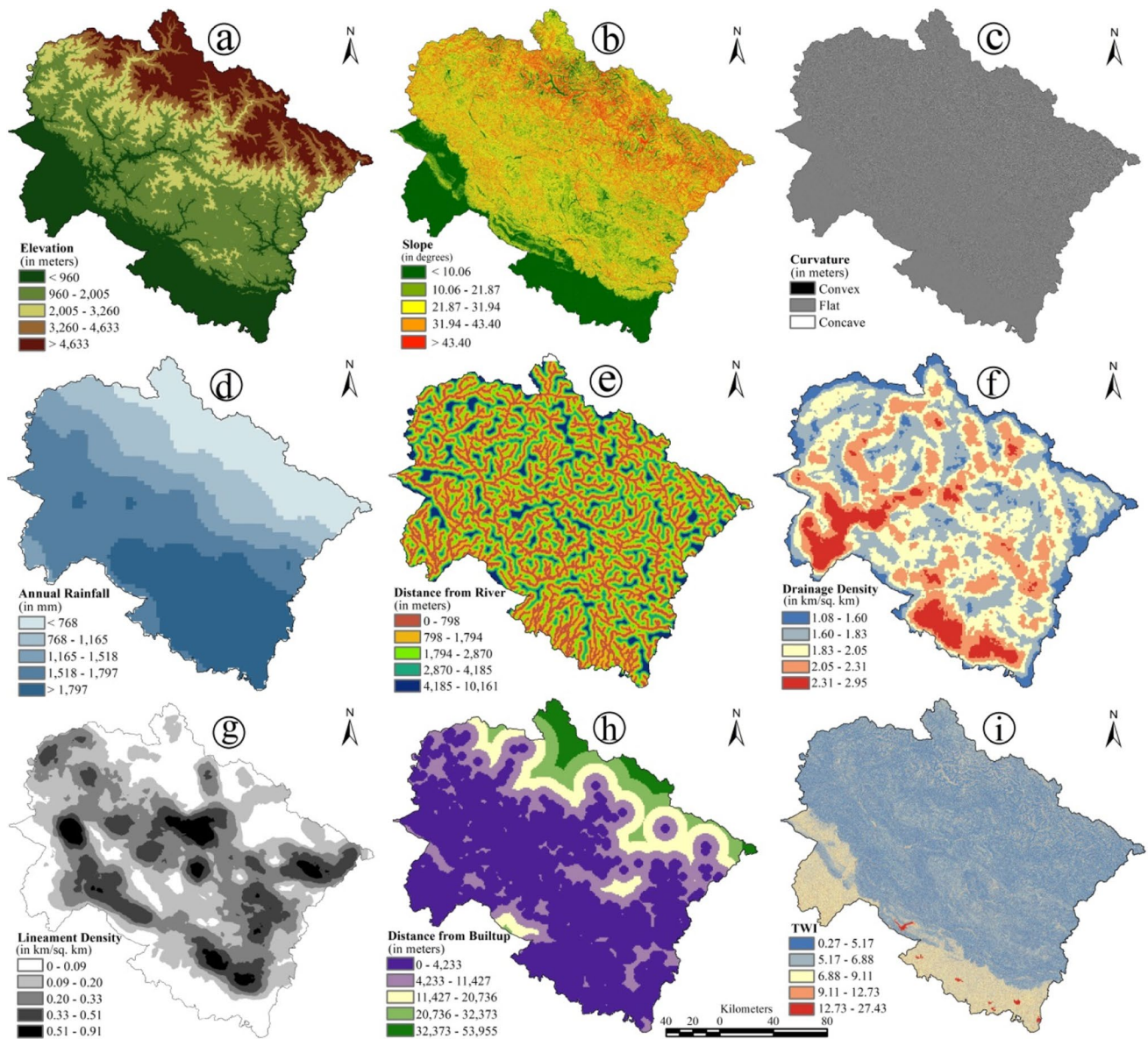
**Table 1** Overview of data sources and their utilization in landslide susceptibility modeling

Data type	Source	Purpose
Historical landslide inventory	Geological Survey of India ( <a href="https://bhukosh.gsi.gov.in/Bhukosh/Public">https://bhukosh.gsi.gov.in/Bhukosh/Public</a> )	To obtain the historical landslide inventory for the study
Landsat-9 (OLI-2/TIRS-2) satellite images	United States Geological Survey ( <a href="https://earthexplorer.usgs.gov/">https://earthexplorer.usgs.gov/</a> )	To generate a land use and land cover (LULC) map and NDVI
SRTM map	SRTM map with a spatial resolution of 30 m	To extract topographic and hydrological features of the study area
Soil depth and texture	Bhuvan Portal, India ( <a href="https://bhuvan.nrsc.gov.in/home/index.php">https://bhuvan.nrsc.gov.in/home/index.php</a> )	To gather soil depth and texture information
Annual rainfall data	India Meteorological Department (IMD)	To collect annual rainfall data for the study area



area’s elevation map was divided into six groups using the natural break categorization tool, namely below 960, 960 to 2005, 2005 to 3260, 3260 to 4633, and above 4633 m (Fig. 2a). Secondly, the slope angle is one of the most significant factors in predicting landslide susceptibility as steeper slopes are more prone to failure (Bui et al., 2020). The slope angle was manually grouped into five categories, namely < 10.06°, 10.06–21.87°, 21.87–31.94°, 31.94–43.40°, and > 43.40° (Fig. 2b). Thirdly, curvature refers to the shape of a surface created by the intersection

of random planes (Canavesi et al., 2020; Gautam, 2022). Plan curvature controls flow divergence and convergence and is perpendicular to the direction of the steepest slope (Sameen et al., 2020). The study area was divided into one of three groups based on curvature values using the natural break classification tool, namely convex, flat, and concave (Fig. 2c). Fourthly, aspect is another significant factor that affects landslide susceptibility in the study area as it reveals the directions of slopes and regulates the quantity of water on hillsides, both of which can



**Fig. 2** Multifaceted geospatial parameters for landslide susceptibility mapping, such as **a** elevation variance across the study area, **b** gradient slope categorization, **c** surface curvature classification, **d** spatial distribution of annual rainfall, **e** proximity to river networks, **f** drainage density heatmap, **g** lineament density overlay, **h** distance from built-up areas, and **i** topographic wetness index (TWI). Integrated

environmental and anthropogenic factors for landslide susceptibility assessment, such as **j** road density distribution, **k** land use and land cover (LULC) classification, **l** normalized differential vegetation index (NDVI) analysis, **m** soil depth variation, **n** soil texture types, **o** geological formations with fault lines, and **p** aspect orientation

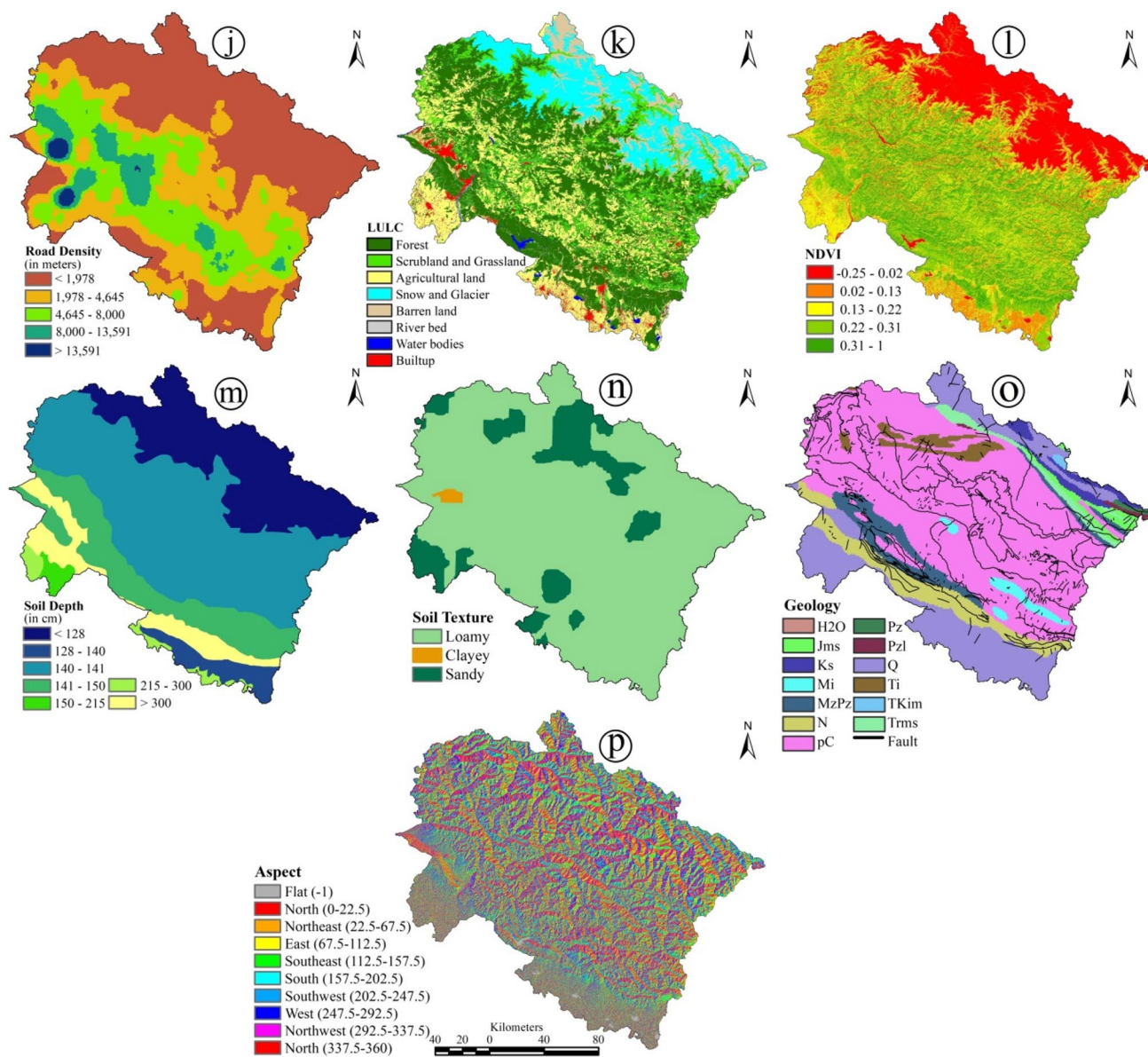


Fig. 2 (continued)

contribute to slope instability (Bui et al., 2020). The present study manually classified aspects into nine categories (Fig. 2p). Fifthly, annual rainfall is a crucial determinant of landslide risk as it increases soil wetness and pore water pressure, which decreases soil cohesiveness and triggers landslides (Yang et al., 2020; Zhu et al., 2020; Chen et al., 2020). The average annual rainfall in the study area ranges from 768 to 1797 mm (Fig. 2d). Then, groundwater flow toward streams and rivers is known to influence undercutting processes, which in turn cause landslides on valley slopes (Zaruba and Mencl, 2014; Tang et al., 2011). The distance from the river was calculated using ArcGIS and divided into five classes, namely 0–798 m, 798–1794

m, 1794–2870 m, 2870–4185 m, and 4185–10161 m (Fig. 2e). Next, drainage density fluctuates based on rock permeability, soil infiltration, precipitation intensity, and slope gradient (Arulbalaji et al., 2019), which significantly affects vegetation cover and landforms in a region (Gao et al., 2022) and contributes to landslides. In this study, the drainage density was calculated in ArcGIS and classified into five classes: 1.08–1.60 km<sup>2</sup>, 1.60–1.83 km<sup>2</sup>, 1.83–2.05 km<sup>2</sup>, 2.05–2.31 km<sup>2</sup>, and 2.31–2.95 km<sup>2</sup> (Fig. 2f).

LD is a measure of the density of lineaments in a particular region. To calculate the LD, the number of lineaments recorded in the region is divided by the total area under consideration. The lineaments are identified



by analyzing the Geological Survey of India website, which helps to identify fault lines and fractures. The presence of lineaments is then used to create a density map, which is classified into five different classes based on their density in  $\text{km}^2$  (Fig. 2g). Distance from built-up (DFB) is an important parameter for landslide susceptibility calculations and estimates. The distance between the research area and built-up areas is crucial for determining the region's vulnerability to landslides. The distance from built-up areas in the research region varies from 0 to 53,955 m (Fig. 2h). Topographic wetness index (TWI) is a measure of the geographical scale on hydrological processes. Topographical variables play a significant role in governing hydrology, which in turn affects pore water pressure and slope stability. In this study, the TWI value varies from 0.27 to 27.43 (Fig. 2i). The proximity of roadways to slopes can increase the susceptibility of the slopes to landslides. In this study, the road density was calculated using ArcGIS and divided into five classes based on distance: < 1978 m, 1978–4645 m, 4645–8000 m, 8000–13,591 m, and > 13,591 m (Fig. 2j). Land use/land cover (LULC) is an important factor in determining slope stability. Landslides are commonly associated with the development of roads, houses, bridges, dams, deforestation, and agricultural land growth. For this study, the LULC map was classified into eight classes: forest, scrubland and grassland, agricultural land, snow and glaciers, barren land, river bed, water bodies, and built-up (Fig. 2k). Vegetation is a critical factor in slope stability. The NDVI is used to quantify vegetation cover and was extracted from Landsat 9 satellite data using the NDVI equation (Eq. 1). The high value (0.22 to 1) represents vegetation, while the low value (−0.25 to 0.22) indicates a lack of vegetation (Fig. 2l).

