Comparison of optimized data-driven models for landslide susceptibility mapping

Armin Ghayur Sadigh[1] · Ali Asghar Alesheikh[1] · Sayed M. Bateni[2] · Changhyun Jun[3] · Saro Lee[4,5] · Jeffrey R. Nielson[6] · Mahdi Panahi[7,8] · Fatemeh Rezaie[4,5]

Abstract

Locations prone to landslides must be identified and mapped to prevent landslide-related damage and casualties. Machine learning approaches have proven effective for such tasks and have thus been widely applied. However, owing to the rapid development of data-driven approaches, deep learning methods that can exhibit enhanced prediction accuracies have not been fully evaluated. Several researchers have compared different methods without optimizing them, whereas others optimized a single method using different algorithms and compared them. In this study, the performances of different fully optimized methods for landslide susceptibility mapping within the landslide-prone Kermanshah province of Iran were compared. The models, i.e., convolutional neural networks (CNNs), deep neural networks (DNNs), and support vector machine (SVM) frameworks were developed using 14 conditioning factors and a landslide inventory containing 110 historical landslide points. The models were optimized to maximize the area under the receiver operating characteristic curve (AUC), while maintaining their stability. The results showed that the CNN (accuracy = 0.88, root mean square error (RMSE) = 0.37220, and AUC = 0.88) outperformed the DNN (accuracy = 0.79, RMSE = 0.40364, and AUC = 0.82) and SVM (accuracy = 0.80, RMSE = 0.42827, and AUC = 0.80) models using the same testing dataset. Moreover, the CNN model exhibiting the highest robustness among the three models, given its smallest AUC difference between the training and testing datasets. Notably, the dataset used in this study had a low spatial accuracy and limited sample points, and thus, the CNN approach can be considered useful for susceptibility assessment in other landslide-prone regions worldwide, particularly areas with poor data quality and quantity. The most important conditioning factors for all models were rainfall and the distances from roads and drainages.

✉ Changhyun Jun
cjun@cau.ac.kr

✉ Fatemeh Rezaie
rezaie@kigam.re.kr

Extended author information available on the last page of the article

🖄 Springer

1 Introduction

Landslides are natural hazards that frequently occur worldwide and threaten the safety of individuals, buildings, and other infrastructures (Iverson, 2000; Kjekstad & Highland, 2009). Population growth, urban development, and indiscriminate use of natural resources have increased the susceptibility of many regions to landslides. It has been estimated that landslides are responsible for 1000 deaths and 4 billion USD of damage annually (Lee & Pradhan, 2007).

Accurate estimation of the spatial distribution of landslides and the generation of susceptibility maps are essential for hazard mitigation and urban development planning in landslide-prone areas. Classic methods, such as direct mapping through field surveying, typically involve measurements of mass displacements, which may be time-consuming, costly, and impractical over large areas (Kovács et al., 2019). Alternative methods include indirect mapping through the use of modeling techniques, relationships between landslide conditioning factors (e.g., land use, slope class, distance from roads or streams, and other site-specific properties that can be predictors of landslide susceptibility), and recorded locations of historical landslides (Guzzetti et al., 2006). The premise for using such conditioning factors with historical data is that future landslides will likely occur in locations that have similar properties as the regions in which past landslides occurred (Guzzetti et al., 2006). In general, alternative approaches to landslide susceptibility mapping (LSM) can overcome the limitations of direct mapping techniques and accelerate the production of susceptibility maps to alleviate the hazards associated with landslides worldwide.

In recent years, many such alternative approaches have been developed, ranging in type from physical models to machine learning (ML) methods (Merghadi et al., 2020). For example, deterministic models evaluate landslide susceptibility with physical laws and data including rock and soil properties, topography, and hydrological conditions (Yilmaz, 2009). However, these models may oversimplify landslide processes, and the required data are usually expensive or difficult to obtain, especially when the study area is large and heterogeneous. Heuristic approaches for LSM establish classes for judging the relative contributions of multiple landslide variables (Dahal et al., 2008a, 2008b; Dai & Lee, 2002). Although these ranking or rating methods are effective, they may be highly subjective compared with other data-driven approaches (Lee et al., 2001). Probabilistic models exploit the statistical properties of landslide factors in locations of past landslides (Constantin et al., 2011; Jaafari et al., 2014; Shirzadi et al., 2017). Other statistical models apply logistic regression, binary logistic regression (Shahabi et al., 2015; Tien Bui et al., 2016), fuzzy logic (Sur et al., 2021, 2022), and knowledge-based methods (Althuwaynee et al., 2016; Kumar & Anbalagan, 2016).

ML models typically use general-purpose learning algorithms that can identify patterns in data, including complex or nonlinear data. Diverse ML techniques have been applied to LSM and noted to achieve excellent results. For instance, a highly optimized support vector machine (SVM) workflow was noted to attain prediction accuracies > 90% (Dou et al., 2020a). Ensemble methods using random forest and decision tree achieved excellent area under the receiver operating characteristic curve (AUC) values of 0.91 and 0.98, respectively (Arabameri et al., 2020; Chowdhuri et al., 2020). However, ML workflows may be complicated, and their optimization is often challenging. Moreover, owing to the rapid development of ML techniques, many methods with potential to improve LSM predictions have not been fully evaluated.

Recent studies have applied deep learning (DL) methods for LSM. The increased complexity or numbers of layers and nodes, which makes these frameworks "deep," render them well-suited for predicting the complicated relationships between landslide conditioning factors and geographic landslide likelihoods (Schmidhuber, 2015). For example, Dou et al. (2020b) compared the performance of a deep neural network (DNN) against logistic regression (LR) and an artificial neural network (ANN) for LSM in Japan. The DNN model outperformed the LR and ANN during training and testing. Thi Ngo et al. (2021) compared convolutional neural network (CNN) and recurrent neural network (RNN) models for LSM in Iran on a national scale and reported that both models achieved AUC values higher than 0.85. These studies highlighted the potential of DL methods for LSM.

Previous studies that assessed and compared the performances of different DL methods did not consider optimization process, and thus, their conclusions may be unreliable. Although several researchers attempted to optimize their models using different optimization algorithms, they focused on a single method and compared the effects of different optimization algorithms on the model performance. In this study, we combined both objectives. Specifically, we performed a reliable and comprehensive comparison among three methods and investigated the possible performance advantage of DL methods (DNN and CNN) over ML (SVM) methods. These three methods were selected because although they have separately demonstrated strong potential for LSM, their landslide susceptibility assessment capabilities have not been comprehensively compared yet. In addition to following the methodologies of existing LSM studies that separately used SVM, DNN, and CNN methods, we optimized the model hyperparameters that control the learning process to ensure that each model reaches its maximum potential and the results can be reliably compared. The performances of the three methods in performing LSM in a landslide-prone province in Iran were assessed using a landslide inventory to evaluate the locations of previous landslides and a set of landslide conditioning factors. The AUC, which has been commonly used in ML and DL studies, was adopted as the performance metric to facilitate the comparison of the obtained results with those reported in the existing studies, for identifying the most effective approaches for LSM.

2 Study area

Iran is a mountainous country with many major population centers located on sloping terrains that are exposed to landslide hazards. In this work, the Kermanshah province in western Iran (Fig. 1), which is one of the most landslide-prone provinces was selected as the study site. Kermanshah has a total area of 95970 km^2 and is located between 33°40′ N–35°20′ N and 45°20′ E–48°10′ E in the Zagros Mountain range. The elevation in the province ranges from 116 to 3359 m above sea level. The region has average low and high temperatures of 20.61 °C and 36.1 °C, respectively, and an average annual rainfall of 500 mm, making it one of the wettest provinces in the country (http://www.kermanshahmet.ir/met/amar).

The province is partially covered with high-density vegetation, agricultural lands, regions of sparse vegetation, and plains connecting mountains and valleys. Kermanshah is seismically active, positioned over the High Zagros Fault (HZF), which is the most active fault in the area. The HZF is 1375 km long with a NW-SE bearing. The occurrence of 14 earthquakes (magnitude 4 or higher) was recorded between August 2019 and December 2021, which resulted in landslides and severe damage to infrastructure. The area consists of two geological zones,

Fig. 1 Location of the study area (Kermanshah, Iran)

covering the northeast and southwest regions of the area: Sanandaj–Sirjan, consists of sedimentary and igneous/metamorphic rock zones owing to high volcanic activity and Zagros, which covers a much larger portion of the province, consists of mostly sedimentary rock, with some ophiolites (Ao et al., 2016; Arian & Aram, 2014).

Landslides in Kermanshah province are mostly triggered by intense rainfall or seismic activities such as earthquakes. Some of the most devastating landslides in Iran's history have occurred in Kermanshah, including the largest landslide recorded in the past 20 years, which occurred in Mela Kabood with an area of effect of 4.61 km^2. This landslide was triggered by a 7.3 magnitude earthquake in November 2017. Many other regions of the province were also damaged by landslides triggered by this earthquake, such as North Dalahoo and Zanganeh outpost (Fig. 2a–c). Multiple occurrences of rockfall were also reported in Babayadegar, Northeast Dalahoo, and Piran (Fig. 2d–f). These phenomena damaged critical infrastructures, disrupted water supply lines and roads, and led to severe injuries (Bordbar et al., 2022). According to reports from Tasnim news agency (https://www.tasnimnews.com), 20 villages were evacuated, and the residents had to be relocated.