Soil depth—soil depth is an important factor in predicting LSM. The soil depth map of the study region shows that the depth ranges from about 128 to 300 cm (Fig. 2m). Soil texture—the soil in the research region has a combination of loamy, clayey, and sandy characteristics, with loamy soils being the most dominant, followed by sandy soils in various patches (Fig. 2n). Geology plays a significant role in determining the frequency of landslides due to variations in the strength and permeability of rocks and soils caused by lithological and structural variables. The geological map of the study area is shown in Fig. 2o.

## Method for feature selection

Feature selection is a crucial step in building effective machine learning models, as it involves identifying the most relevant features that contribute to the predictive power of the model. This subsection details the implementation of various feature selection methods used in our study.

### Decision tree

Decision trees are employed for feature selection by evaluating the importance of each feature in making the split decisions. During the tree construction, the algorithm calculates the decrease in node impurity (e.g., Gini impurity for classification, variance for regression) for each feature at every split. Features that result in significant impurity reduction are considered more important. In our implementation, we utilize the feature importance scores generated by the decision tree to select the top features that contribute most to the model's predictive accuracy.

### Random forest

Random forest extends the feature selection capability of decision trees by averaging the feature importance scores across all trees in the forest. Since each tree is built on a random subset of features, this method provides a more robust estimation of feature importance. We aggregate the feature importance scores from all the trees within the random forest to identify and select the most significant features for our model.

### Bagging

While bagging itself is more of an ensemble method than a direct feature selection technique, we leverage the individual base estimators (e.g., decision trees) within the bagging ensemble to assess feature importance. Similar to random forest, we average the feature importance scores across all base estimators in the bagging ensemble to determine the key features. This approach helps in reducing the variance of the feature importance estimates, making the feature selection process more stable.

### LightGBM

LightGBM employs a novel feature selection approach by constructing gradient-boosted trees using a leaf-wise growth strategy rather than a level-wise strategy. Feature importance in LightGBM is determined based on the number of times a feature is used in splitting the nodes, weighted by the gain of each split. This method allows us to identify features that not only frequently contribute to splits but also significantly improve the model's performance.

### XGBoost

XGBoost provides a quantitative measure of feature importance based on the number of times a feature appears in the trees across all the boosting rounds, weighted by the gain

associated with each feature split. This approach allows for a comprehensive understanding of how each feature contributes to the model's predictions, enabling the selection of the most impactful features.

## SVM

Feature selection with SVM involves analyzing the weight coefficients associated with each feature in the hyperplane equation. In linear SVM, features with larger absolute weight values are considered more important as they have a greater influence on the decision boundary. For non-linear classification, the selection of features can be more complex and may involve techniques such as recursive feature elimination (RFE) with SVM where features are recursively removed based on their weights in the SVM model.

## Method for LSM using Bayesian-optimized deep learning models

### Bayesian optimization

Bayesian optimization is a probabilistic model-based approach for global optimization of black-box functions that are expensive to evaluate. It is particularly effective for hyperparameter tuning in complex machine-learning models where direct optimization is computationally infeasible. Bayesian optimization employs the Gaussian process (GP) to model the objective function and utilizes an acquisition function to decide where to sample next. This process iteratively updates the GP model based on observed evaluations and selects new hyperparameters to evaluate by optimizing the acquisition function. In the context of landslide susceptibility modeling, Bayesian optimization is used to systematically explore the hyperparameter space of ANN, DNN, and ENN models, aiming to find the optimal configuration that maximizes model performance, typically measured in terms of prediction accuracy or minimizing loss on a validation dataset.

### Artificial neural network

Artificial neural networks (ANNs) are computational models inspired by the structure and functional aspects of biological neural networks. They consist of layers of interconnected nodes or neurons where each connection represents a weight. In the context of landslide susceptibility modeling, an ANN takes various geographical and environmental features as input, such as slope angle, aspect, soil type, vegetation cover, and hydrological factors. These inputs are processed through one or more hidden layers where the network learns to identify complex patterns and relationships through the adjustment of weights during training. The final layer produces

a binary output indicating the susceptibility to landslides. Training an ANN involves using backpropagation and optimization algorithms like stochastic gradient descent (SGD) or Adam to minimize a loss function, typically binary cross-entropy for classification problems.

### Deep neural network

Deep neural networks (DNNs) extend ANNs by incorporating multiple hidden layers, allowing for the hierarchical extraction and learning of high-level features from input data. In landslide susceptibility applications, DNNs can model the intricate, non-linear interactions between various factors contributing to landslide occurrence, such as geological, topographical, and meteorological conditions. Each layer in a DNN transforms its input data into a slightly more abstract and composite representation, capturing complex patterns that simpler models might miss. Implementing DNNs for landslide susceptibility involves careful design considerations, such as the number of layers, the number of neurons in each layer, activation functions, and regularization techniques like dropout to prevent overfitting. The training process uses advanced optimization techniques to adjust the weights and biases to minimize prediction errors.

### Elman neural network

Elman neural networks (ENN), a type of recurrent neural network (RNN), are particularly suited for modeling temporal or sequential data, making them ideal for incorporating time-dependent factors into landslide susceptibility modeling, such as sequential rainfall data or temporal land-use changes. ENNs feature a set of recurrently connected hidden units that provide memory capabilities, allowing the network to maintain information about previous inputs in its internal state. This feature enables ENNs to capture dynamic temporal behaviors in the data, offering a significant advantage over traditional feedforward neural networks when dealing with time-series or spatial-temporal data. Implementing ENNs for landslide susceptibility involves selecting the appropriate architecture, including the number of hidden layers and units, and training the model to learn temporal patterns that are indicative of landslide occurrences.

### Implementation

The implementation of ANN, DNN, and ENN models for landslide susceptibility modeling involves a systematic approach that begins with data preprocessing, including normalization or standardization of input features to ensure effective model training. The models are then trained using the training dataset, comprising various landslide-related features and labels indicating landslide occurrences. The



training process involves forward propagation to make predictions, calculation of loss using a suitable loss function, and backward propagation to update the model weights. Model validation is conducted using a separate validation set to tune hyperparameters and avoid overfitting. Once trained, the models can classify unseen geographical areas into categories based on their susceptibility to landslides. Model performance is evaluated using metrics such as accuracy, precision, recall, and F1 score. The trained models serve as powerful tools for identifying potential landslide-prone areas, contributing to risk assessment, early warning systems, and informed decision-making for disaster management and mitigation strategies.

### Validation and comparison of the models

Model validation and comparison are critical steps in assessing the performance and reliability of machine learning models. In the context of landslide susceptibility modeling using ANN, DNN, and ENN, various metrics and graphical analyses are employed to validate and compare model outputs.

#### Model validation using ROC and precision-recall curves

The receiver operating characteristic (ROC) curve and the precision-recall curve are two widely used tools for evaluating the performance of classification models. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings, providing insight into the model's ability to distinguish between classes. The area under the ROC curve (AUC) serves as a summary measure, with values closer to 1 indicating better model performance. Precision-recall curves, on the other hand, plot precision (the proportion of true positive predictions in the positive predictive class) against recall (the proportion of actual positives correctly identified by the model), which is particularly useful in scenarios with imbalanced datasets. Average precision (AP) summarizes the precision-recall curve as the weighted mean of precisions achieved at each threshold, emphasizing the contribution of high-recall regions. By comparing these curves and summary metrics for ANN, DNN, and ENN models, researchers can gauge the models' discriminative power and their trade-offs between sensitivity (recall) and specificity.

#### Quantitative metrics for model comparison

Beyond graphical methods, a set of quantitative metrics provides a comprehensive view of model performance. Accuracy measures the proportion of true results (both true positives and true negatives) among the total number of cases examined. Precision assesses the model's accuracy in predicting

positive labels, while recall (or sensitivity) evaluates how well the model identifies actual positives. Specificity and negative predictive value (NPV) are metrics focusing on the model's performance in predicting and identifying negative outcomes, respectively. The F1 score offers a balance between precision and recall, helpful in comparing model performance in situations of uneven class distributions. The Matthews correlation coefficient (MCC) and Cohen's Kappa provide overall performance indicators, taking into account true and false positives and negatives, ideal for imbalanced datasets. By computing these metrics for each model, researchers can conduct a thorough comparison, identifying strengths and weaknesses in the context of landslide susceptibility prediction.

### Graphical analysis for in-depth model evaluation

Graphical analyses, such as Bland-Altman plots and cumulative distribution function (CDF) plots, offer deeper insights into model predictions relative to actual observations. Bland-Altman plots highlight the agreement between predicted and true values, illustrating the mean difference (bias) and limits of agreement, which can indicate systematic errors or biases in model predictions. CDF plots, by visualizing the cumulative probability distribution of model predictions, allow researchers to assess the overall distribution and variance of predictions, providing a visual comparison of model performance across the entire range of predicted values. These graphical analyses complement traditional metrics, offering a nuanced view of model behavior, predictive accuracy, and reliability in classifying areas based on their susceptibility to landslides. Together, these validation and comparison methods form a robust framework for evaluating and refining ANN, DNN, and ENN models, ensuring their effectiveness in landslide susceptibility modeling and contributing to informed decision-making in disaster risk management.

### Behavioral assessment using game theory

Game theory offers a mathematical framework for strategic decision-making among multiple agents, and it has found applications in machine learning for analyzing how individual features or parameters influence a model's predictions. SHAP (SHapley Additive exPlanations), grounded in cooperative game theory, provides a systematic method to quantify the contribution of each feature to the prediction of a model, even for complex models like deep neural networks (DNNs), Elman neural networks (ENNs), and generalized linear models (GLMs).

The essence of SHAP values can be understood through the Shapley value concept in cooperative game theory, which assigns a value to each player (feature) that reflects their contribution to the total payout (prediction). Mathematically, the Shapley value for a feature  $i$  in a prediction model can be expressed as follows:

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [v(S \cup \{i\}) - v(S)]$$

where  $N$  is the set of all features,  $S$  is a subset of features excluding  $i$ ,  $v(S)$  is the prediction model's payout (or predicted value) with features in set  $S$ , and  $|S|$  is the cardinality of set  $S$ . This formula calculates the average marginal contribution of feature  $i$  over all possible combinations of features.

In the context of landslide susceptibility modeling, SHAP values can demystify the impact of various parameters like topographic variables (e.g., slope, elevation, aspect), soil properties, land use, and geological characteristics on model predictions. SHAP plots visually represent these contributions, highlighting key factors influencing landslide occurrences. Beyond SHAP, other game theory constructs like Nash equilibrium and game payoff matrices further enrich the analysis of parameter behavior within landslide susceptibility models. For instance, a game payoff matrix can model the interplay among different parameters affecting landslide likelihood, with Nash equilibrium pinpointing parameter configurations most associated with landslide events. Incorporating game theory and SHAP values into landslide

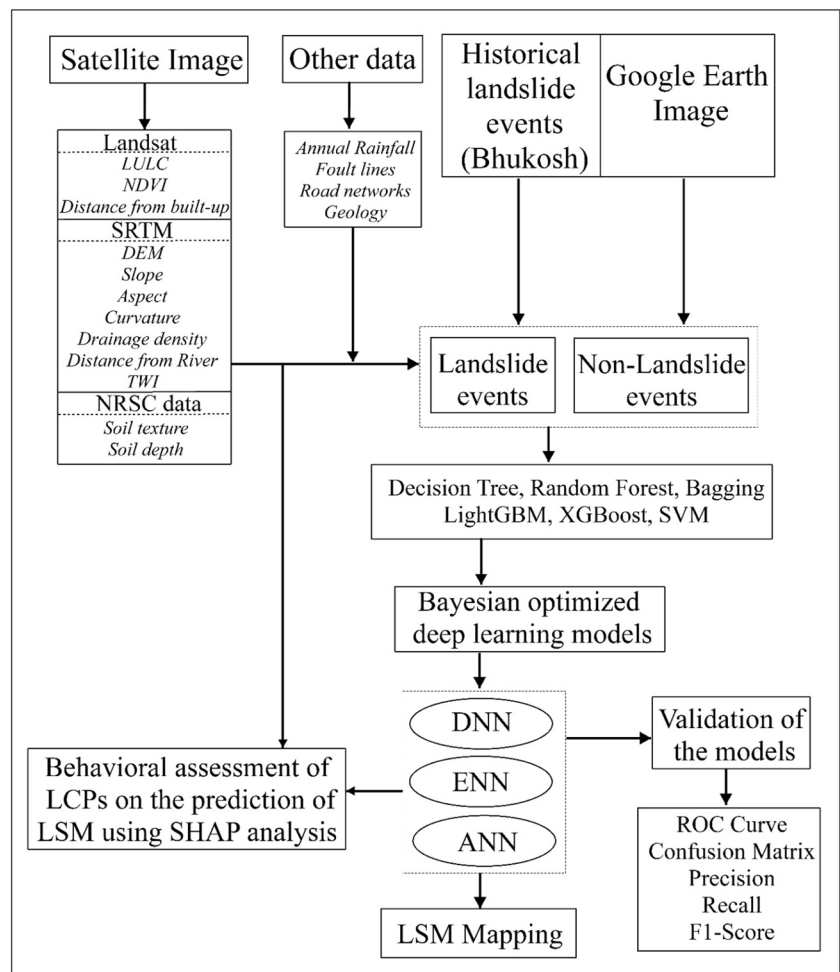
susceptibility models not only aids in understanding the significance of different factors but also enhances model transparency and interpretability. This approach facilitates the development of more accurate, reliable models for landslide hazard assessment and risk management, ultimately contributing to informed decision-making in geohazard mitigation strategies. The methods utilized in the entire work has been presented in a methodological flow chart (Fig. 3).