In April 2019, multiple regions in Kermanshah province were damaged by landslides caused by intense rainfall and flooding (Fig. 3). According to the Islamic Republic News Agency (https://www.irna.ir), 281 villages, 134 km of roads, 46 buildings, and the feeding pipeline of an oil refinery were damaged due to landslides.
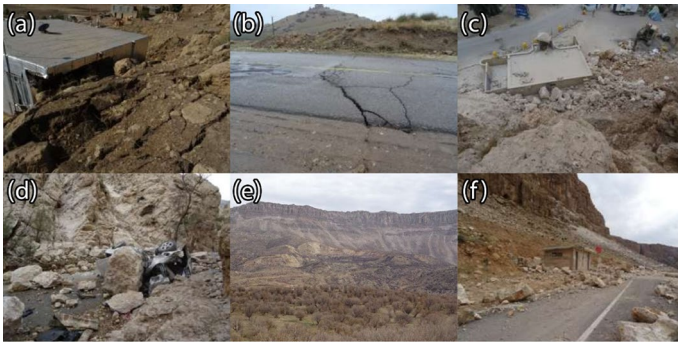
Fig. 2 Photographs of landslide and rockfall damage across Kermanshah province (Haghshenas et al., 2017) **a** Mela Kabood **b** Dalahoo **c** Zanganeh outpost **d** Babayadegar **e** Northeast Dalahoo **f** Piran (https://www.tasnimnews.com)

**Fig. 3** Roads damaged by landslides due to intense rainfall (https://www.irna.ir)

3 Materials and methods

3.1 Overview

The creation of landslide susceptibility maps through SVM, DNN, and CNN modeling consisted of a multi-step workflow, established with reference to previous studies (Fig. 4). First, the sampling points were split into testing and training groups, followed by data sourcing and processing to develop a geodatabase. Using the training sample, we extracted different formats of data from the geodatabase for model optimization and training. Multiple batches of models were trained using different sets of hyperparameters for each method. The AUC and root mean square error (RMSE) per epoch were recorded for each of the models as a graph for future comparison. After training multiple batches of models per method, the model with the highest AUC and stable RMSE graph in each batch was selected. Next, the AUC and RMSE graphs of the final set of chosen models were assessed, and the best model resulting from the optimization of the corresponding method was selected. The performances of these optimized models were validated using additional metrics. If the models exhibit satisfactory validation results, they were used to generate susceptibility maps of the area.

3.2 Geospatial dataset

Two essential datasets were acquired for this study: (1) a landslide inventory, and (2) a set of landslide conditioning factors. These data were sourced from national authorities (e.g.,
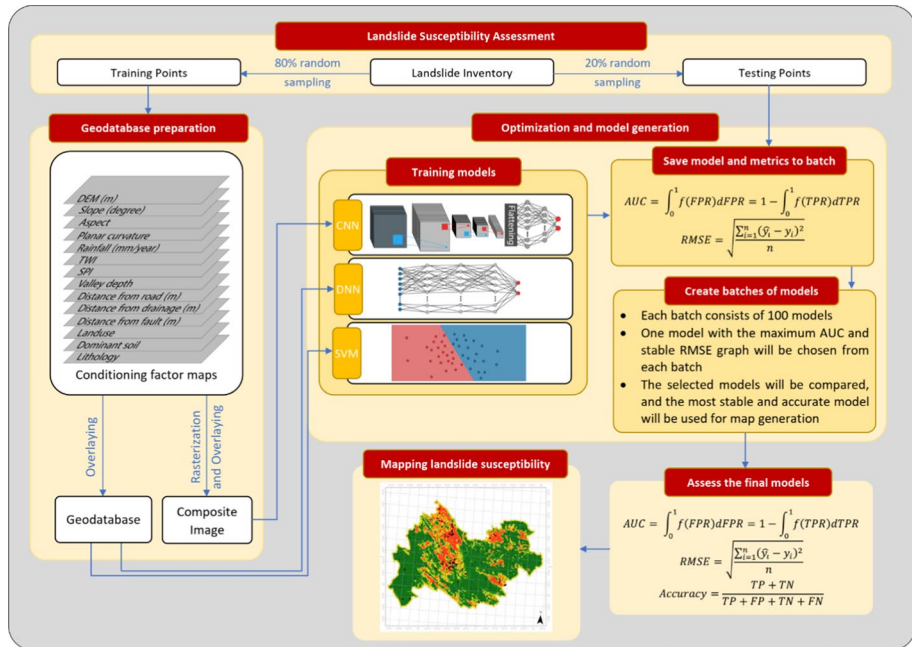
Fig. 4 Workflow of this study

geological survey and mineral exploration of Iran and Ministry of Agriculture Jihad) and various online sources as digital maps. An elevation map and its derivatives (e.g., slope, aspect, and curvature) were created using ALOS World elevation tiles that were resampled from 30 m to 85 m resolution raster images to ensure uniform resolution across the dataset.

3.2.1 Landslide inventory

The landslide inventory included 110 landslide location records from across the study area. Points for these landslides were generated from the centroid of the landslide scarp. This inventory was used to create a landslide density map, which was divided into five density classes or regions (very low, low, moderate, high, and very high). Most of the 110 landslide locations lay within the "high" and "very high" density regions. A second set of 110 locations was created to represent non-landslide locations, by randomly sampling points in the very-low-density regions. Figure 5 shows the distributions of both types of points and the density map.

This dataset was split into two groups of training and testing points. Such sets are typically partitioned to include 70–80% of the data for training and 20–30% for testing (Nefeslioglu et al., 2008; Pourghasemi & Rahmati, 2018). In this study, an 80–20% split was chosen. Stratified sampling was performed to create the training and testing sets to ensure equal numbers of points from each group of landslide and non-landslide points. The geospatial data were used to generate training data as image patches for the CNN model and as
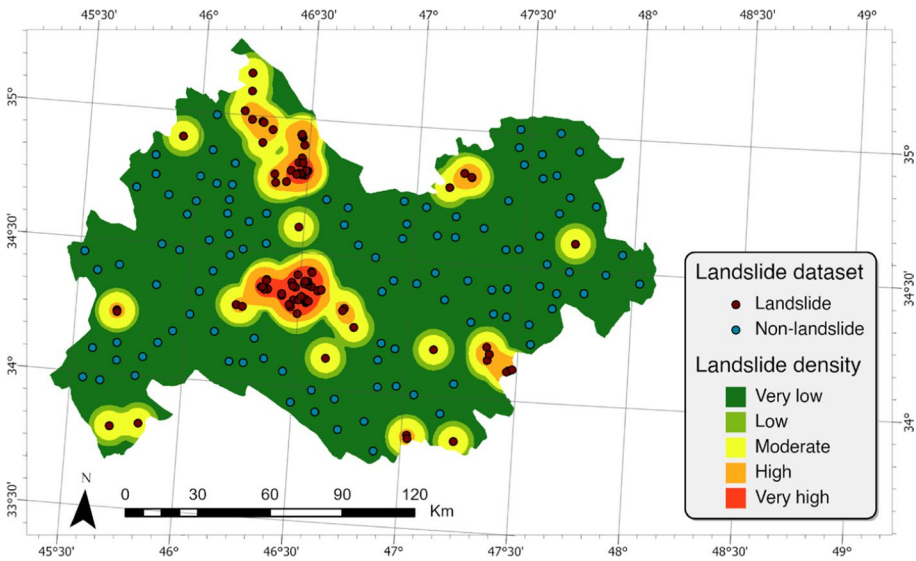
Fig. 5 Landslide density and point distribution across Kermanshah province

data tables extracted from both vector and raster datasets for the SVM and DNN models. The input datasets were extracted at the locations of landslide and non-landslide points.

3.2.2 Landslide conditioning factors

Four major groups of factors affect landslide susceptibility: geomorphological, hydrological, geological, and environmental factors. In this study, 14 landslide conditioning factors from these groups were considered: elevation; slope; aspect; planar curvature; rainfall; topographic wetness index (TWI); stream power index (SPI); valley depth; land use; dominant soil; lithology; and distances from roads, drainages, and faults. Maps was obtained for each factor, which were of different types, such as ordinal data in which the order is relevant (e.g., elevation and slope), and nominal data without any ranking or order (aspect, land use, dominant soil, and lithology). Table 1 summarizes the properties and importance of the factor maps.

Geomorphological factors play a crucial role in landslide susceptibility assessment. Elevation is one of the most commonly used factors in landslide modeling, as elevated steep slopes affect the surface reliability and stability (Umar et al., 2014). The slope angle has been widely used as a key factor in landslide modeling as it can represent the sheer stress and force and considerably affects hydrological processes (Nohani et al., 2019; Pourghasemi & Rahmati, 2018). The slope aspect, which indicates the azimuth of maximum slope, affects the amount of sunlight and rainfall received, which influences the precipitation and vegetation and root development (Jaafari, 2018; Kavzoglu et al., 2015). Curvature indicates the concavity or convexity of the surface, which is another morphological factor that can affect erosive processes and their intensity to potentially destabilize the surface. The planar curvature is the amount of curvature in a horizontal plane that determines the convergence or divergence of flows and runoffs (Fallah-Zazuli et al., 2019; Jaafari et al., 2015; Pourghasemi & Rahmati, 2018). Valley depth, which indicates the difference in elevation between

Table 1 Description of conditioning factors

Factor	Source	Resolution/scale	Significance
Elevation	DEM	85 m	Rock strength and dehydrogenase activity Ali et al., (2021)
Aspect	DEM	85 m	Amount of rainfall, morphologic structure, and vegetation composition and development Jaafari, (2018), Kavzoglu et al., (2015)
Slope	DEM	85 m	Shear stress, hydrological processes, and infrastructure development Nohani et al., (2019), Pourghasemi and Rahmati, (2018)
Planar curvature	DEM	85 m	Erosion processes and flows Fallah-Zazuli et al., (2019), Jaafari et al., (2015), Pourghasemi & Rahmati, (2018)
Rainfall	National map	1:100,000	Erosion processes, vegetation density, and runoff generation Dou et al., (2019b) Mondini et al., (2011)
TWI	DEM	85 m	Influence of topography on soil conditions and spatial wetness patterns Arabameri et al., (2019), Lee and Pradhan, (2007), Panahi et al., (2022)
SPI	DEM	85 m	Erosive power of flowing water and soil conditions Sestraş et al., (2019), Sevgen et al., (2019)
Valley depth	DEM	85 m	Stability of slopes and soil pore water pressure Hadji et al., (2018), Hakim et al., (2022), Pourghasemi et al., (2020)
Distance from roads	National map	1:250,000	Road construction and infrastructure development Fallah-Zazuli et al., (2019), Sestraş et al., (2019)
Distance from drainage	National map	1:100,000	Erosion processes, slope erosion, and slope stability Dou et al., (2019a, 2019b), Fallah-Zazuli et al., (2019), Kadirhodjaev et al., (2020)
Distance from fault	National map	1:100,000	Geomorphological processes and tectonic activities Fallah-Zazuli et al., (2019); Nguyen et al., (2019); Pham et al., (2020)
Land use	National map	1:25,000	Infrastructure development, soil stability, and vegetation density Nasiri et al., (2019), Nedbal & Brom, (2018)
Dominant soil	National map	1:500,000	Strenjlvrgth and permeability of soils, erosion processes, and shear stress Sharma et al., (2012)
Lithology	National map	1:25,000	Mineral characteristics, rock types, permeability, and surface runoff Hong et al., (2015, Reneau, (2000), Yilmaz and Ercanoglu, (2019)

the valley and upstream ridge, affects the slope stability and soil pore water pressure, which influence the landslide occurrence probabilities (Hadji et al., 2018; Hakim et al., 2022; Pourghasemi et al., 2020).