## Results

### Feature selection analysis

In this study, six feature selection methods, namely decision tree, random forest, bagging, LightGBM, XGBoost, and SVM, were used to identify critical factors for landslide susceptibility prediction. Methods such as decision tree, random forest, bagging, LightGBM, and XGBoost generate direct measures of feature importance based on their algorithmic structures, using metrics such as information gain, Gini importance, or gain scores. In contrast, the SVM

**Fig. 3** Flowchart of the overall methodology



model, particularly with linear kernels, does not inherently provide a direct metric for feature importance, as its mechanism is focused on identifying the optimal hyperplane for classification. In this context, the coefficients associated with each feature in the SVM decision function are considered approximations of the feature’s importance. The magnitude of these coefficients indicates the relative importance of each feature in shaping the decision boundary, with larger absolute values indicating a stronger influence on the model’s predictions. In the study, the absolute values of the coefficients of a linear SVM model, scaled to the range 0–1, were used to estimate the importance of the features. This method facilitated the inclusion of SVM in the comparative analysis of feature selection techniques alongside other models that provide a direct output of importance metrics. The results of each method are presented in Fig. 4, providing a holistic overview of the most influential factors for landslide susceptibility identified by the different feature selection strategies.

The decision tree analysis showed the overriding importance of features such as slope and elevation, while factors such as soil texture and land use/land cover (LULC) were found to be less influential. Similarly, the random forest and bagging methods emphasized the importance of variables such as distance from rivers (DFR) and lineament density (LD), while soil texture and geology were considered less

crucial. The LightGBM and XGBoost analyses were consistent with these observations and indicated that geology and soil properties play a lesser role in landslide susceptibility. The results of the SVM model, which indicated less influence of topographic wetness index (TWI) and curvature, confirmed this trend. These findings led to the exclusion of geology, soil texture, and curvature from further modeling, as their influence on predictive capabilities was consistently minimal. This decision, based on the observation of their limited influence on model predictions, aims to improve the efficiency and accuracy of landslide susceptibility models and ensure their robustness and focus. Such a targeted approach is intended to optimize decision-making and planning in landslide-prone regions by concentrating resources and analytical efforts on the factors that are most important for understanding landslide susceptibility.

### Implementation of Bayesian-optimized deep learning models

#### Optimization of DL models

The optimization of the ANN model in this study was meticulously performed using Bayesian optimization, a powerful strategy for fine-tuning hyperparameters. The search space

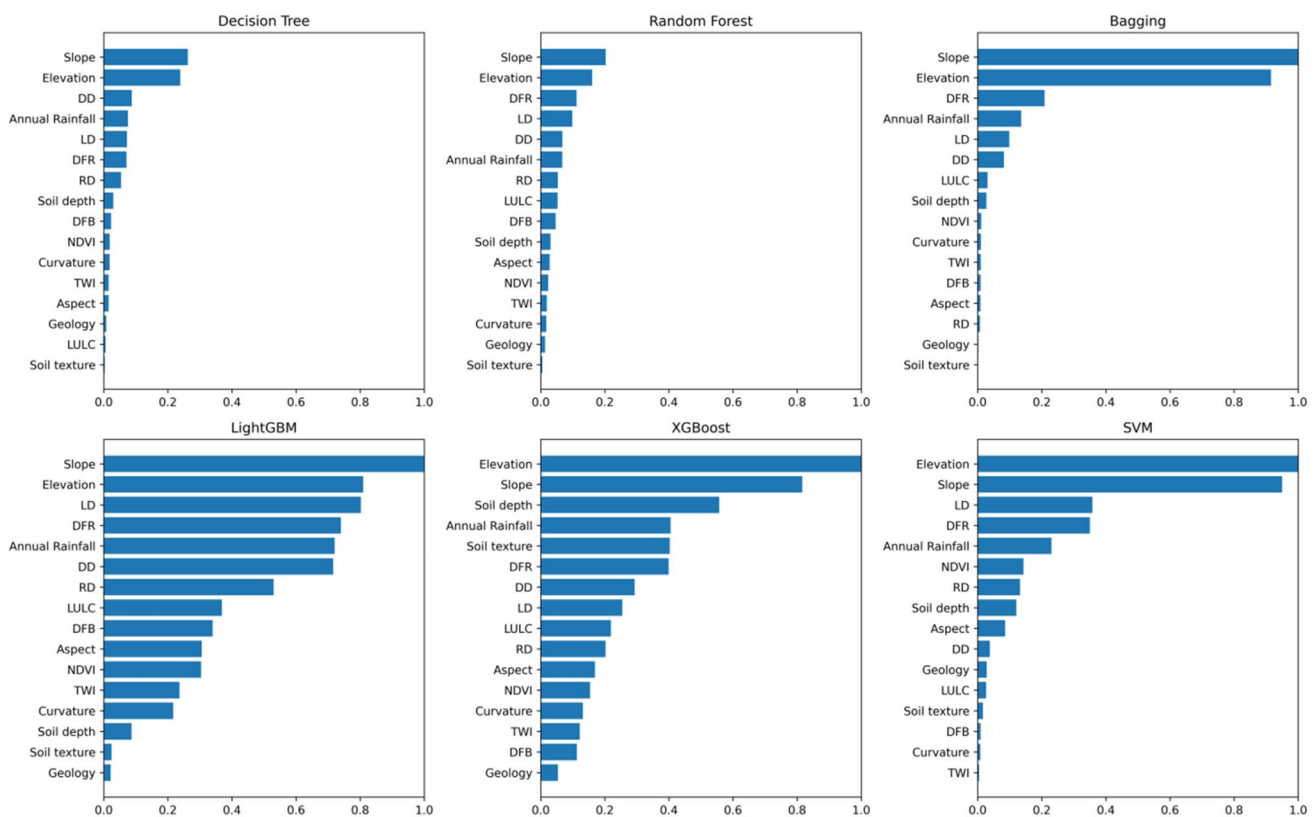


Fig. 4 Feature selection using six machine learning models for appropriate parameter selection

for the hyperparameters was defined to include three layers' unit counts (ranging from 32 to 512 for each layer), learning rate (ranging from 0.0001 to 0.1), dropout rate (ranging from 0.1 to 0.5), the number of layers (from 1 to 5), and batch size (ranging from 32 to 256). These hyperparameters were selected based on their known influence on the training process and overall model performance, considering the depth and complexity of the model, the speed and stability of learning, regularization effects to combat overfitting, and the efficiency of the training process. The hyperparameter

tuning results were plotted on a 2-dimensional surface using the function "plot\_objective(res\_gp)" (Fig. 5). The objective plot, comprising various subplots, offers a detailed view of the hyperparameter search. Each subplot shows the interplay between two hyperparameters and their influence on the model's performance. The different shades of color represent the model's performance, with darker shades typically indicating better performance (higher accuracy or lower loss). The black dots represent combinations of hyperparameters that were evaluated, while the red star marks

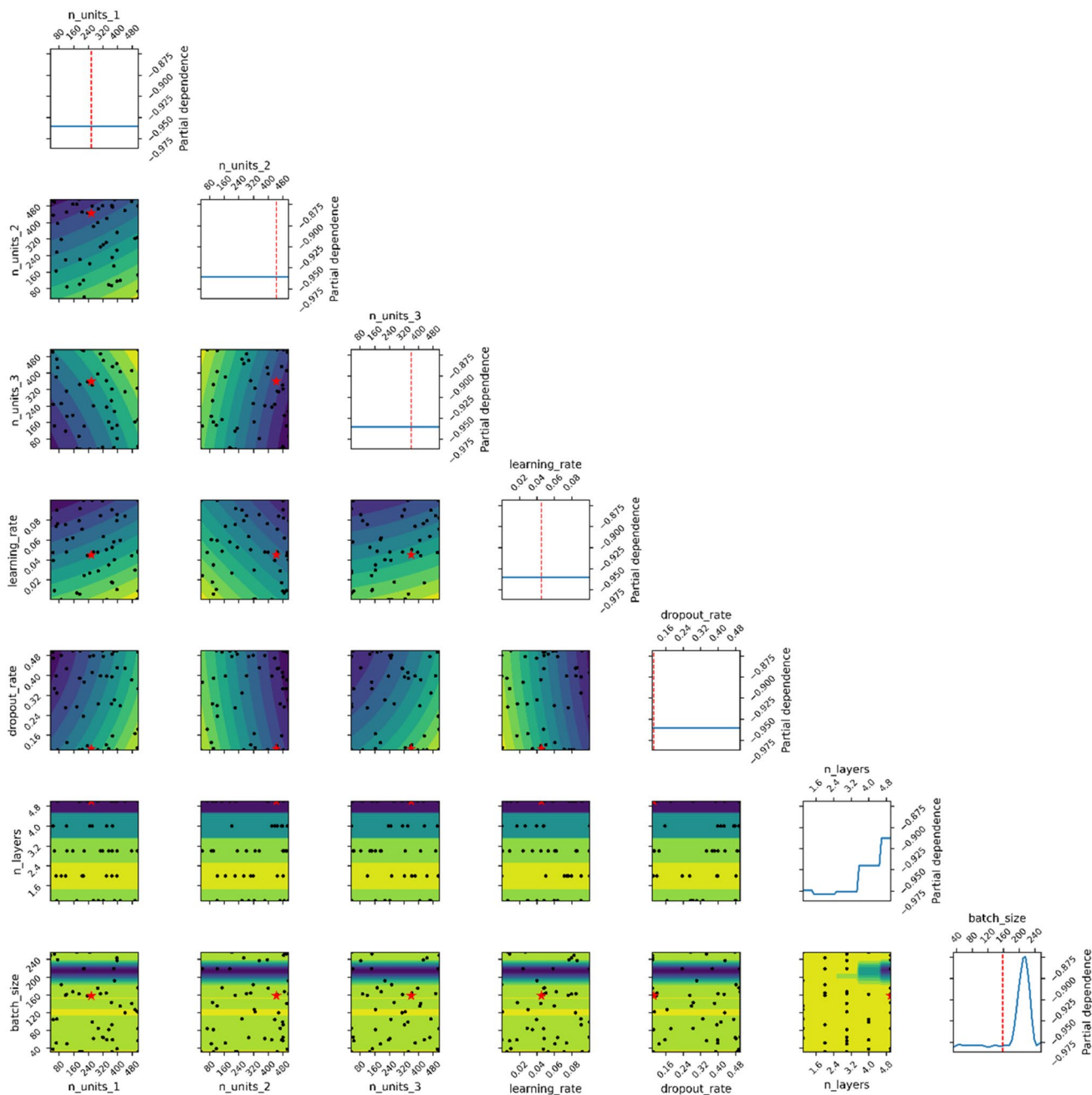


Fig. 5 Objective function plot to show hyperparameter search of ANN model for finding best hyperparameters



the combination of hyperparameters that resulted in the best performance. This graphical representation allows for an intuitive understanding of how hyperparameter values relate to model success and where the optimization algorithm focused its search. The objective plot demonstrated that the optimal hyperparameters were found at [256, 445, 359] for the neuron counts across the layers, with a learning rate of approximately 0.045, a dropout rate just above 0.1, indicating a preference for model complexity moderated by regularization to prevent overfitting. The model complexity was further substantiated by the optimal number of layers set to the maximum of 5 and a mid-range batch size of 158, balancing the computational load with model performance. The convergence plot reinforces the effectiveness of the optimization process, where the minimum negative accuracy score rapidly stabilizes after the initial iterations, signifying a swift approach to the optimal hyperparameter region (Fig. 6a). This stabilization, depicted as a flat line in the convergence plot, suggests that additional iterations beyond

the initial ones provided diminishing returns, quantitatively supporting the adequacy of the optimization procedure and the robustness of the identified hyperparameters in predicting landslide susceptibility.

The search space parameters of DNN model were set to include neuron counts for three layers (32 to 512), learning rate (0.0001 to 0.1), dropout rate (0.1 to 0.5), number of layers (5 to 20), and batch size (32 to 256). The objective plot revealed the intricate landscape of hyperparameter efficacy, with darker shades indicating regions of higher model performance and the red stars pinpointing the optimal hyperparameter values (Fig. 7). These optimal values were determined as [32, 512, 32] for the neuron counts, a learning rate at the upper limit of 0.1, a moderate dropout rate of 0.1, a depth of 5 layers for the network, and a batch size of 124. The convergence plot further validated the effectiveness of the optimization process, depicting a rapid initial decrease in the minimum negative accuracy score, followed by a plateau, which indicates

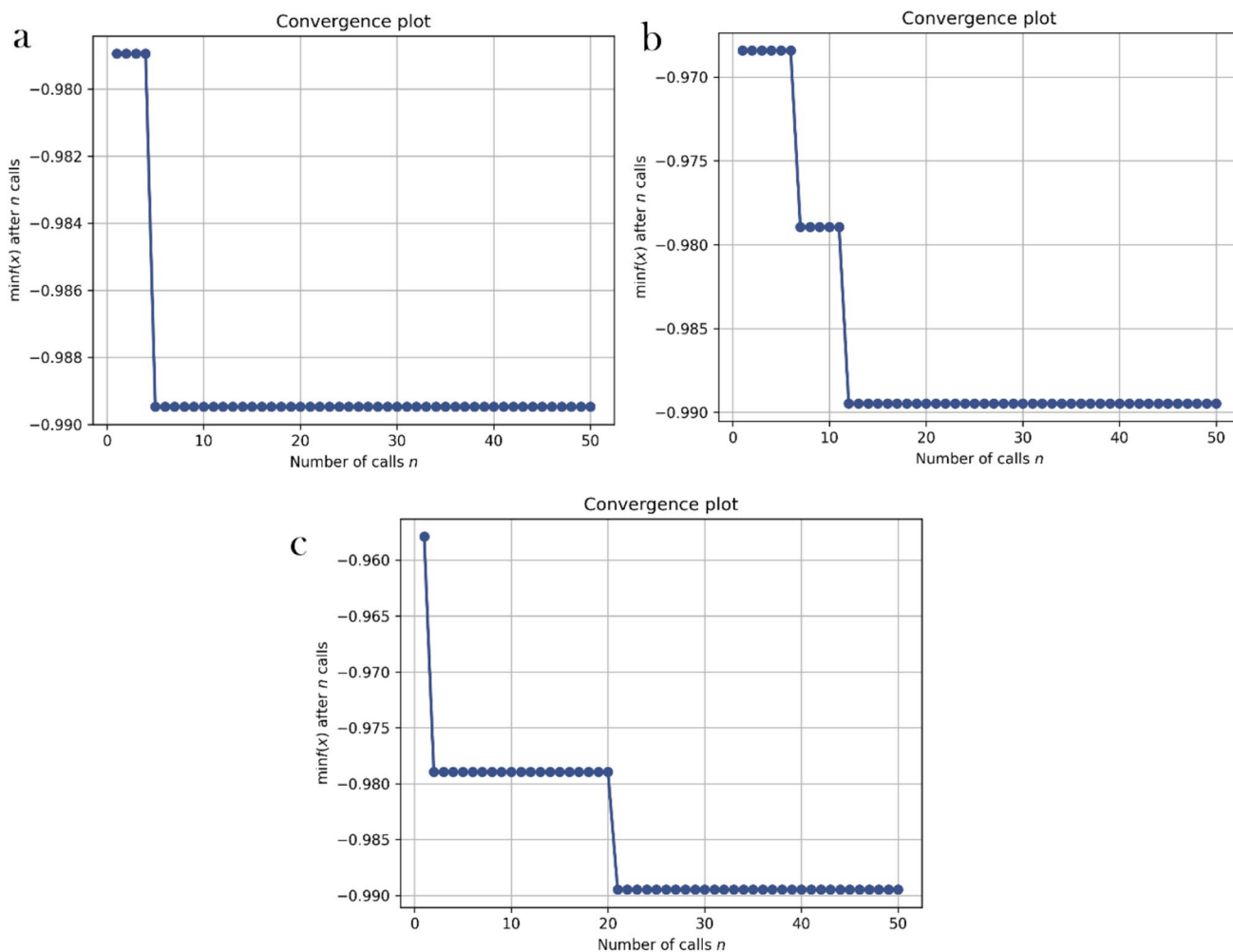
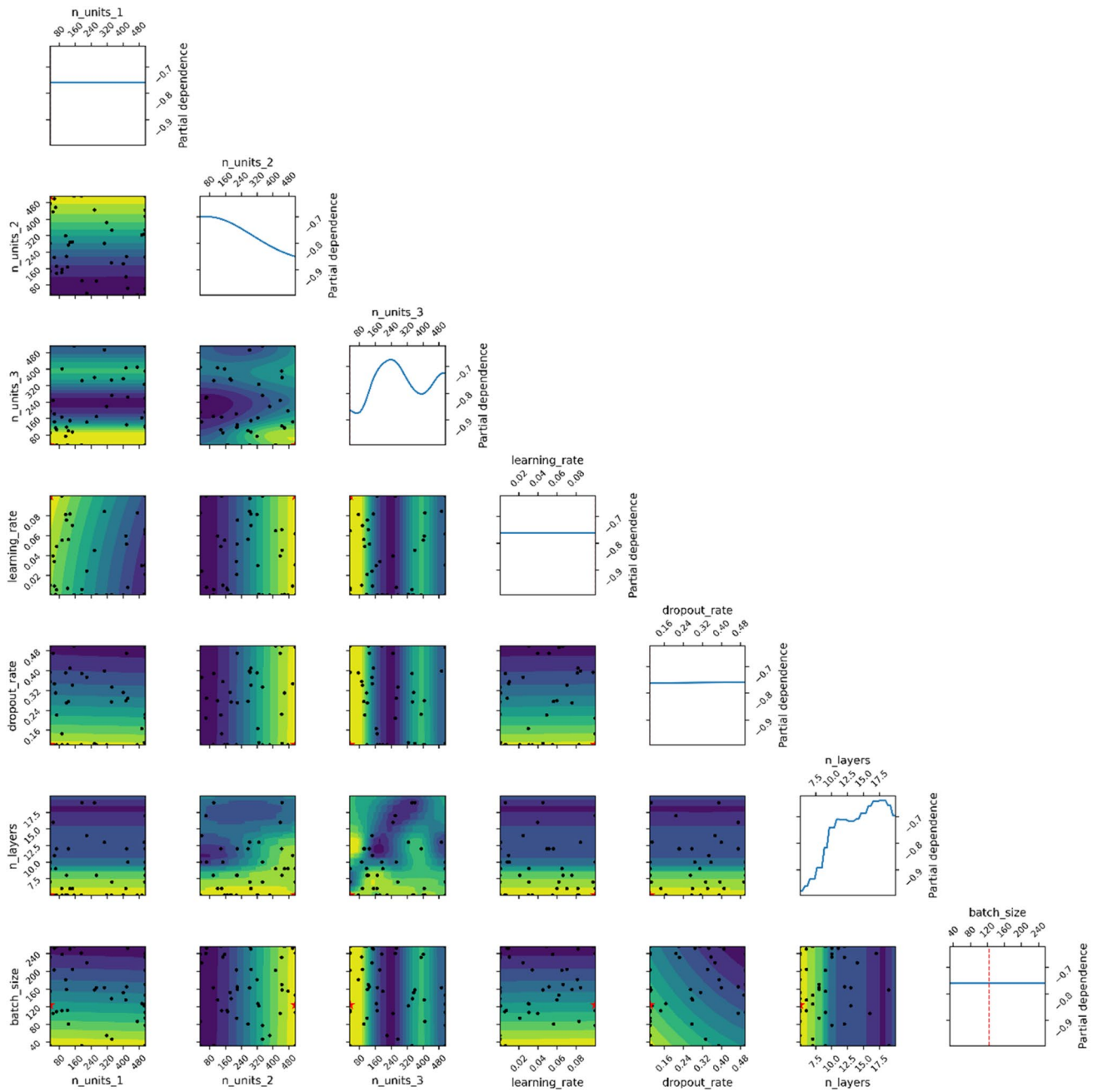


Fig. 6 Convergence plot to show the performance of hyperparameter search for a ANN, b DNN, and c ENN

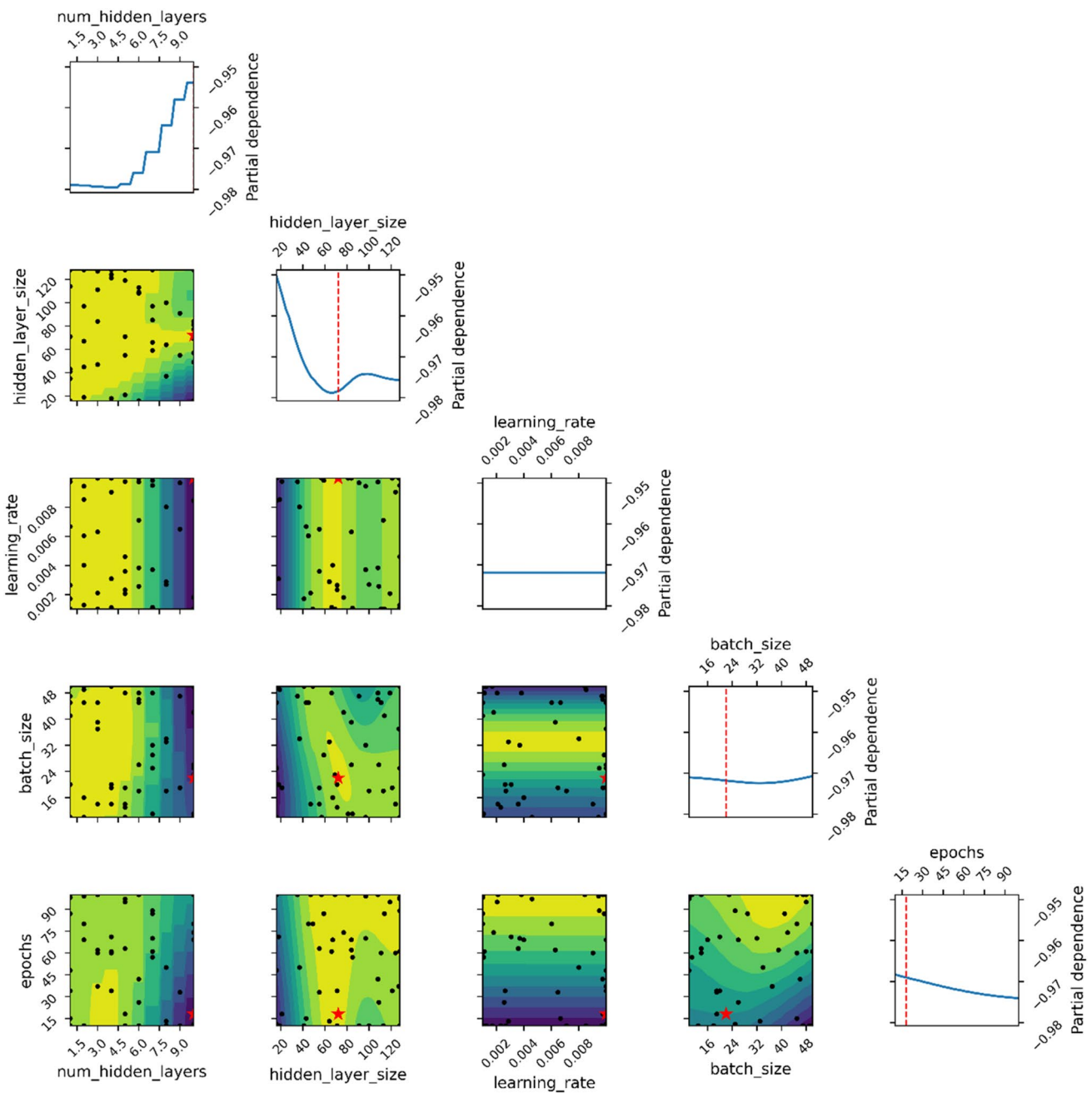


**Fig. 7** Objective function plot to show hyperparameter search of DNN model for finding best hyperparameters

that the optimal hyperparameters were identified early in the 50 iterations (Fig. 6b). This early stabilization suggests that the chosen hyperparameters are robust, providing a DNN model that is well-suited for the complexity of landslide susceptibility prediction.