Among hydrological factors, rainfall is a notable landslide-triggering factor in the study area, which can also cause erosion, affect vegetation density, and promote the generation of runoff (Dou et al., 2019b; Mondini et al., 2011). TWI is another factor that has been commonly used in similar studies as it can clarify the influence of the topography and flow accumulation on soil conditions and spatial wetness patterns (Arabameri et al., 2019; Lee & Pradhan, 2007; Panahi et al., 2022). The SPI can reflect the erosive power of flows and its effect on the surface (Sestraş et al., 2019; Sevgen et al., 2019). Drainages and other water flow significantly influence the erosive processes, which then affect the slope stability and landslide probability (Dou et al., 2019a, 2019b; Fallah-Zazuli et al., 2019; Kadirho-djaev et al., 2020); therefore, the distance from these sources was selected as a hydrological factor.

Geological factors are another set of commonly used factors in landslide modeling. The soil texture indicates the strength and permeability of soils, which affect the erosion processes and shear stress (Sharma et al., 2012). We used the dominant soil texture or textures present in each unit to reclassify and simplify the available dataset. As another key factor, the lithology determines the mineral characteristics of different rock types, permeability of rocks, and their contribution to the generation of surface runoff (Hong et al., 2015; Reneau, 2000; Yilmaz & Ercanoglu, 2019). Faults represent notable geological factors as they cause tectonic activities that can trigger landslides. Moreover, faults affect the geomorphology of the surface by deforming it, thereby influencing the slope stability (Fallah-Zazuli et al., 2019; Nguyen et al., 2019; Pham et al., 2020). Therefore, the distance from faults was introduced as a factor.

The last set of impact factors used in landslide susceptibility assessment is environmental factors. In this study, land use and distance from roads were used as environmental factors. Land being used for different purposes has different properties, with some being more susceptible to landslides. Depending on the type of land use or land cover, different units can indicate industrial development status or expectations, their effect on soil stability, and current or expected vegetation density (Nasiri et al., 2019; Nedbal & Brom, 2018), which can alter the landslide probability. The distance from roads and transportation networks was used as an impact factor as their development may have destabilized the soil in their vicinity, and the force applied to the ground by traversing vehicles can increase the chance of landslide occurrence (Fallah-Zazuli et al., 2019; Sestraş et al., 2019).

All the conditioning factor maps were rasterized with a resolution of 85 m. The raster layers were resampled, and polygon layers were rasterized using appropriate toolboxes to create a raster with the same grid as the other layers. Figure 6 shows an overview of the conditioning factor maps.

3.2.3 Multicollinearity analysis

Before the modeling phase, any multicollinearity among selected parameters must be analyzed and identified. Removing factors with high correlation helps decrease the data dimensionality and model complexity, which can shorten the training phase and prevent models from becoming biased toward certain factors (Tehrany et al., 2019). In this study, the extent of correlation between the conditioning factors was evaluated using the variance inflation factor (VIF), calculated as
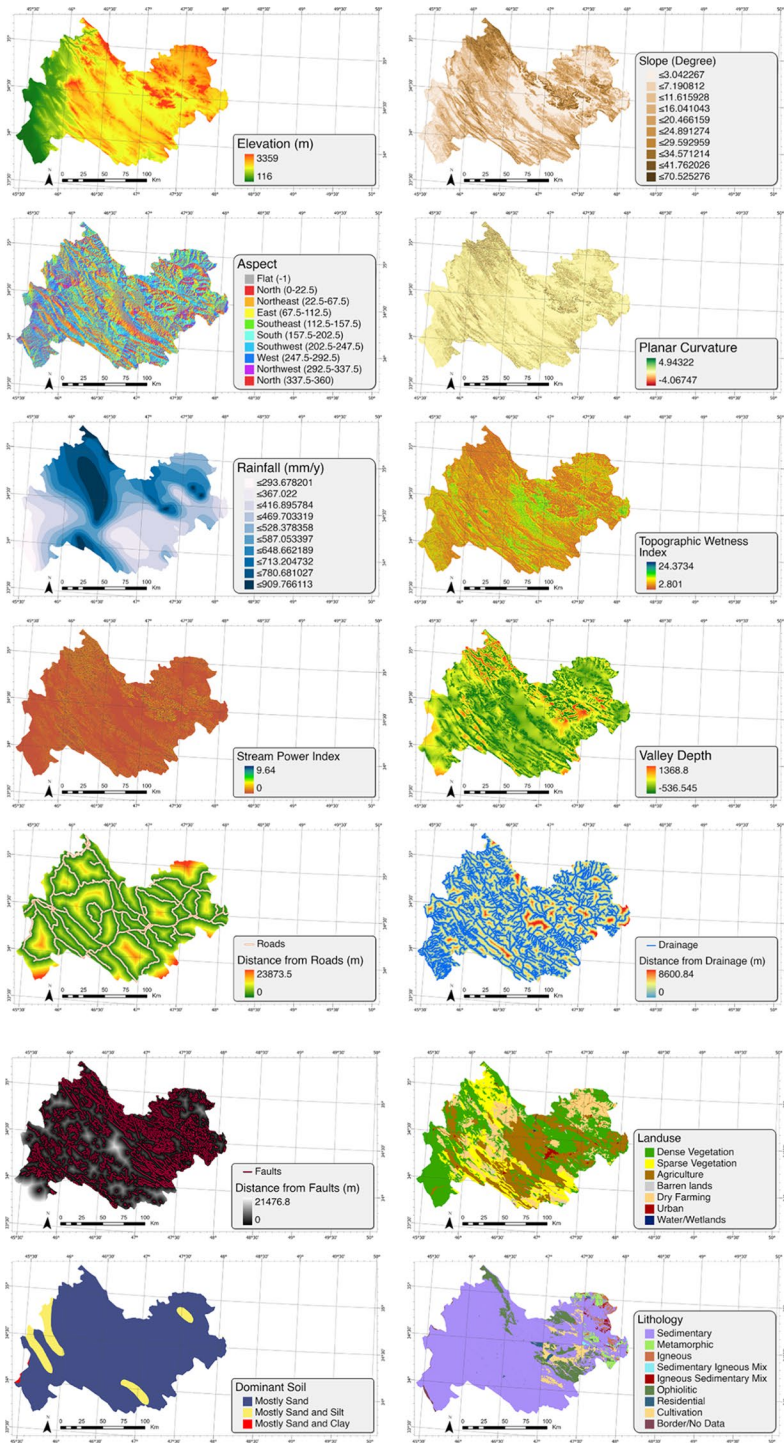
Fig. 6 Conditioning factor maps

$$VIF_i = \frac{1}{1 - R_i^2} \tag{1}$$

where R_i is the multi-correlation coefficient between the ith factor and other conditioning factors. According to the literature (Kalantar et al., 2019; Roy & Saha, 2019), factors with VIF > 5 are considered to have high multicollinearity and should be removed or combined with another related variable into a single index (O'brien, 2007).

3.3 Numerical modeling methods

3.3.1 SVM

SVM is a supervised learning method that originated from statistical learning theory and the structural risk minimization principle (Lee et al., 2017b; Vapnik, 1995). This method can be used for both classification and regression (Cristianini & Shawe-Taylor, 2000; Vapnik, 1995). SVM determines a line or hyper-plane in the multi-dimensional space of training samples to separate the samples into two classes with optimal margins (distances from the separating surface and closest point) (Xu et al., 2012; Yao et al., 2008). Larger margins have been noted to be more resistant to noise (Kanevski, 2009). The initial space of an SVM can be transformed into a feature space with a higher dimensionality by using a kernel function. This transformation can increase the linear separation between the points (Abe, 2010; Chang & Lin, 2001) and allow the SVM to function as a nonlinear classifier as well. The commonly used kernel functions are the radial basis function (RBF) and linear, sigmoid, and polynomial kernels.

The RBF kernel, defined in Eq. (2), has been applied successfully in similar nonlinear regression problems, such as flood modeling (Tehrany et al., 2015), and was thus used in this work after numerous optimization trials.

$$K\left(\vec{x}, \overrightarrow{x'}\right) = \exp\left(\left\|-\frac{x - x'^2}{2\sigma^2}\right\|\right) \tag{2}$$

where $\|x - x'\|^2$ is the squared Euclidean distance between the two feature vectors, and σ is a tuning parameter (Tien Bui et al., 2016). In addition to the kernel function, the SVM has two other key hyperparameters: gamma ($\gamma = \frac{1}{2\sigma^2}$) and regularization. Gamma determines the spread of classification boundaries, thereby affecting the flexibility of the model in classifying new data samples close to the classification boundaries. The regularization parameter is used for error control and relates to the tolerance of misclassification. With lower tolerance, the boundaries become stricter, resulting in lower errors and higher accuracies. However, the model reliability may not necessarily increase and must be assessed using the test dataset after training.

3.3.2 DNN

The DNN model is an ANN with more than one hidden layer (LeCun et al., 2015). This deeper structure allows the model to extract more complex features and patterns from the input data (Schmidhuber, 2015). DNN models alter data from one depiction to another and are therefore widely used for pattern detection and classification in nonlinear problems such as landslide zoning.