The optimization of the ENN model was executed through Bayesian optimization within a defined hyperparameter search space, which included the number of hidden layers (1 to 10), hidden layer size (16 to 128),

learning rate (0.001 to 0.01), batch size (10 to 50), and number of epochs (10 to 100). The optimal hyperparameters, determined by this process, were found to be a configuration with 10 hidden layers, a hidden layer size of 72, a learning rate of 0.01, a batch size of 22, and 18 epochs. These parameters were discerned through the interpretation of the objective plot, which presents a visual exploration of hyperparameter performance, where the color gradient indicates varying levels of model



**Fig. 8** Objective function plot to show hyperparameter search of ENN model for finding best hyperparameters

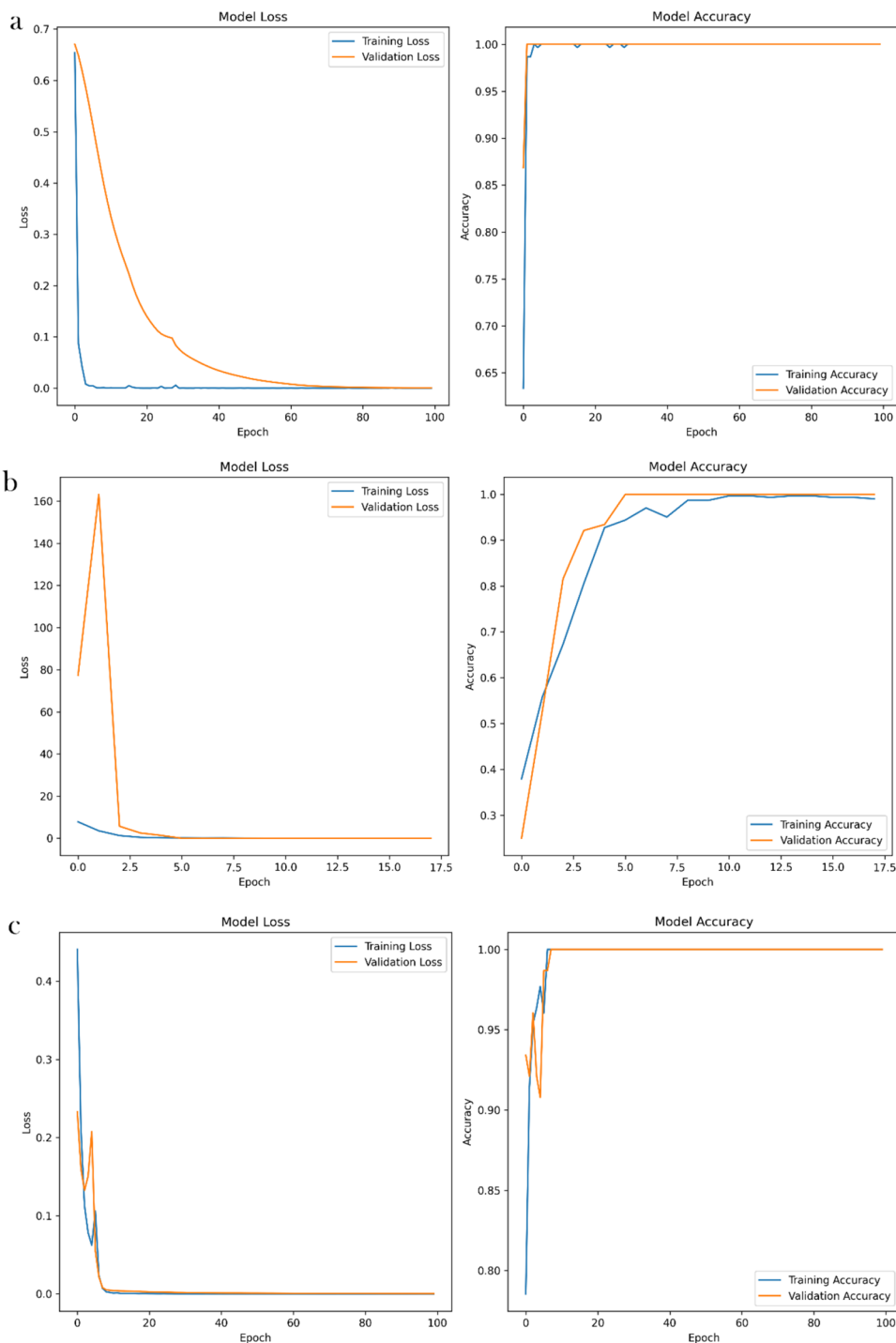
accuracy and the red star marks the peak of model performance (Fig. 8). The convergence plot solidifies the effectiveness of the chosen hyperparameters, exhibiting an initial sharp improvement in model accuracy that plateaus with subsequent iterations, suggesting early attainment of an optimal hyperparameter set (Fig. 6c). This plateau, where further calls do not significantly change the minimum negative accuracy score, quantitatively affirms the efficiency of the optimization process and

the robustness of the resulting ENN model for predicting landslide susceptibility.

**Evaluation of optimized deep learning models through learning curves**

Learning curves are essential for understanding a model’s learning dynamics and diagnosing issues related to its learning process, such as underfitting or overfitting. By examining these

**Fig. 9** Learning curves for the optimized **a** ANN, **b** DNN, and **c** ENN model, displaying the training and validation loss and accuracy over epochs



curves, we can assess how well the model is generalizing from the training data to unseen data and ensure the robustness of the model’s predictive capabilities. The learning curve for the optimized ANN model shows a characteristic rapid decrease in training loss within the initial epochs, indicating that the model is quickly learning from the training dataset (Fig. 9a). The validation loss decreases in tandem with the training loss but starts to plateau, suggesting the model is generalizing well

to unseen data. Both training and validation loss curves converge to a low value, which indicates a minimal gap between the two. This is indicative of a well-fitted model with a balance between bias and variance. The training and validation accuracy curves mirror this behavior, rapidly ascending to high values and converging, with the validation accuracy slightly below but closely following the training accuracy, reaching a stable state after about 20 epochs. For the DNN model, the



learning curve depicts a different pattern. The initial training loss is substantially higher, followed by a steep decline, which then stabilizes (Fig. 9b). The validation loss mirrors the sharp descent but then shows a slight increase, which could be indicative of overfitting as the model starts to learn noise from the training data. However, the training and validation accuracy both show a steep increase, with the training accuracy reaching near-perfect levels and the validation accuracy trailing at a high level but not plateauing as smoothly as the training accuracy, suggesting that the model is optimal, but might benefit from further tuning of regularization parameters or training on more diverse data to improve generalization. Lastly, the ENN model's learning curves reveal an efficient learning process, with the training loss rapidly dropping to a low level and the validation loss closely following, indicating good model generalization (Fig. 9c). The slight divergence between the training and validation loss curves implies a small amount of overfitting. Nonetheless, the accuracy curves for both training and validation quickly ascend to high levels, maintaining a narrow gap between them throughout the training process. This closeness of the training and validation accuracy curves, maintaining high values, suggests the ENN model has learned the patterns within the data well and generalizes effectively to new data, potentially making it a robust model for predicting landslide susceptibility.

### Accuracy assessment of the models

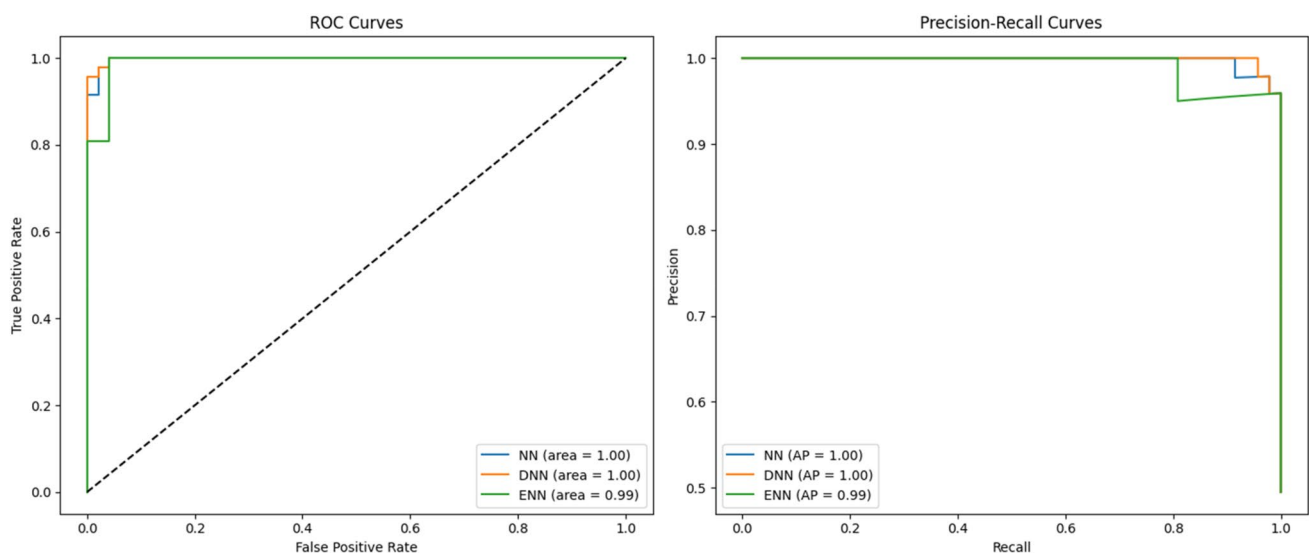
The ROC and precision-recall curves provide a visual and quantitative assessment of model performance for the ANN, DNN, and ENN models (Fig. 10). The ROC curve illustrates the trade-off between the TPR and the FPR, with both ANN

and DNN achieving a perfect AUC of 1.00 and ENN slightly behind at 0.99. In the precision-recall space, which is often more informative in unbalanced datasets, all models perform well, with ANN and DNN achieving a perfect AP score of 1.00 and ENN just below 0.99, indicating a high degree of reliability in predicting positive classes.

In terms of quantitative metrics, ANN and DNN have identical scores across the board, with an accuracy, precision, and F1 score of approximately 0.979 and a perfect recall of 1.000 (Table 2). These results indicate that both models correctly identified every instance of the positive class. The specificity and NPV are 0.958 and 1.000, respectively, with a very low FPR of 0.0417, indicating few false positives. MCC and Cohen's Kappa, which both consider true and false positives and negatives, are also high at around 0.959,

**Table 2** Performance metrics for ANN, DNN, and ENN models

Metric	ANN	DNN	ENN
Accuracy	0.9789	0.9789	0.9684
Precision	0.9592	0.9592	0.9583
Recall	1.0000	1.0000	0.9787
Specificity	0.9583	0.9583	0.9583
F1 score	0.9792	0.9792	0.9684
NPV	1.0000	1.0000	0.9787
FPR	0.0417	0.0417	0.0417
FNR	0.0000	0.0000	0.0213
FDR	0.0408	0.0408	0.0417
Balanced accuracy	0.9792	0.9792	0.9685
MCC	0.9588	0.9588	0.9371
Cohen's Kappa	0.9579	0.9579	0.9368



**Fig. 10** Comparative ROC and precision-recall curves for ANN, DNN, and ENN models. The curves demonstrate the performance of each model in terms of true positive rate vs. false positive rate (left) and precision vs. recall (right)

emphasizing the strong performance of the models. The ENN model lags slightly behind with a precision of 0.9684 and an F1 score of the same value, but still shows high precision and specificity, albeit with a slightly lower recall of 0.9787, suggesting that some positive instances were missed. Overall, the ANN and DNN models exhibit superior performance metrics, making them the best candidates for landslide susceptibility prediction in this study, with the DNN perhaps offering a slight advantage due to its more complex architecture that may capture more nuanced patterns in the data.

### Comparative analysis

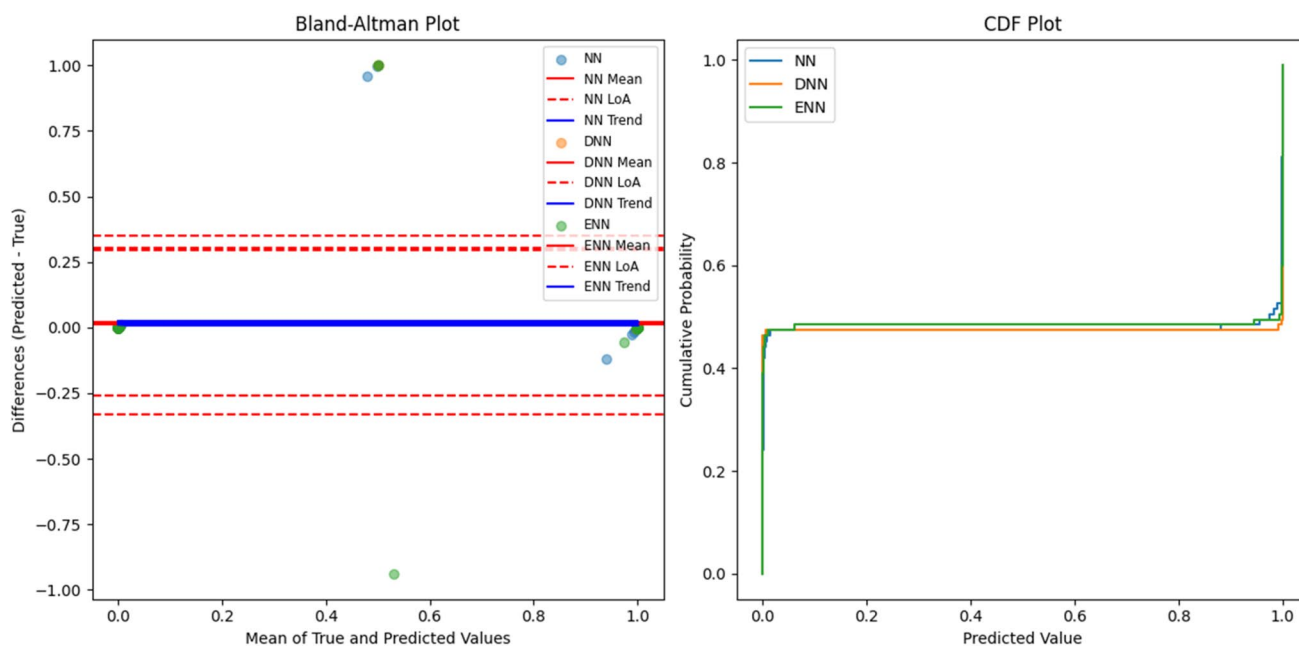
The Bland-Altman plot and CDF diagram provide insights into the comparative analysis of the predictive performance of the ANN, DNN, and ENN models. Together, these plots quantify the predictive accuracy and reliability of the models in the assessment of landslide susceptibility. The Bland-Altman plot, which assesses the agreement between the predicted and true values, shows that the mean differences for all three models are close to zero—ANN at 0.019157, DNN at 0.021028, and ENN at 0.010684, indicating that the predictions are, on average, closely aligned with the true values. The standard deviations of these differences are slightly larger for the ENN model (0.173863) than for the ANN (0.141141) and DNN (0.143528), suggesting greater variability in the predictions of the ENN model. The limits of agreement, calculated as the mean difference plus or minus 1.96 times the standard

deviation, are closest for the ANN model and widest for the ENN model, further indicating that the ANN and DNN models have closer agreement with the true values.

The trend lines in the Bland-Altman plots show negligible slopes for all models, confirming that there is no obvious proportional bias over the entire range of predictions. In the CDF plot, which illustrates the cumulative probability of the predicted values, all models quickly approach a cumulative probability of 1, indicating high confidence in their predictions (Fig. 11). However, the Bland-Altman results provide a more accurate assessment, with the ANN and DNN models showing slightly better agreement with the true values than the ENN model (Fig. 11). Overall, the ANN and DNN models appear to perform best, with the DNN model having the smallest lead due to the slightly lower mean difference and tighter limits of agreement, suggesting that its predictions are in better agreement with the true values.

### LSM modeling using optimized models

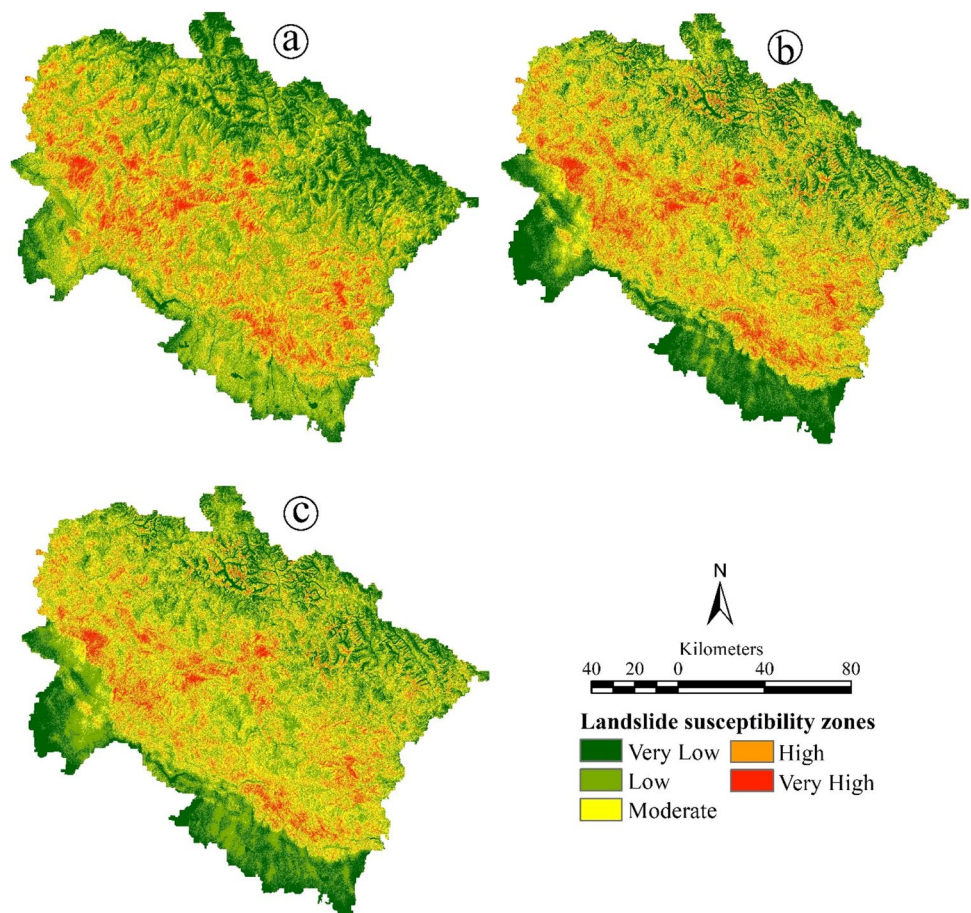
In this study, the researchers have used Bayesian optimization to train three deep learning models (ENN, ANN, and DNN) to generate a LSM that ranges from 0 to 1. The LSM values were then classified into five categories, i.e., very high, high, moderate, low, and very low, using a natural break algorithm. The area (km<sup>2</sup>) under the five categories for all three models was computed, and the results were presented in Fig. 12.



**Fig. 11** Comparative analysis of model predictions. The Bland-Altman plot (left) illustrates the agreement between the true values and the predictions made by ANN, DNN, and ENN models, with mean differences and limits of agreement indicating the precision of each

model. The CDF plot (right) displays the cumulative distribution of the predicted probabilities, reflecting the confidence in the models' predictive performances

**Fig. 12** Landslide susceptibility mapping using Bayesian-optimized **a** ENN, **b** DNN, and **c** ANN models



The results show that the DNN model outperformed the other two models in terms of area covered under each category. For example, the DNN model covered 7839.5497 km<sup>2</sup> under the very low landslide susceptibility zone, whereas the ENN and ANN models covered 8524.828 and 8802.5497 km<sup>2</sup>, respectively. Similarly, the DNN model covered 3613.4442 km<sup>2</sup> under the very high landslide susceptibility zone, whereas the ENN and ANN models covered 4018.0906 and 4613.4442 km<sup>2</sup>, respectively.

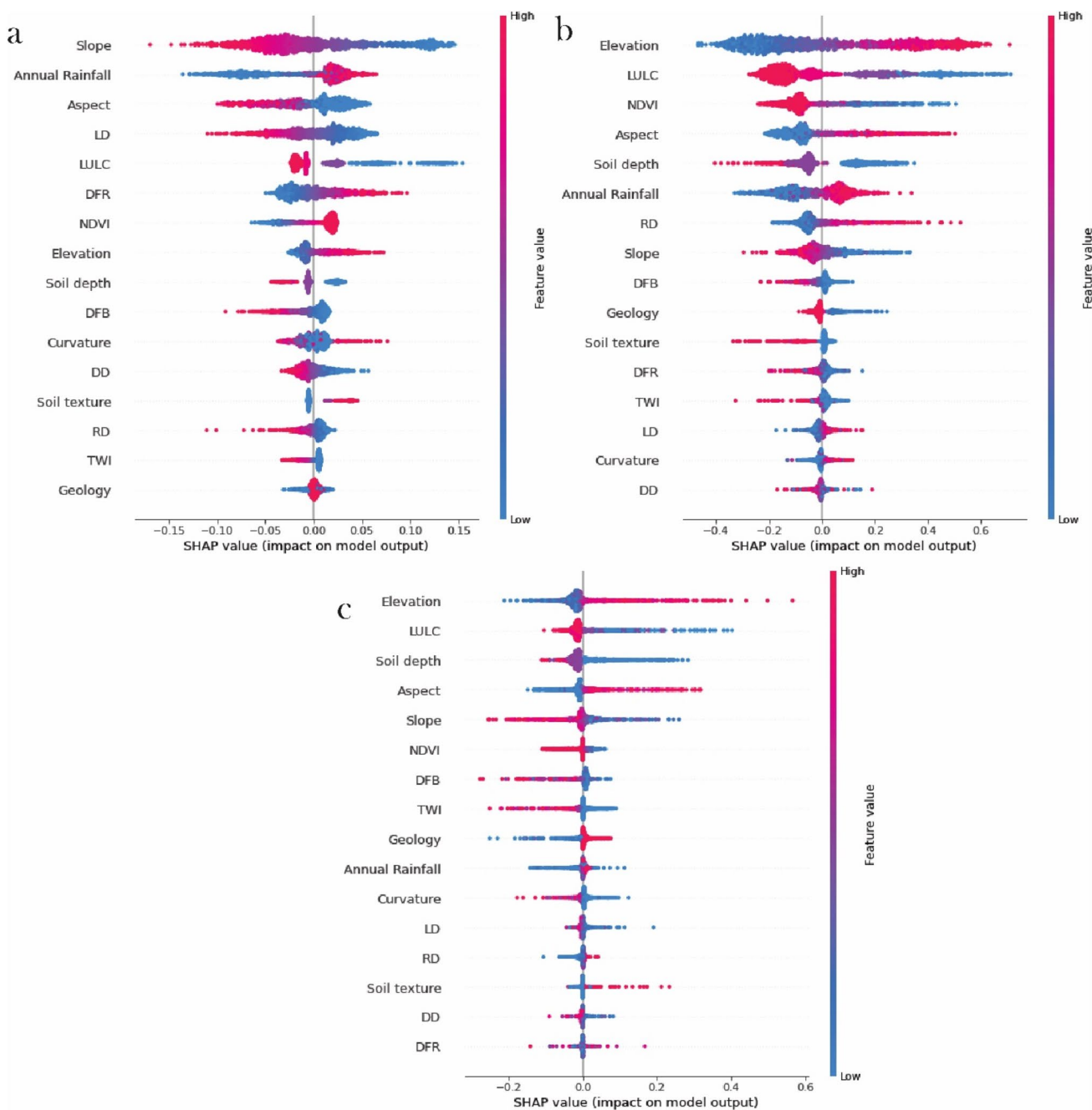
These findings suggest that the DNN model is better suited for generating landslide susceptibility maps, as it can classify areas with higher accuracy compared to the other two models. This information can be used by decision-makers and planners to prioritize resources for areas that are more susceptible to landslides. For example, areas under the very high and high susceptibility zones may require more attention and resources to reduce the risk of landslides.

### Behavioral assessment using game theory and SHAP analysis

The application of SHAP values for interpretability in machine learning offers a compelling insight into the contribution of each feature toward the model's predictions. This study has

employed the SHAP method to elucidate the factors influencing landslide susceptibility in Uttarakhand, India, incorporating the valuable precedents set by Pradhan et al. (2023) and Zhou et al. (2022) in leveraging SHAP for explainable AI in landslide modeling. By integrating SHAP with our Bayesian-optimized DNN, ENN, and ANN models, we have derived a nuanced understanding of the parameter relationships contributing to landslide occurrences, depicted in Fig. 13.