The first and last layers of a DNN are the input and output layers, respectively. Several hidden layers are present between the input and output layers. Each node acts as a variable, containing the value calculated from the previous layer's nodes and an activation function that transforms the previous values into a new range. Given the structure of a DNN, it is necessary to determine the proper combination of the number of layers, number of nodes in each layer, and activation function in each layer. Increasing the node and layer counts can enable the recognition of more complex data relationships but also increases the number of parameters to be calculated, which can hinder the learning process. The activation function determines the data transformation between layers, which influences the feature classification quality. Various activation functions, such as the rectified linear unit (ReLU) and sigmoid, linear, and tanh functions, are available to accomplish diverse modeling objectives.

DNNs are large neural networks and are thus susceptible to overfitting and deteriorate performance when provided new data samples. Overfitting causes the network to mimic sample properties, thereby reducing the model flexibility. Dropout layers are typically used to prevent this phenomenon. Through the introduction of dropout layers, a random number of layer nodes are ignored, preventing weight update. The dropout rate, as a percentage, determines the number of dropout nodes selected. Notably, dropout rates that are too low will not prevent overfitting, whereas excessively high rates may result in underfitting and prevent convergence.

Other parameters, such as the batch size and learning rate, also affect the DNN training and convergence speeds. During DNN training, optimal model weights are determined over a series of training epochs in which the weights are iteratively updated to decrease the overall model error. The batch size controls the number of sample points per epoch and can thus be appropriately selected to balance the training accuracy and speed. The learning rate controls how much the weights are allowed to change when updated. A high learning rate will result in overshooting while updating weights, which may retard or prevent model convergence or destabilize the learning process. In contrast, a low learning rate will drastically slow down model convergence and increase the number of training epochs needed to reach adequate metrics.

3.3.3 CNN

CNNs are modified DNNs that specialize in processing images or gridded data with convolutional, pooling, flattening, and fully connected layers (Li et al., 2021). Without these layers, image processing with a DNN may require excessive data samples and result in a slow and process-heavy training phase (Lee et al., 2017a). Convolution layers apply filters or kernels over regions of images and then pass results to the next layer, thereby creating feature maps. Pooling layers decrease the number of pixels, and thus, the image size and overall parameter count, while preserving important features. Flattening layers convert the resulting feature map to a fully connected layer with an equal number of neurons. The product is the input layer of an ANN or DNN, which is responsible for the classification of the features extracted from the initial image data.

The number of convolution layers influences feature detection in the input data, in addition to the learning speed and parameter count. Similar to the node count in the DNNs, the number of filters in each convolution layer determines the amount of data transferred from one layer to another, which defines the balance between speed and performance. Two other parameters that affect information preservation are the kernel size and stride. The kernel size determines the data window used in each step for feature

detection. The optimal choice for this parameter depends on the problem: Choosing a small kernel size may result in the loss of possible spatial data patterns or error mitigation, whereas a large kernel size may increase the number of parameters and negatively affect the training process and convergence. In this study, 3×3 and 5×5 pixel kernels were tested for each layer to optimize the model performance. The stride relates to the travel of the kernel or pooling window over the input data. Increased stride may decrease the parameter count but result in overshooting and the loss of important data features. In this study, stride values of 1, 2, and 3 pixels were tested for the max pooling layers.

3.4 Hyperparameter optimization

The hyperparameters of the CNN, DNN, and SVM models (Table 2) affect their learning rate and stability, as discussed in the previous sections. Therefore, we determined the optimal combination of these hyperparameters for each algorithm.

Hyperparameter tuning was conducted during the training phase. All the processing and modeling steps were programmed in Python (Release 3.7). Keras library was used for the CNN and DNN model development, and SciKit-learn library was used for SVM modeling. The CNN and DNN models were optimized using the Optuna library (Akiba et al., 2019), and the SVM model was optimized using SciKit-learn's built-in grid search tool. Table 3 summarizes the software and hardware used for the computational process.

Table 2 Hyperparameters of the considered models

Model	Hyperparameters
SVM	Kernel
	Gamma
	Regularization
DNN	Number of hidden layers
	Number of nodes in each hidden layer
	Activation function of each hidden layer
	Dropout rate of each hidden layer
	Learning rate
	Batch size
CNN	Number of convolution layers
	Filter count of each convolution layer
	Kernel size for each convolution layer
	Activation function for each convolution layer
	Stride of each max pooling layer
	Number of hidden fully connected layers
	Number of nodes in each fully connected layer
	Activation function for each fully connected layer
	Learning rate
	Batch size

Table 3 Software and hardware specifications	Software/hardware	Parameter
	Software environment	Python (Release 3.7)
		TensorFlow (2.3.0) + Keras (2.4.3) + SciKit-learn (0.23.2)
		Optuna (2.9.1)
		ArcGIS Pro 2.5 + SAGA GIS 7.9.1
	CPU	Intel Core i7-6700HQ@3.1 GHz
	GPU	Nvidia GTX 960M
	RAM	16 GB DDR4

3.5 Evaluation of results

The model performances were evaluated using the training and testing datasets. The performance metrics used during the testing step were the classification accuracy and AUC calculated as follows:

$$AUC = \int_0^1 f(FPR)dFPR = 1 - \int_0^1 f(TPR)dTPR \tag{3}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

where TPR and FPR are the true positive rate and false positive rate, respectively. TPR indicates the percentage of landslide samples correctly classified, and FPR represents the percentage of non-landslide samples misclassified. FN, TN, TP, and FP denote the false negative, true negative, true positive, and false positive, respectively (Darabi et al., 2021). The accuracy was calculated as the ratio of correct predictions among all predictions. The error values for all models were calculated as the RMSE.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n \left(Y_{\text{Actual}} - Y_{\text{Predicted}} \right)^2}{n}} \tag{5}$$

where Y_{Actual} and $Y_{\text{Predicted}}$ are the real and predicted values, respectively, and n is the total number of samples. Following Khosravi et al. (2016), AUC values in the ranges 0.5–0.6, 0.6–0.7, 0.7–0.8, 0.8–0.9, and 0.9–1.0 were used to indicate poor, moderate, good, very good, and excellent performances, respectively.

3.6 Susceptibility mapping

After optimizing the hyperparameters to maximize the model performance, the final models were used to generate the susceptibility maps. This process resulted in a raster map of landslide susceptibility indexes per model, enumerated with values ranging from 0 to 1, where 0 indicated similarity to non-landslide samples, and 1 indicated similarity to historical landslide occurrences. The resulting map values were then reclassified into five categories corresponding to very low, low, medium, high, and very high

landslide-prone areas. This classification was performed using the quantile method, which has been commonly used for threshold value determination with skewed data (Akgun et al., 2012).

4 Results

4.1 Multicollinearity analysis

As discussed, the VIF method was used to identify any multicollinearity among the proposed impact factors. The results indicated the lack of any multicollinearity among the factors, as the VIF values of all the conditioning factors were lower than 5 (Table 4), and therefore, all 14 factors were used in the modeling process.

4.2 Optimization results

The optimized hyperparameters for the CNN, DNN, and SVM models are listed in Table 5. The CNN optimization process recommended the use of one or two hidden layers in most candidate models, resulting in a simpler model. Moreover, introduction of the dropout did not benefit the CNN model, and including a dropout rate resulted in fewer candidate models in each test batch.

The DNN model was not sensitive to activation functions. All three functions (Softmax, ReLU, and linear) were interchangeably used across layers among the candidate models. Optimized dropout rates rarely exceeded 10% in the candidate models, potentially because of the small sample size. The DNN depth was noted to be more important than the layer size (neuron count). Models with more layers and fewer neurons in each layer were more common candidates in test batches than shallow models with neuron counts higher than 50.

Table 4 Variance inflation factor (VIF) for landslide conditioning factors

Factor	VIF
Elevation	1.78
Aspect	1.24
Slope	2.24
Planar curvature	1.20
Rainfall	1.45
TWI	2.07
SPI	1.31
Valley depth	1.76
Distance from roads	1.30
Distance from drainage	1.27
Distance from fault	1.40
Land use	1.30
Soil	1.15
Lithology	1.40

Table 5 Hyperparameter optimization results

Hyperparameters	CNN	DNN	SVM
Kernel	–	–	RBF
Gamma	–	–	0.125
Regularization	–	–	4
Number of hidden layers	1	4	–
Number of nodes in hidden layers	32	22 – 24 – 24 – 22	–
Activation function of layers	Softmax	Softmax – Linear – ReLU – Softmax	–
Dropout rate of hidden layers	0%	16% – 4% – 2% – 6%	–
Learning rate	0.0006	0.0104	–
Batch size	25	35	–
Number of convolution layers	2	–	–
Filter count of convolution layers	49–37	–	–
Kernel size of convolution layers	5–3	–	–
Activation function of convolution layers	Linear–ReLU	–	–
Stride of each max pooling layer	1–1	–	–
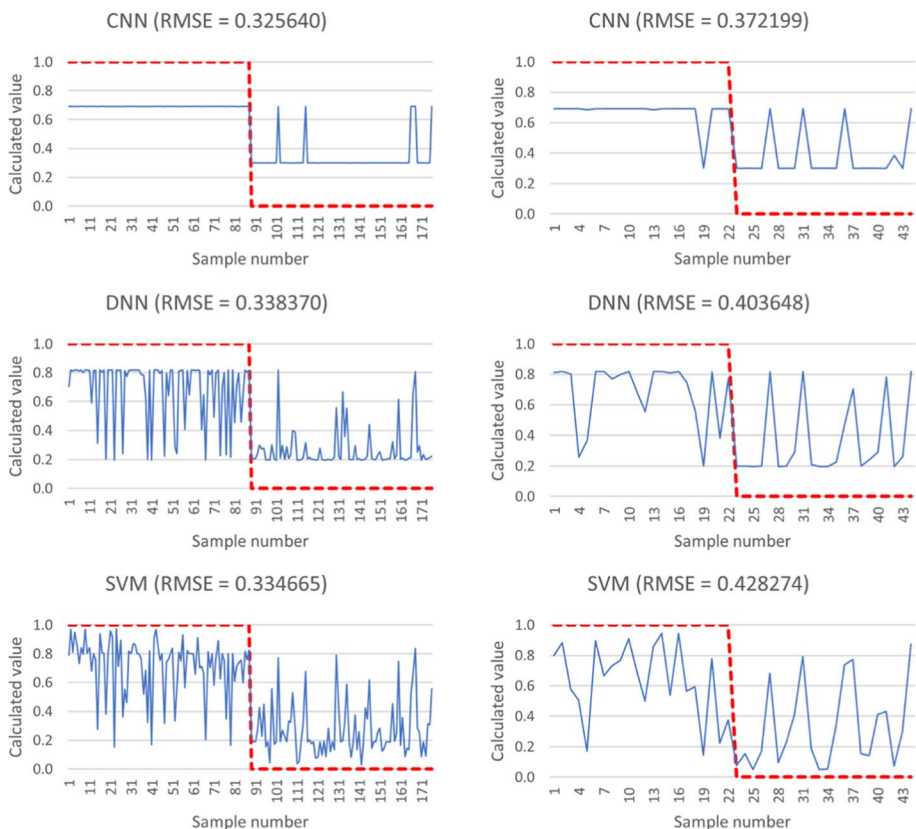
Fig. 7 Expected (red lines) and predicted (blue lines) values during training (left) and testing (right)

4.3 Model validation

Figure 7 shows model predictions of landslide susceptibility, during both training and testing, and their error values. The plots show the expected values (red lines) and calculated values (blue lines) of each sample in the training and testing sets. RMSE values were calculated for each dataset and model. Despite more instances of predictions being closer to the expected values (0 and 1), the DNN and SVM models had more instances with larger errors than the CNN model, resulting in higher overall RMSE errors over both the training and testing datasets.

All models achieved satisfactory classification accuracy, RMSE, and AUC values (Table 6). The CNN model (88% classification accuracy during testing) was more accurate than the DNN model (79% classification accuracy during testing) and SVM model (80% classification accuracy during testing). The SVM model was slightly more accurate than the DNN model, despite having a slightly higher RMSE error (approximately 0.43 and 0.40 in testing, respectively). In terms of the AUC values, during testing, the CNN model (AUC = 0.88) outperformed both the DNN (AUC = 0.82) and SVM (AUC = 0.80) models. Moreover, the CNN model exhibited a higher robustness than the other models, owing to its smaller difference in the testing and training AUC values. Specifically, the CNN was slightly more robust than the DNN model, but considerably more robust than the SVM model.
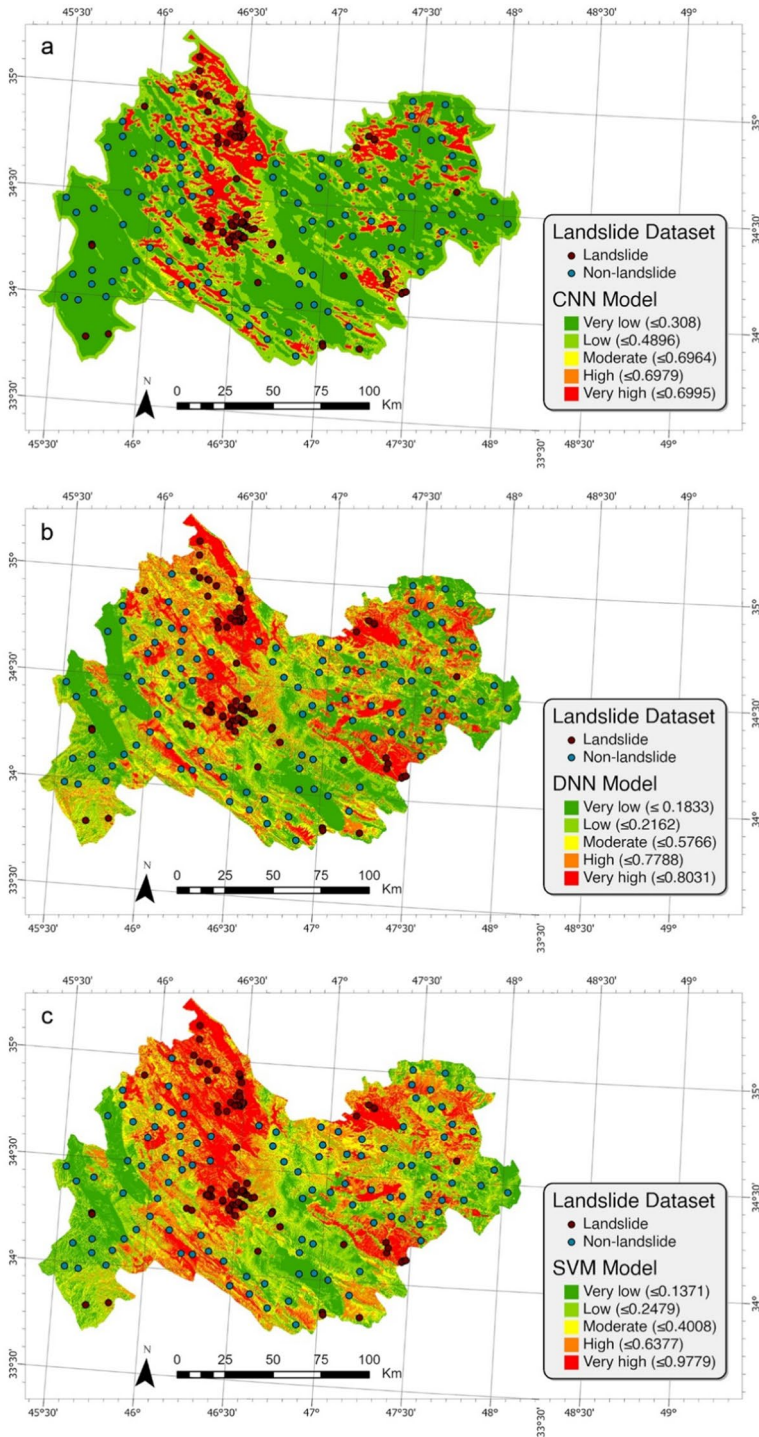
4.4 Susceptibility maps

Landslide susceptibility maps were generated with each model (Fig. 8). The maps contained comparable land-area proportions of the five different susceptibility classes (Figs. 8 and 9). The CNN predicted considerably more land area in the very low susceptibility class and considerably less land area in the moderate and high classes. These results could be explained by the fact that the CNN used the neighboring class properties in each pixel. Because very high and very low index values determined the error (which is the loss metric to be minimized in the training phase), the model tended to prioritize them over intermediate values. Therefore, fewer intermediate values were predicted. For example, pixels of high susceptibility were close to pixels of very high susceptibility and were thus classified as having very high susceptibility. The percentages for each susceptibility class for each model are shown in Fig. 9.

4.5 Factor importance analysis

To identify the conditioning factors that most notably affected the LSM, the Relief-F method was used to calculate the importance factors. Figure 10 presents a comparison of

Table 6 Model performance metrics

Model	Accuracy		RMSE		AUC	
	Training	Testing	Training	Testing	Training	Testing
CNN	0.96	0.88	0.32564	0.37220	0.96	0.88
DNN	0.81	0.79	0.33837	0.40364	0.91	0.82
SVM	0.86	0.80	0.33466	0.42827	0.93	0.80

Fig. 8 Landslide susceptibility maps produced by the **a** CNN, **b** DNN, and **c** SVM models
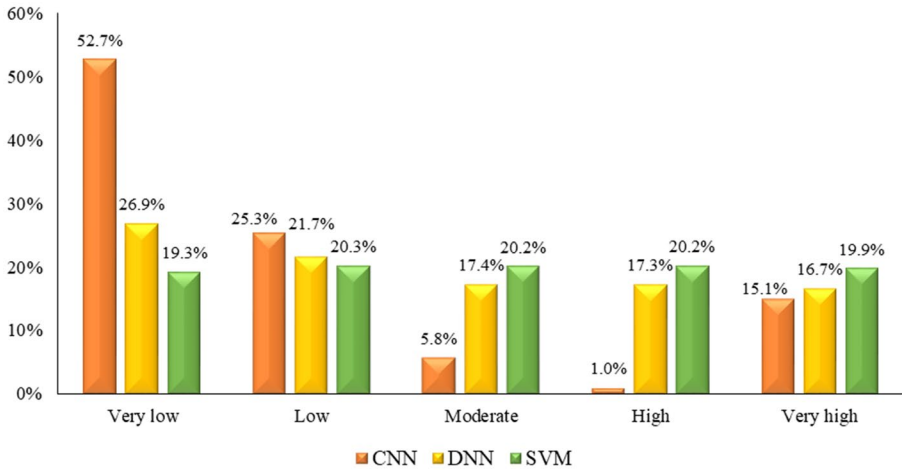
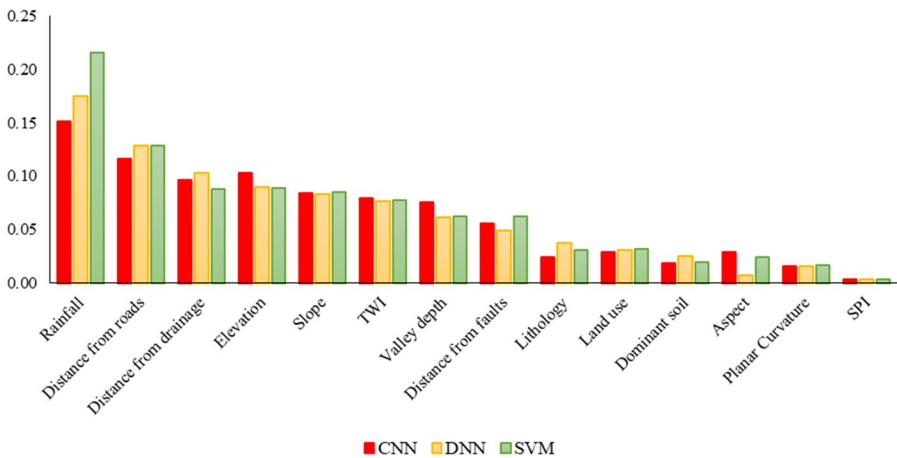**Fig. 9** Percentages of each landslide susceptibility class

Fig. 10 Comparison of the importance of conditioning factors between models

the importance factors for each model. The most important conditioning factors among all models were the rainfall and distances from roads and drainages, followed closely by elevation, slope, TWI, valley depth, and distance from faults. The significance of the remaining factors was considerably lower than these factors.