The SHAP summary plots reveal the relative impact of each variable on the model's prediction. The plots exhibit a range of color intensities, with red indicating a higher impact and blue a lower impact on the prediction outcome. The SHAP summary for the ENN model indicates that slope and annual rainfall hold the most substantial influence on predictions, consistent with established landslide literature (Pradhan et al., 2023) (Fig. 13a). The high SHAP values for slope and annual rainfall underscore their critical role in landslide occurrences. Factors like aspect and land use, specifically built-up and agricultural areas, also show significant impact, depicted by their red shading. Interestingly, the impact of soil depth and distance from the river (DFR) is less pronounced compared to the other factors, as indicated by the narrower spread of SHAP values. The distribution of SHAP values for soil texture and the topographic wetness



**Fig. 13** SHAP summary plot for behavioral assessment of LCPs on the prediction of LSM using **a** ENN, **b** DNN, and **c** ANN models

index (TWI) shows a mix of positive and negative impacts, suggesting a complex interaction with landslide susceptibility. The DNN model’s SHAP plot reveals a slightly different hierarchy of feature importance (Fig. 13b). Elevation and land use/land cover (LULC) exhibit the most substantial positive impact on model predictions, aligning with the notion that higher elevations and certain land uses are more susceptible to landslides. NDVI and aspect follow closely, suggesting that vegetation density and slope orientation are

also key contributors to landslide risk. Soil depth, while important, shows a wide distribution of impact across samples, indicating variability in how this factor affects landslide susceptibility across different regions. The features like geology and distance from the river (DFR) present a lower impact, as evidenced by their less intense SHAP values. For the ANN model, elevation stands out as the most influential factor, with a high density of positive SHAP values, suggesting that areas of higher elevation are more prone to



landslides (Fig. 13c). Land use (LULC) follows closely, particularly highlighting the importance of human activities and land management in landslide susceptibility. Soil depth and aspect are also notable factors but with less intensity compared to elevation and LULC. Interestingly, the plot indicates that while slope is a significant factor, its impact is less dominant in the ANN model compared to the ENN model. Additionally, the SHAP values for NDVI and distance from the river bank (DFB) are relatively lower, suggesting these factors have a moderate influence on the model's output.

The cumulative impact of these factors, as detailed in the SHAP analysis, provides a scientific foundation for practical landslide disaster prevention and mitigation, offering a quantifiable approach to prioritize interventions. By applying the SHAP method to deep learning models, we gain not just a predictive model for landslide susceptibility but also a valuable interpretive tool that imparts more profound scientific guidance. The SHAP analysis bridges the gap between predictive accuracy and practical applicability, guiding targeted interventions for disaster prevention and enhancing the models' reliability. The discussion is thus expanded to emphasize the superior performance of the DNN model, as suggested by the model evaluation metrics and SHAP analysis, asserting its suitability for generating accurate and interpretable landslide susceptibility maps.

## Discussion

Landslides are one of the most catastrophic natural disasters that have caused severe damage to property and human life worldwide. In India, landslides are a common phenomenon, and the state of Uttarakhand is highly prone to landslides due to its geographical location and topography. Uttarakhand has been hit by several landslide disasters in the past, which have caused enormous loss of life and property. Therefore, predicting landslide susceptibility zones in Uttarakhand is of paramount importance to minimize the damage caused by landslides. Here, three important issues have been discussed, such as landslide susceptibility assessment, understanding of the novelty, and policy implication.

### Landslide susceptibility assessment

In this study, six different feature selection techniques were employed, namely decision tree, random forest, bagging, LightGBM, XGBoost, and SVM, to identify the most influential variables for predicting landslide susceptibility zones in Uttarakhand. Based on the results of feature selection, three variables—geology, soil texture, and curvature were excluded, and eight variables were selected, including slope, elevation, DFR, annual rainfall, LD, DD, LULC, and soil depth.

Next, three different models, namely DNN, Elman neural network, and ANN, were implemented using a Bayesian optimization technique to predict the landslide susceptibility zones in Uttarakhand. To understand the behavior of the model and how each variable contributes to the prediction of landslide susceptibility, the SHAP (SHapley Additive exPlanations) method was used. The DNN model projected that 7839.5497 km<sup>2</sup> would fall under the “very low” landslide susceptibility zone, while the ENN and ANN models covered 8524.828 and 8802.5497 km<sup>2</sup>, respectively. On the other hand, the DNN model predicted that 3613.4442 km<sup>2</sup> would be classified under the “very high” susceptibility zone, while the ENN and ANN models covered 4018.0906 and 4613.4442 km<sup>2</sup>, respectively. The study found that the areas predicted to have high landslide susceptibility are primarily located in the southern foothills and Ganga river valley in Uttarakhand where there is rapid urban expansion and infrastructure development (Gupta et al., 2022). The increase in urban growth and new construction in these areas has a significant impact on triggering landslides. The “very high” and “high” susceptibility zones are also located in urban clusters such as Dehradun, Haridwar, Mussoorie, Tehri, Haldwani, Pauri, Nainital, Pithoragarh, and Almora. These observations on susceptibility to landslides are consistent with previous studies conducted by Ram et al. (2020) in the Mussoorie region and Pham et al. (2017) in the Tehri urban cluster region. Thus, the results of this study on landslide susceptibility in Uttarakhand are in line with previous findings and provide valuable insights for land use planning and disaster management in the region.

Understanding the factors that contribute to landslide susceptibility is critical for developing effective landslide risk management strategies. The susceptibility of landslides is caused by the interaction of several land cover parameters (LCPs), such as topographic and human influences (Gupta et al., 2022). A number of studies have addressed the complex problem of quantifying the relationship between landslide susceptibility and the parameters that influence it (Pham et al., 2017; Ram et al., 2020; Tran et al., 2021). To understand the behavior of the model and how each variable contributes to the prediction of landslide susceptibility, the SHAP (SHapley Additive exPlanations) method was used in this study. The results showed that high elevation, built-up and agricultural land use, less vegetation, aspect (north and northwest), soil depth less than 140 cm, and annual rainfall above 1500 mm significantly influence and initiate landslides in Uttarakhand. These findings are consistent with previous studies that have identified slope and rainfall as the most important parameters for landslide triggering (Pham et al., 2017; Tanyu et al., 2021; Gupta et al., 2022).

## Understanding the novelty

Landslide susceptibility mapping is an important task for reducing the risk of landslides and ensuring the safety of people and infrastructure in landslide-prone areas. Many previous studies have focused on the use of DNN models for this task, and various optimization techniques such as grid search, particle swarm optimization, and genetic algorithms have been used to improve the performance of these models.

However, there has been little research on the use of ENN and ANN for landslide susceptibility mapping, and even fewer studies have focused on optimizing these models (Zhang et al., 2023; Liu et al., 2023; Wang et al., 2024; Yu & Chen, 2024). This is where the current study stands out, as it employs Bayesian model optimization to optimize DNN, ENN, and ANN models for landslide susceptibility mapping, and derives SHAP values from these models. The use of Bayesian model optimization is a novel approach in this field, and the results obtained are very promising. Bayesian optimization is a powerful tool for hyperparameter tuning, and its application to the optimization of DNN, ENN, and ANN models for landslide susceptibility mapping could have a significant impact on future research in this field. This study has a unique aspect that has not been explored before: it derives SHAP values from three optimized models (DNN, ENN, and ANN), which is different from the methods used by other researchers who typically use XGBoost and random forest. Moreover, the derivation of SHAP values from these optimized models is another novel aspect of this study. SHAP values are an important tool for interpreting the results of machine learning models and can help to identify the factors that contribute most to landslide susceptibility. The fact that the SHAP values were computed using optimized DNN, ENN, and ANN models further adds to the reliability of the results obtained.

Therefore, the novel approaches taken in this study, namely the use of Bayesian model optimization and the derivation of SHAP values from optimized DNN, ENN, and ANN models, make a significant contribution to the field of landslide susceptibility mapping. This is important because the results of landslide susceptibility mapping can have significant implications for management and safety purposes. The use of these optimized models and the derivation of SHAP values from them can help stakeholders make more informed decisions and take appropriate actions to reduce the risk of landslides and ensure the safety of people and infrastructure in landslide-prone areas.

## Policy implication

The results of this study have important implications for landslide risk management in Uttarakhand. The identified LCPs should be considered in land use planning and

development activities to reduce the risk of landslides. For example, strategies to reduce the impact of urbanization and agricultural land use on the susceptibility of landslides could include enforcing zoning regulations, adopting appropriate building codes and designing retaining structures, and establishing effective drainage systems. In addition, there is a need to increase the resilience of the society against landslides. This can be achieved by raising public awareness on the dangers of landslides and educating communities on how to prepare for and respond to landslide events. Effective communication strategies, early warning systems, and emergency response plans should be developed and implemented to minimize the impact of landslides on society. Furthermore, research efforts should focus on developing more accurate and efficient landslide susceptibility models that incorporate a wide range of LCPs and other relevant factors such as land use, soil properties, and vegetation cover. These models can help in predicting and mitigating the impact of landslides and inform policymakers and stakeholders on the most effective landslide risk management strategies to adopt in the region. Therefore, the results of this study demonstrate the importance of considering a range of LCPs in landslide susceptibility analysis and provide valuable insights for landslide risk management and mitigation strategies in Uttarakhand. A comprehensive and collaborative approach that integrates land use planning, community preparedness, and effective communication strategies is essential for developing a landslide-resilient society in the region.

## Conclusion

This study employs Bayesian model optimization to optimize three deep learning models (ENN, ANN, and DNN) for landslide susceptibility mapping and derive SHAP values from these optimized models. The study uses six different feature selection techniques to identify the most important variables for predicting landslide susceptibility. The results indicate that slope, elevation, and drainage density are the most influential variables, while geology and soil texture have less influence. The DNN model outperformed the other two models in terms of the area covered under each landslide susceptibility zone, suggesting that it is better suited for generating landslide susceptibility maps. However, this study also highlights the need for further research on the use of Bayesian optimization and the derivation of SHAP values from optimized models, as well as the need to apply these techniques in different regions with varying geological and environmental conditions. Moreover, the study highlights the importance of evaluating the performance of models using various performance metrics to ensure their accuracy and reliability.

The study highlights the need for further research on the application of these techniques in different regions with varying geological and environmental conditions. Additionally, there is a need to investigate how the use of these optimized models and SHAP values can be translated into practical decision-making for landslide risk management. The current issue with SHAP values, which require separate runs with XGBoost models, can be overcome by integrating the computation with the landslide susceptibility models. This study's innovative approaches provide a foundation for future research and management decisions in the field of landslide susceptibility mapping. Further research can explore the use of different deep learning models, feature selection techniques, and optimization methods to improve the accuracy and reliability of landslide susceptibility mapping. Ultimately, this can contribute to reducing the risk of landslides and ensuring the safety of people and infrastructure in landslide-prone areas.

**Acknowledgements** The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Research Group under grant no. R.G.P2/68/44. The authors are also thankful to the USGS Earth Explorer for making the LANDSAT data freely available.

**Author contribution** Conceptualization: Javed Mallick and Hoang Thi Hang; data curation: Javed Mallick and Hoang Thi Hang; formal analysis: Javed Mallick; funding acquisition: Javed Mallick; methodology: Javed Mallick and Meshel Alkahtani; project administration: Meshel Alkahtani and Chander Kumar Singh; resources: Meshel Alkahtani; software: Javed Mallick and Hoang Thi Hang; supervision: Javed Mallick; validation: Javed Mallick; writing—original draft: Javed Mallick; writing—review and editing: Chander Kumar Singh

**Funding** Funding for this research was given under award no. R.G.P2/68/44 by the Deanship of Scientific Research; King Khalid University, Ministry of Education, Kingdom of Saudi Arabia.