5 Discussion

5.1 Importance of localized susceptibility assessment

Landslide susceptibility in Iran has previously been studied with various approaches, such as fuzzy analytic network process (Alilou et al., 2019) and ML techniques (Pourghasemi & Rahmati, 2018; Shirzadi et al., 2019). The existing studies were conducted in different

regions with different climates and geological properties. Although the models attained high accuracy, they yielded differing results regarding the importance of conditioning factors. In other words, although ML and DL models can be successfully applied to diverse regions, they must represent the regional differences associated with landslide factors and triggering mechanisms. For example, although Thi Ngo et al. (2021) used the CNN and RNN methods to perform a national-scale LSM of Iran, their results are not necessarily reliable or correct for every region in the country, and the use of DL methods necessitates more comprehensive research.

5.2 Model performance

The ML and DL methods evaluated in this study have been separately used in existing landslide studies and noted to yield promising results (Dou et al., 2020a, 2020b; Thi Ngo et al., 2021). A direct comparison of the three methods combined with parameter optimization suggested that the CNN model has a performance advantage over the more commonly used SVM and DNN models. Specifically, the CNN model exhibited higher AUC and accuracy values across the database and the smallest RMSE error. Moreover, the CNN's AUC during testing (0.88) was comparable to the CNN performance (0.85) reported by Thi Ngo et al. (2021). However, the AUC was slightly lower than the values for a DNN (0.90–0.92) reported by Dou et al. (2020b) and an SVM model (0.74–0.91) reported by Dou et al. (2020a).

5.3 Factor importance

We observed that the rainfall, distance from roads, distance from drainage, and elevation were the most notable factors affecting LSM. Rainfall is known to be a key factor affecting landslide occurrence in the area. In particular, Kermanshah experiences intense rainfalls and has a high amount of annual rainfall compared to the national average, which promotes soil destabilization. Road construction and usage lead to further destabilization, which renders the land ready to collapse and cause a landslide. Proximity to drainage network affects the land through erosion and the provision of unstable layers of runoffs, which can lead to mass movements. These movements and drainage patterns have also been known to affect the landslide probability. Finally, elevation influences the landslide probability by indirectly affecting the precipitation and vegetation cover. Many researchers have highlighted the significance of the distances from roads (Dao et al., 2020) and drainage (Dao et al., 2020; Kalantar et al., 2020), slope and elevation (Dou et al., 2020a; Liang et al., 2021; Mandal et al., 2021), rainfall (Mandal et al., 2021), and TWI (Liang et al., 2021; Panahi et al., 2020) in LSM. In contrast, other researchers have indicated that rainfall (Liang et al., 2021), slope (Dao et al., 2020), and land use (Panahi et al., 2020) are not highly important. The variable importance results obtained in this study are moderately different from those reported by Thi Ngo et al. (2021) based on their national-scale LSM of Iran. While we determined the rainfall, distance from roads, distance from drainage, and elevation to be the most important factors, Thi Ngo et al. (2021) deemed these factors moderately important and mentioned slope as the most important factor. Nevertheless, there are several similarities between the results of the two studies, such as the aspect and elevation having low and moderate importance values, respectively. These comparisons

emphasize that LSM should be conducted locally in countries with diverse climate and geological properties.

5.4 Limitations and scope for future studies

The two main limitations of this study compared with similar previous studies are the smaller sample size of the landslide inventory and lower accuracy and resolution of the input data and certain factor maps. As such, the results of developing a single model and comparing its performance with that of the existing studies is not reliable. To address these limitations, we developed multiple models using the same limited dataset available and assessed their potential in modeling such data. Although all the developed models exhibited satisfactory performance, model performance can be enhanced by using larger and more accurate datasets.

The sample size is a key factor affecting model training. In this study, 110 landslides locations were combined with 110 randomly sampled non-landslide locations and split into 80–20% subsets for training and testing data, respectively. The resulting training (176) and testing (44) points represented small samples for DL algorithms. For DL models, such as CNNs and DNNs, large sample sizes are recommended, especially if the models are configured with high levels of complexity (e.g., high hidden layer or convolution layer count), as shown by previous similar studies, where 440 total points were used on average (Dao et al., 2020; Yao et al., 2020). A small sample size can limit the complexity and thus suppress the advantages of DL approaches. As a structured learning method, the SVM was less affected by sample size compared with the CNN and DNN and exhibited a performance comparable with that of the DNN model.

In future work, several strategies can be implemented to alleviate the effects of small samples sizes. Additional data on landslide locations can likely be acquired through other methods, such as remote sensing, thereby increasing the sample size (Kalantar et al., 2020; Liang et al., 2021). Moreover, semi-supervised learning can be introduced to add more samples to each class and update uncertain labels in training iterations (Yao et al., 2020). Lastly, RNNs, which use neural nodes that keep historical information from previous samples and steps, may be a promising model for solving the problem of interest, having been shown to improve the performance of models with small sample sizes (Xiao et al., 2018).

The accuracy of input data may also limit the accuracy of the models. In this study, certain conditioning factor maps had scales ranging from 1:100,000 to 1:500,000, smaller than those used in similar previous studies. The larger pixel size of the output maps led to a low spatial resolution, and potentially, a low accuracy. The data resolution can be enhanced by using more detailed mapping survey data or remote sensing products, which can help improve data resolution, and possibly model accuracy.

6 Conclusions

The performances of CNN, DNN, and SVM algorithms for LSM in Kermanshah, Iran were evaluated and compared. The hyperparameters were optimized to ensure that the models achieve their peak performance values to conduct a reliable comparison. The results indicated that the CNN (AUC = 88%) outperformed the DNN (AUC = 82%) and

SVM (AUC = 80%) models for LSM. Moreover, the CNN model was more robust than the other models, given the smaller difference in its AUC values for the training and testing datasets. In addition, the CNN predicted the locations of landslide and non-landslide points with the lowest overall RMSE. The superiority of the CNN was attributable to the use of a dataset with lower spatial accuracy and limited number of sample points compared with those used in similar studies conducted worldwide. In other words, the CNN model could more effectively handle datasets with low data quality and quantity than the other proposed models in similar situations. Although these three data-driven techniques had not been directly optimized and compared for LSM prior to this study, their individual performances, in terms of the AUC values, were comparable to those reported previously. Therefore, the CNN may be a valuable tool for LSM to support future planning and development in other landslide-prone regions worldwide, especially in areas with limited data availability or quality.

Data availability Some or all of the data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abe, S. (2010). *Support vector machines for pattern classification*. Springer: Advances in Pattern Recognition.

Akgun, A., Sezer, E. A., Nefeslioglu, H. A., Gokceoglu, C., & Pradhan, B. (2012). An easy-to-use MATLAB program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm. *Computers & Geosciences, 38*, 23–34. https://doi.org/10.1016/j.cageo.2011.04.012

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2623–2631.

Ali, S. A., Parvin, F., Vojteková, J., Costache, R., Linh, N. T. T., Pham, Q. B., Vojtek, M., Gigović, L., Ahmad, A., & Ghorbani, M. A. (2021). GIS-based landslide susceptibility modeling: A comparison between fuzzy multi-criteria and machine learning algorithms. *Geoscience Frontiers, 12*, 857–876. https://doi.org/10.1016/j.gsf.2020.09.004

Alilou, H., Rahmati, O., Singh, V. P., Choubin, B., Pradhan, B., Keesstra, S., Ghiasi, S. S., & Sadeghi, S. H. (2019). Evaluation of watershed health using Fuzzy-ANP approach considering geo-environmental and topo-hydrological criteria. *J Environ Manag, 232*, 22–36. https://doi.org/10.1016/j.jenvman.2018.11.019

Althuwaynee, O. F., Pradhan, B., & Lee, S. (2016). A novel integrated model for assessing landslide susceptibility mapping using CHAID and AHP pair-wise comparison. *International Journal of Remote Sensing, 37*, 1190–1209. https://doi.org/10.1080/01431161.2016.1148282

Ao, S., Xiao, W., Khalatbari Jafari, M., Talebian, M., Chen, L., Wan, B., Ji, W., & Zhang, Z. (2016). U-Pb zircon ages, field geology and geochemistry of the Kermanshah ophiolite (Iran): From continental rifting at 79Ma to oceanic core complex at ca. 36Ma in the southern Neo-Tethys. *Gondwana Research, 31*, 305–318. https://doi.org/10.1016/j.gr.2015.01.014

Arabameri, A., Karimi-Sangchini, E., Pal, S. C., Saha, A., Chowdhuri, I., Lee, S., & Tien Bui, D. (2020). Novel credal decision tree-based ensemble approaches for predicting the landslide susceptibility. *Remote Sens (basel), 12*, 3389. https://doi.org/10.3390/rs12203389

Arabameri, A., Pradhan, B., Rezaei, K., Sohrabi, M., & Kalantari, Z. (2019). GIS-based landslide susceptibility mapping using numerical risk factor bivariate model and its ensemble with linear multivariate regression and boosted regression tree algorithms. *Journal of Mountain Science, 16*, 595–618. https://doi.org/10.1007/s11629-018-5168-y

Arian, M., & Aram, Z. (2014). Relative tectonic activity classification in the Kermanshah area, western Iran. *Solid Earth, 5*, 1277–1291. https://doi.org/10.5194/se-5-1277-2014

Bordbar, M., Aghamohammadi, H., Pourghasemi, H. R., & Azizi, Z. (2022). Multi-hazard spatial modeling via ensembles of machine learning and meta-heuristic techniques. *Science and Reports, 12*, 1451. https://doi.org/10.1038/s41598-022-05364-y

Chang, C.-C., & Lin, C.-J. (2001). Training *v*-support vector classifiers: theory and algorithms. *Neural Computation, 13*, 2119–2147. https://doi.org/10.1162/089976601750399335

Chowdhuri, I., Pal, S. C., Arabameri, A., Ngo, P. T. T., Chakrabortty, R., Malik, S., Das, B., & Roy, P. (2020). Ensemble approach to develop landslide susceptibility map in landslide dominated Sikkim Himalayan region India. *Environmental Earth Science, 79*, 476. https://doi.org/10.1007/s12665-020-09227-5

Constantin, M., Bednarik, M., Jurchescu, M. C., & Vlaicu, M. (2011). Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). *Environment and Earth Science, 63*, 397–406. https://doi.org/10.1007/s12665-010-0724-y

Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press.