**Data availability** The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Ethics approval and consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

## References

- Achour Y, Pourghasemi HR (2020) How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? *Geosci Front* 11(3):871–883
- Alqadhi S, Mallick J, Talukdar S, Ahmed M, Khan RA, Sarkar SK, Rahman A (2022c) Assessing the effect of future landslide on ecosystem services in Aqabat Al-Sulbat region, Saudi Arabia. *Nat Hazards* 113(1):641–671
- Alqadhi S, Mallick J, Talukdar S, Bindajam AA, Saha TK, Ahmed M, Khan RA (2022a) Combining logistic regression-based hybrid optimized machine learning algorithms with sensitivity analysis to achieve robust landslide susceptibility mapping. *Geocarto Int* 37(25):9518–9543
- Alqadhi S, Mallick J, Talukdar S, Bindajam AA, Van Hong N, Saha TK (2022b) Selecting optimal conditioning parameters for landslide susceptibility: an experimental research on Aqabat Al-Sulbat, Saudi Arabia. *Environ Sci Pollut Res* 29(3):3743–3762
- Althuwaynee OF, Pradhan B, Park HJ, Lee JH (2014) A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping. *Catena* 114:21–36
- Anamika K, Mehra R, Malik P (2020) Assessment of radiological impacts of natural radionuclides and radon exhalation rate measured in the soil samples of Himalayan foothills of Uttarakhand, India. *J Radioanal Nucl Chem* 323(1):263–274
- Arabameri A, Chandra Pal S, Rezaie F, Chakraborty R, Saha A, Blaschke T et al (2022) Decision tree based ensemble machine learning approaches for landslide susceptibility mapping. *Geocarto Int* 37(16):4594–4627
- Arulbalaji P, Padmalal D, Sreelash K (2019) GIS and AHP techniques based delineation of groundwater potential zones: a case study from southern Western Ghats, India. *Sci Rep* 9(1):1–17
- Awasthi S, Varade D, Bhattacharjee S, Singh H, Shahab S, Jain K (2022) Assessment of land deformation and the associated causes along a rapidly developing Himalayan Foothill Region using multi-temporal Sentinel-1 SAR datasets. *Land* 11(11):2009
- Azarafza M, Azarafza M, Akgün H, Atkinson PM, Derakhshani R (2021) Deep learning-based landslide susceptibility mapping. *Sci Rep* 11(1):1–16
- Billah M, González PA, Delgado RC (2019) Patterns of mortality caused by natural disasters and human development level: a South Asian analysis. *Indian J Public Health Res Dev* 10(2):312–316
- Bui DT, Tsangaratos P, Nguyen VT, Van Liem N, Trinh PT (2020) Comparing the prediction performance of a deep learning neural network model with conventional machine learning models in landslide susceptibility assessment. *Catena* 188:104426
- Canavesi V, Segoni S, Rosi A, Ting X, Nery T, Catani F, Casagli N (2020) Different approaches to use morphometric attributes in landslide susceptibility mapping based on meso-scale spatial units: a case study in Rio de Janeiro (Brazil). *Remote Sens* 12(11):1826
- Chen L, Mei L, Zeng B, Yin K, Shrestha DP, Du J (2020) Failure probability assessment of landslides triggered by earthquakes and rainfall: a case study in Yadong County, Tibet, China. *Sci Rep* 10(1):1–12
- Dey J, Sakhre S, Vijay R, Bherwani H, Kumar R (2021) Geospatial assessment of urban sprawl and landslide susceptibility around the Nainital Lake, Uttarakhand, India. *Environ Dev Sustain* 23(3):3543–3561
- Dobriyal MJR, Bijalwan A (2017) Forest fire in western Himalayas of India: a review. *New York Sci J* 10
- Gao H, Liu F, Yan T, Qin L, Li Z (2022) Drainage density and its controlling factors on the eastern margin of the Qinghai–Tibet Plateau. *Front Earth Sci* 9:1280
- Gautam N (2022) Landslide susceptibility mapping of Kinnaur District in Himachal Pradesh, India using probabilistic frequency ratio model. *J Geol Soc India* 98(11):1595–1604
- Gupta V, Nautiyal H, Kumar V, Imlirenla J, Tondon RS (2016) Landslide hazards around Uttarkashi township, Garhwal Himalaya, after the tragic flash flood in June 2013. *Nat Hazards* 80:1689–1707. <https://doi.org/10.1007/s11069-015-2048-4>

- Gupta V, Kumar S, Kaur R, Tandon RS (2022) Regional-scale landslide susceptibility assessment for the hilly state of Uttarakhand, NW Himalaya, India. *J Earth Syst Sci* 131(1):1–18
- Haque U, Blum P, Da Silva PF, Andersen P, Pilz J, Chalov SR et al (2016) Fatal landslides in Europe. *Landslides* 13(6):1545–1554
- Kaur H, Gupta S, Parkash S, Thapa R (2018) Knowledge-driven method: a tool for landslide susceptibility zonation (LSZ). *Geology, Ecology, and Landscapes* 7(1):1–15
- Kavzoglu T, Colkesen I, Sahin EK (2019) Machine learning techniques in landslide susceptibility mapping: a survey and a case study. *Landslides: Theory, practice and modelling* 283–301
- Kumar A, Asthana AKL, Priyanka RS, Jayagondaperumal R, Gupta AK, Bhakuni SS (2017) Assessment of landslide hazards induced by extreme rainfall event in Jammu and Kashmir Himalaya, north-west India. *Geomorphology* 284:72–87
- Liao M, Wen H, Yang L (2022) Identifying the essential conditioning factors of landslide susceptibility models under different grid resolutions using hybrid machine learning: a case of Wushan and Wuxi counties, China. *Catena* 217:106428. <https://doi.org/10.1016/j.catena.2022.106428>
- Liu W, Sun J, Liu G, Fu S, Liu M, Zhu Y, Gao Q (2023) Improved GWO and its application in parameter optimization of Elman neural network. *PLoS One* 18(7):e0288071
- Liu Z, Qiu H, Ma S, Yang D, Pei Y, Du C et al (2021) Surface displacement and topographic change analysis of the Changhe landslide on September 14, 2019. *China Landslides* 18(4):1471–1483
- Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. *Advances in neural information processing systems* 30
- Mahato S, Pal S, Talukdar S, Saha TK, Mandal P (2021) Field based index of flood vulnerability (IFV): a new validation technique for flood susceptible models. *Geosci Front* 12(5):101175
- Mallick J, Alqadhi S, Talukdar S, AlSubih M, Ahmed M, Khan RA et al (2021) Risk assessment of resources exposed to rainfall induced landslide with the development of GIS and RS based ensemble metaheuristic machine learning algorithms. *Sustainability* 13(2):457
- Mallick J, Alqadhi S, Talukdar S, Sarkar SK, Roy SK, Ahmed M (2022) Modelling and mapping of landslide susceptibility regulating potential ecosystem service loss: an experimental research in Saudi Arabia. *Geocarto Int* 37(25):10170–10198
- Martha TR, Roy P, Jain N, Khanna K, Mrinalni K, Kumar KV, Rao PVN (2021) Geospatial landslide inventory of India—an insight into occurrence and exposure on a national scale. *Landslides* 18(6):2125–2141
- Nwazelibie VE, Unigwe CO, Egbueri JC (2022) Integration and comparison of algorithmic weight of evidence and logistic regression in landslide susceptibility mapping of the Orumba north erosion-prone region, Nigeria. *Model Earth Syst Environ* 9(1):967–986
- Pham BT, Tien BD, Pourghasemi HR, Prakash I, Dholakia MB (2017) Landslide Susceptibility Assessment in the Uttarakhand Area (India) Using GIS: a comparison study of prediction capability of naïve Bayes, multilayer perceptron neural networks, and functional trees methods. *Theor Appl Climatol* 128:255–273. <https://doi.org/10.1007/s00704-015-1702-9>
- Pradhan B, Dikshit A, Lee S, Kim H (2023) An explainable AI (XAI) model for landslide susceptibility modeling. *Appl Soft Comput* 142:110324
- Ram P, Gupta V, Devi M, Vishwakarma N (2020) Landslide susceptibility mapping using bivariate statistical method for the hilly township of Mussoorie and its surrounding areas, Uttarakhand Himalaya. *J Earth Syst Sci* 129(1):1–18
- Razavi S, Jakeman A, Saltelli A, Prieur C, Iooss B, Borgonovo E et al (2021) The future of sensitivity analysis: an essential discipline for systems modeling and policy support. *Environ Model Softw* 137:104954
- Rihan M, Bindajam AA, Talukdar S, Naikoo MW, Mallick J, Rahman A (2023) Forest fire susceptibility mapping with sensitivity and uncertainty analysis using machine learning and deep learning algorithms. *Adv Space Res* 72(2):426–443
- Saha A, Saha S (2020) Comparing the efficiency of weight of evidence, support vector machine and their ensemble approaches in landslide susceptibility modelling: a study on Kurseong region of Darjeeling Himalaya, India. *Rem Sens Appl: Soc Environ* 19:100323
- Sahin EK (2022) Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. *Geocarto Int* 37(9):2441–2465
- Sameen MI, Sarkar R, Pradhan B, Drukpa D, Alamri AM, Park HJ (2020) Landslide spatial modelling using unsupervised factor optimisation and regularised greedy forests. *Comput Geosci* 134:104336
- Shi ZM, Xiong X, Peng M, Zhang LM, Xiong YF, Chen HX, Zhu Y (2017) Risk assessment and mitigation for the Hongshiyuan landslide dam triggered by the 2014 Ludian earthquake in Yunnan, China. *Landslides* 14(1):269–285
- Sun D, Gu Q, Wen H, Xu J, Zhang Y, Shi S et al (2023) Assessment of landslide susceptibility along mountain highways based on different machine learning algorithms and mapping units by hybrid factors screening and sample optimization. *Gondwana Res* 123:89–106. <https://doi.org/10.1016/j.gr.2022.07.013>
- Sun D, Shi S, Wen H, Xu J, Zhou X, Wu J (2021a) A hybrid optimization method of factor screening predicated on GeoDetector and Random Forest for landslide susceptibility Mapping. *Geomorphology* 379:107623. <https://doi.org/10.1016/j.geomorph.2021.107623>
- Sun D, Wen H, Wang D, Xu J (2020a) A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology* 362:107201. <https://doi.org/10.1016/j.geomorph.2020.107201>
- Sun D, Xu J, Wen H, Wang D (2021b) Assessment of landslide susceptibility mapping based on Bayesian hyperparameter optimization: a comparison between logistic regression and random forest. *Eng Geol* 281:105972. <https://doi.org/10.1016/j.enggeo.2020.105972>
- Sun D, Xu J, Wen H, Wang Y (2020c) An optimized random forest model and its generalization ability in landslide susceptibility mapping: application in two areas of Three Gorges Reservoir, China. *J Earth Sci* 31:1068–1086. <https://doi.org/10.1007/s12583-020-1072-9>
- Tang C, Zhu J, Qi X, Ding J (2011) Landslides induced by the Wenchuan earthquake and the subsequent strong rainfall event: a case study in the Beichuan area of China. *Eng Geol* 122(1–2):22–33
- Tanyu BF, Abbaspour A, Alimohammadlou Y, Tecuci G (2021) Landslide susceptibility analyses using random forest, C4. 5, and C5. 0 with balanced and unbalanced datasets. *Catena* 203:105355
- Tran TH, Dam ND, Jalal FE, Al-Ansari N, Ho LS, Phong TV et al (2021) GIS-based soft computing models for landslide susceptibility mapping: a case study of Pithoragarh district, Uttarakhand state, India. *Math Probl Eng* 2021. <https://doi.org/10.1155/2021/9914650>
- Wang X, Wang D, Li X, Zhang M, Cheng S, Li S et al (2024) The prediction of cross-regional landslide susceptibility based on pixel transfer learning. *Remote Sens* 16(2):347
- Yang H, Yang T, Zhang S, Zhao F, Hu K, Jiang Y (2020) Rainfall-induced landslides and debris flows in Mengdong Town, Yunnan Province, China. *Landslides* 17(4):931–941
- Yu X, Chen H (2024) Research on the influence of different sampling resolution and spatial resolution in sampling strategy on landslide susceptibility mapping results. *Sci Rep* 14(1):1549
- Zaruba Q, Mencl V (2014) *Landslides and their control*. Elsevier



- Zhang G, Wang M, Liu K (2019) Forest fire susceptibility modeling using a convolutional neural network for Yunnan province of China. *Int J Disaster Risk Sci* 10(3):386–403
- Zhang MY, Han Y, Yang P, Wang CL (2023) Landslide displacement prediction based on optimized empirical mode decomposition and deep bidirectional long short-term memory network. *J Mt Sci* 20(3):637–656
- Zhang Y, Wen H, Xie P, Hu D, Zhang J, Zhang W (2021) Hybrid-optimized logistic regression model of landslide susceptibility along mountain highway. *Bull Eng Geol Environ* 80(10):7385–7401. <https://doi.org/10.1007/s10064-021-02415-y>
- Zhou X, Wen H, Li Z, Zhang H, Zhang W (2022) An interpretable model for the susceptibility of rainfall-induced shallow landslides based on SHAP and XGBoost. *Geocarto Int* 37(26):13419–13450
- Zhou X, Wen H, Zhang Y, Xu J, Zhang W (2021) Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. *Geosci Front* 12(5):101211. <https://doi.org/10.1016/j.gsf.2021.101211>
- Zhu C, Zhang J, Liu Y, Ma D, Li M, Xiang B (2020) Comparison of GA-BP and PSO-BP neural network models with initial BP model for rainfall-induced landslides risk assessment in regional scale: a case study in Sichuan, China. *Nat Hazards* 100(1):173–204

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.