Dahal, R. K., Hasegawa, S., Nonomura, A., Yamanaka, M., Dhakal, S., & Paudyal, P. (2008). Predictive modelling of rainfall-induced landslide hazard in the Lesser Himalaya of Nepal based on weights-of-evidence. *Geomorphology, 102*, 496–510. https://doi.org/10.1016/j.geomorph.2008.05.041

Dahal, R. K., Hasegawa, S., Nonomura, A., Yamanaka, M., Masuda, T., & Nishino, K. (2008). GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environmental Geology, 54*, 311–324. https://doi.org/10.1007/s00254-007-0818-3

Dai, F. C., & Lee, C. F. (2002). Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology, 42*, 213–228. https://doi.org/10.1016/S0169-555X(01)00087-3

Darabi, H., Torabi Haghighi, A., Rahmati, O., Jalali Shahrood, A., Rouzbeh, S., Pradhan, B., & Tien Bui, D. (2021). A hybridized model based on neural network and swarm intelligence-grey wolf algorithm for spatial prediction of urban flood-inundation. *J Hydrol (Amst), 603*, 126854. https://doi.org/10.1016/j.jhydrol.2021.126854

Dou, J., Yunus, A. P., Bui, D. T., Merghadi, A., Sahana, M., Zhu, Z., Chen, C.-W., Han, Z., & Pham, B. T. (2020). Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides, 17*, 641–658. https://doi.org/10.1007/s10346-019-01286-5

Dou, J., Yunus, A. P., Merghadi, A., Shirzadi, A., Nguyen, H., Hussain, Y., Avtar, R., Chen, Y., Pham, B. T., & Yamagishi, H. (2020). Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning. *Science of The Total Environment, 720*, 137320. https://doi.org/10.1016/j.scitotenv.2020.137320

Dou, J., Yunus, A. P., Tien Bui, D., Merghadi, A., Sahana, M., Zhu, Z., Chen, C.-W., Khosravi, K., Yang, Y., & Pham, B. T. (2019). Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Science of the Total Environment, 662*, 332–346. https://doi.org/10.1016/j.scitotenv.2019.01.221

Dou, J., Yunus, A. P., Xu, Y., Zhu, Z., Chen, C.-W., Sahana, M., Khosravi, K., Yang, Y., & Pham, B. T. (2019). Torrential rainfall-triggered shallow landslide characteristics and susceptibility assessment using ensemble data-driven models in the Dongjiang Reservoir Watershed, China. *Natural Hazards and Earth System Sciences, 97*, 579–609. https://doi.org/10.1007/s11069-019-03659-4

Fallah-Zazuli, M., Vafaeinejad, A., Alesheykh, A. A., Modiri, M., & Aghamohammadi, H. (2019). Mapping landslide susceptibility in the Zagros Mountains, Iran: A comparative study of different data mining models. *Earth Science Informatics, 12*, 615–628. https://doi.org/10.1007/s12145-019-00389-w

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., & Galli, M. (2006). Estimating the quality of landslide susceptibility models. *Geomorphology, 81*, 166–184. https://doi.org/10.1016/j.geomorph.2006.04.007

Hadji, R., Achour, Y., Hamed, Y., 2018. Using GIS and RS for Slope Movement Susceptibility Mapping: Comparing AHP, LI and LR Methods for the Oued Mellah Basin, NE Algeria, In *Advances in Science, Technology and Innovation*. Springer, Cham, pp. 1853–1856. https://doi.org/10.1007/978-3-319-70548-4_536

Haghshenas, E., Ashayeri, I., Mousavi, S.M., Beiglari, M., 2017. November 2017 Sarpol Zahab earthquake. report (5th edition). Kermanshah. http://www.iiees.ac.ir/. (Last access 21 July 2022)

Hakim, W. L., Rezaie, F., Nur, A. S., Panahi, M., Khosravi, K., Lee, C.-W., & Lee, S. (2022). Convolutional neural network (CNN) with metaheuristic optimization algorithms for landslide susceptibility mapping in Icheon, South Korea. *Journal of environmental management, 305*, 114367. https://doi.org/10.1016/j.jenvman.2021.114367

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

Iverson, R. M. (2000). Landslide triggering by rain infiltration. *Water Resources Research, 36*, 1897–1910. https://doi.org/10.1029/2000WR900090

Jaafari, A. (2018). LiDAR-supported prediction of slope failures using an integrated ensemble weights-of-evidence and analytical hierarchy process. *Environment and Earth Science, 77*, 42. https://doi.org/10.1007/s12665-017-7207-3

Jaafari, A., Najafi, A., Pourghasemi, H. R., Rezaeian, J., & Sattarian, A. (2014). GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. *International Journal of Environmental Science and Technology, 11*, 909–926. https://doi.org/10.1007/s13762-013-0464-0

Jaafari, A., Najafi, A., Rezaeian, J., Sattarian, A., & Ghajar, I. (2015). Planning road networks in landslide-prone areas: A case study from the northern forests of Iran. *Land Use Policy, 47*, 198–208. https://doi.org/10.1016/j.landusepol.2015.04.010

Kadirhodjaev, A., Rezaie, F., Lee, M.-J., & Lee, S. (2020). Landslide susceptibility assessment using an optimized group method of data handling model. *ISPRS Int J Geoinf, 9*, 566. https://doi.org/10.3390/ijgi9100566

Kalantar, B., Ueda, N., Lay, U.S., Al-Najjar, H.A.H., Halin, A.A., 2019. Conditioning factors determination for landslide susceptibility mapping using support vector machine learning, In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. pp. 9626–9629. IEEE. https://doi.org/10.1109/IGARSS.2019.8898340

Kalantar, B., Ueda, N., Saeidi, V., Ahmadi, K., Halin, A. A., & Shabani, F. (2020). Landslide susceptibility mapping: machine and ensemble learning based on remote sensing big data. *Remote Sens (basel), 12*, 1737. https://doi.org/10.3390/rs12111737

Kanevski, M. (2009). *Machine learning for spatial environmental data theory, applications, and software*. EPFL Press. https://doi.org/10.1201/9781439808085

Kavzoglu, T., Kutlug Sahin, E., & Colkesen, I. (2015). Selecting optimal conditioning factors in shallow translational landslide susceptibility mapping using genetic algorithm. *Engineering Geology, 192*, 101–112. https://doi.org/10.1016/j.enggeo.2015.04.004

Khosravi, K., Pourghasemi, H. R., Chapi, K., & Bahri, M. (2016). Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: a comparison between Shannon's entropy, statistical index, and weighting factor models. *Environmental Monitoring and Assessment, 188*, 656. https://doi.org/10.1007/s10661-016-5665-9

Kjekstad, O., Highland, L., 2009. Economic and Social Impacts of Landslides. In *Landslides–Disaster Risk Reduction*. Springer Berlin Heidelberg, Berlin, Heidelberg, (pp. 573–587). https://doi.org/10.1007/978-3-540-69970-5_30

Kovács, I. P., Czigány, Sz., Dobre, B., Fábián, Sz. Á., Sobucki, M., Varga, G., & Bugya, T. (2019). A field survey–based method to characterise landslide development: a case study at the high bluff of the Danube, south-central Hungary. *Landslides, 16*, 1567–1581. https://doi.org/10.1007/s10346-019-01205-8

Kumar, R., & Anbalagan, R. (2016). Landslide susceptibility mapping using analytical hierarchy process (AHP) in Tehri reservoir rim region, Uttarakhand. *Journal of the Geological Society of India, 87*, 271–286. https://doi.org/10.1007/S12594-016-0395-8

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436–444. https://doi.org/10.1038/nature14539

Lee, C. F., Li, J., Xu, Z. W., & Dai, F. C. (2001). Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environmental Geology, 40*, 381–391. https://doi.org/10.1007/s002540000163

Lee, J.-G., Jun, S., Cho, Y.-W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: General overview. *Korean Journal of Radiology, 18*, 570. https://doi.org/10.3348/kjr.2017.18.4.570

Lee, S., Hong, S.-M., & Jung, H.-S. (2017). A support vector machine for landslide susceptibility mapping in Gangwon Province Korea. *Sustainability, 9*, 48. https://doi.org/10.3390/su9010048

Lee, S., & Pradhan, B. (2007). Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides, 4*, 33–41. https://doi.org/10.1007/s10346-006-0047-y

Li, Y., Chen, W., Rezaie, F., Rahmati, O., Davoudi Moghaddam, D., Tiefenbacher, J., Panahi, M., Lee, M.-J., Kulakowski, D., Tien Bui, D., & Lee, S. (2021). Debris flows modeling using geo-environmental factors: Developing hybridized deep-learning algorithms. *Geocarto International.* https://doi.org/10.1080/10106049.2021.1912194

Liang, Z., Wang, C., Duan, Z., Liu, H., Liu, X., & Ullah Jan Khan, K. (2021). A hybrid model consisting of supervised and unsupervised learning for landslide susceptibility mapping. *Remote Sens (basel), 13*, 1464. https://doi.org/10.3390/rs13081464

Mandal, K., Saha, S., & Mandal, S. (2021). Applying deep learning and benchmark machine learning algorithms for landslide susceptibility modelling in Rorachu river basin of Sikkim Himalaya. *India. Geoscience Frontiers, 12*, 101203. https://doi.org/10.1016/j.gsf.2021.101203

Merghadi, A., Yunus, A. P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D. T., Avtar, R., & Abderrahmane, B. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews, 207*, 103225. https://doi.org/10.1016/j.earscirev.2020.103225

Mondini, A. C., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., & Ardizzone, F. (2011). Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. *Remote Sensing of Environment, 115*, 1743–1757. https://doi.org/10.1016/j.rse.2011.03.006

Nasiri, V., Darvishsefat, A. A., Rafiee, R., Shirvany, A., & Hemat, M. A. (2019). Land use change modeling through an integrated Multi-Layer Perceptron Neural Network and Markov Chain analysis (case study: Arasbaran region, Iran). *Journal of Forestry Research, 30*, 943–957. https://doi.org/10.1007/s11676-018-0659-9

Nedbal, V., & Brom, J. (2018). Impact of highway construction on land surface energy balance and local climate derived from LANDSAT satellite data. *Science of the Total Environment, 633*, 658–667. https://doi.org/10.1016/j.scitotenv.2018.03.220

Nefeslioglu, H. A., Gokceoglu, C., & Sonmez, H. (2008). An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Engineering Geology, 97*, 171–191. https://doi.org/10.1016/j.enggeo.2008.01.004

Nguyen, V., Pham, B., Vu, B., Prakash, I., Jha, S., Shahabi, H., Shirzadi, A., Ba, D., Kumar, R., Chatterjee, J., & Tien Bui, D. (2019). Hybrid machine learning approaches for landslide susceptibility modeling. *Forests, 10*, 157. https://doi.org/10.3390/f10020157

Nohani, M., Sharafi, K., Pradhan, P., & Lee, M. (2019). Landslide susceptibility mapping using different GIS-based bivariate models. *Water (basel), 11*, 1402. https://doi.org/10.3390/w11071402

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*, 673–690. https://doi.org/10.1007/s11135-006-9018-6

Panahi, M., Gayen, A., Pourghasemi, H. R., Rezaie, F., & Lee, S. (2020). Spatial prediction of landslide susceptibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various metaheuristic algorithms. *Science of the Total Environment, 741*, 139937. https://doi.org/10.1016/j.scitotenv.2020.139937

Panahi, M., Rahmati, O., Rezaie, F., Lee, S., Mohammadi, F., & Conoscenti, C. (2022). Application of the group method of data handling (GMDH) approach for landslide susceptibility zonation using readily available spatial covariates. *Catena (Amst), 208*, 105779. https://doi.org/10.1016/j.catena.2021.105779

Pham, B. T., Prakash, I., Dou, J., Singh, S. K., Trinh, P. T., Tran, H. T., Le, T. M., van Phong, T., Khoi, D. K., Shirzadi, A., & Bui, D. T. (2020). A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. *Geocarto International, 35*, 1267–1292. https://doi.org/10.1080/10106049.2018.1559885

Pourghasemi, H. R., Kornejady, A., Kerle, N., & Shabani, F. (2020). Investigating the effects of different landslide positioning techniques, landslide partitioning approaches, and presence-absence balances on landslide susceptibility mapping. *Catena (Amst), 187*, 104364. https://doi.org/10.1016/j.catena.2019.104364

Pourghasemi, H. R., & Rahmati, O. (2018). Prediction of the landslide susceptibility: Which algorithm, which precision? *Catena (amst), 162*, 177–192. https://doi.org/10.1016/j.catena.2017.11.022

Reneau, S. L. (2000). Stream incision and terrace development in Frijoles Canyon, Bandelier National Monument, New Mexico, and the influence of lithology and climate. *Geomorphology, 32*, 171–193. https://doi.org/10.1016/S0169-555X(99)00094-X

Roy, J., & Saha, S. (2019). Landslide susceptibility mapping using knowledge driven statistical models in Darjeeling District, West Bengal India. *Geoenviron Disaster, 6*, 11. https://doi.org/10.1186/s40677-019-0126-8

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks, 61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Sestraş, P., Bilaşco, Ş, Roşca, S., Naş, S., Bondrea, M., Gâlgău, R., Vereş, I., Sălăgean, T., Spalević, V., & Cîmpeanu, S. (2019). Landslides susceptibility assessment based on GIS statistical bivariate analysis in the hills surrounding a metropolitan area. *Sustainability, 11*, 1362. https://doi.org/10.3390/su11051362

Sevgen, K., & Nefeslioglu, G. (2019). A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression ANN and random forest. *Sensors, 19*, 3940. https://doi.org/10.3390/s19183940

Shahabi, H., Hashim, M., Ahmad, B., & bin. (2015). Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central Zab basin Iran. *Environmental Earth Sciences, 73*, 8647–8668. https://doi.org/10.1007/s12665-015-4028-0

Sharma, L. P., Patel, N., Debnath, P., & Ghose, M. K. (2012). Assessing landslide vulnerability from soil characteristics—a GIS-based analysis. *Arabian J Geosci, 5*, 789–796. https://doi.org/10.1007/s12517-010-0272-5

Shirzadi, A., Chapi, K., Shahabi, H., Solaimani, K., Kavian, A., Ahmad, B., & bin. (2017). Rock fall susceptibility assessment along a mountainous road: an evaluation of bivariate statistic, analytical hierarchy process and frequency ratio. *Environment and Earth Science, 76*, 152. https://doi.org/10.1007/s12665-017-6471-6

Shirzadi, A., Solaimani, K., Roshan, M. H., Kavian, A., Chapi, K., Shahabi, H., Keesstra, S., Ahmadbin, B., & Bui, D. T. (2019). Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution. *Catena (amst), 178*, 172–188. https://doi.org/10.1016/j.catena.2019.03.017

Sur, U., Singh, P., Meena, S. R., & Singh, T. N. (2022). Predicting landslides susceptible zones in the Lesser Himalayas by ensemble of per pixel and object-based models. *Remote Sens (basel), 14*, 1953. https://doi.org/10.3390/rs14081953

Sur, U., Singh, P., Rai, P. K., & Thakur, J. K. (2021). Landslide probability mapping by considering fuzzy numerical risk factor (FNRF) and landscape change for road corridor of Uttarakhand, India. *Environment, Development and Sustainability, 23*, 13526–13554. https://doi.org/10.1007/s10668-021-01226-1

Tehrany, M. S., Jones, S., & Shabani, F. (2019). Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. *Catena (amst), 175*, 174–192. https://doi.org/10.1016/j.catena.2018.12.011

Tehrany, M. S., Pradhan, B., Mansor, S., & Ahmad, N. (2015). Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. *Catena (amst), 125*, 91–101. https://doi.org/10.1016/j.catena.2014.10.017

Thi Ngo, P. T., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Kariminejad, N., Cerda, A., & Lee, S. (2021). Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geoscience Frontiers, 12*, 505–519. https://doi.org/10.1016/j.gsf.2020.06.013

Tien Bui, D., Tuan, T. A., Klempe, H., Pradhan, B., & Revhaug, I. (2016). 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, 13*, 361–378. https://doi.org/10.1007/s10346-015-0557-6

Umar, Z., Pradhan, B., Ahmad, A., Jebur, M. N., & Tehrany, M. S. (2014). Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. *Catena (amst), 118*, 124–135. https://doi.org/10.1016/j.catena.2014.02.005

van Dao, D., Jaafari, A., Bayat, M., Mafi-Gholami, D., Qi, C., Moayedi, H., van Phong, T., Ly, H. B., Le, T. T., Trinh, P. T., Luu, C., Quoc, N. K., Thanh, B. N., & Pham, B. T. (2020). A spatially explicit deep learning neural network model for the prediction of landslide susceptibility. *Catena (Amst), 188*, 104451. https://doi.org/10.1016/j.catena.2019.104451

Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer. https://doi.org/10.1007/978-1-4757-2440-0

Xiao, L., Zhang, Y., & Peng, G. (2018). Landslide susceptibility assessment using integrated deep learning algorithm along the China-Nepal Highway. *Sensors, 18*, 4436. https://doi.org/10.3390/s18124436

Xu, C., Dai, F., Xu, X., & Lee, Y. H. (2012). GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. *Geomorphology, 145–146*, 70–80. https://doi.org/10.1016/j.geomorph.2011.12.040

Yao, J., Qin, S., Qiao, S., Che, W., Chen, Y., Su, G., & Miao, Q. (2020). Assessment of landslide susceptibility combining deep learning with semi-supervised learning in Jiaohe County, Jilin Province China. *Applied Sciences, 10*, 5640. https://doi.org/10.3390/app10165640

Yao, X., Tham, L. G., & Dai, F. C. (2008). Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of Hong Kong, China. *Geomorphology, 101*, 572–582. https://doi.org/10.1016/j.geomorph.2008.02.011

Yilmaz, I., Ercanoglu, M., 2019. Landslide inventory, sampling and effect of sampling strategies on landslide susceptibility/hazard modelling at a glance. In *Advances in Natural and Technological Hazards Research*. (pp. 205–224). https://doi.org/10.1007/978-3-319-73383-8_9

Yilmaz, I. (2009). A case study from Koyulhisar (Sivas-Turkey) for landslide susceptibility mapping by artificial neural networks. *Bulletin of Engineering Geology and the Environment, 68*, 297–306. https://doi.org/10.1007/s10064-009-0185-2

Authors and Affiliations

Armin Ghayur Sadigh[1] · Ali Asghar Alesheikh[1] · Sayed M. Bateni[2] · Changhyun Jun[3] · Saro Lee[4,5] · Jeffrey R. Nielson[6] · Mahdi Panahi[7,8] · Fatemeh Rezaie[4,5]

Armin Ghayur Sadigh
a.ghayursadigh@email.kntu.ac.ir

Ali Asghar Alesheikh
alesheikh@kntu.ac.ir

Sayed M. Bateni
smbateni@hawaii.edu

Saro Lee
leesaro@kigam.re.kr

Jeffrey R. Nielson
jeffrey.nielson@wsu.edu

Mahdi Panahi
panahi2012@yahoo.com

[1] Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran, Iran

[2] Department of Civil and Environmental Engineering and Water Resources Research Center, University of Hawaii at Manoa, Honolulu, HI, USA

[3] Department of Civil and Environmental Engineering, College of Engineering, Chung-Ang University, Seoul 06974, Republic of Korea

[4] Geoscience Data Center, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124, Gwahak-ro, Yuseong-gu, Daejeon 34132, Republic of Korea

[5] Department of Geophysical Exploration, Korea University of Science and Technology, 217, Gajeong-ro, Yuseong-gu, Daejeon 34113, Republic of Korea

[6] School of the Environment, Washington State University, Vancouver, WA, USA

[7]   Department of Physical Geography and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

[8]   Department of Computer Engineering, Chosun University, Gwangju 61452, Republic of Korea