The Egyptian Journal of Remote Sensing and Space Sciences 24 (2021) 845–855

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

The Egyptian Journal of Remote Sensing and Space Sciences

journal homepage: www.sciencedirect.com

Research Paper

# Improvement of landslide spatial modeling using machine learning methods and two Harris hawks and bat algorithms

Sina Paryani [a], Aminreza Neshat [a,*], Biswajeet Pradhan [b,c,d]

[a] Department of GIS/RS, Faculty of Natural Resources and Environment, Science and Research Branch, Islamic Azad University, Tehran, Iran
[b] The Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, New South Wales, Australia
[c] Center of Excellence for Climate Change Research, King Abdulaziz University, P. O. Box 80234, Jeddah 21589, Saudi Arabia
[d] Earth Observation Center, Institute of Climate Change, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia

ARTICLE INFO

*Article history:*
Received 26 March 2021
Revised 18 August 2021
Accepted 22 August 2021
Available online 4 September 2021

*Keywords:*
Landslide susceptibility
ANFIS
SVR
Harris hawks optimization
Bat algorithm
GIS

ABSTRACT

Landslide is a natural phenomenon that can turn into a natural disaster. The main goal of this research was to spatial prediction of a high-risk region located in the Zagros mountains, Iran, using hybrid machine learning and metaheuristic algorithms, namely the adaptive neuro-fuzzy inference system (ANFIS), support vector regression (SVR), the Harris hawks optimisation (HHO), and the bat algorithm (BA). The landslide occurrences were first divided into training and testing datasets with a 70/30 ratio. Fourteen landslide-related factors were considered, and the stepwise weight assessment ratio analysis (SWARA) were employed to determine the correlation between landslides and factors. After that, the hybrid models of ANFIS-HHO, ANFIS-BA, SVR-HHO and SVR-BA were applied to generate landslide susceptibility maps (LSMs). Finally, in order to validation and comparison of the applied models, two indexes, namely mean square error (MSE) and area under the ROC curve (AUROC), were used. According to the validation results, the AUROC values for the ANFIS-HHO, ANFIS-BA, SVR-HHO and SVR-BA were 0.849, 0.82, 0.895, and 0.865, respectively. The SVR-HHO showed the highest accuracy, with AUROC of 0.895 and lowest MSE of 0.147, and ANFIS-BA showed the least accuracy with an AUROC value of 0.82 and MSE value of 0.218. Based on the results, although four hybrid models with more than 80% accuracy can generate very good results, the SVR is superior to the ANFIS model, whereas the HHO algorithm outperformed the bat algorithm. The map generated in this study can be employed by land use planners in more efficient management.



## 1. Introduction

Gravity causes downward movement of the materials of a slope (including rock, soil, and natural materials), a phenomenon which is called landslide (Highland and Bobrowsky, 2008). According to the estimations, this geological phenomenon accounts for at least 17% of casualties of natural disasters worldwide (Pourghasemi et al., 2012), resulting billions of dollars of financial losses (Yilmaz, 2009); therefore, landslides are a globally serious issue.

The development of remote sensing and geographic information systems allow spatial modelling of many natural disasters such as landslides (Arabameri et al., 2020). Considering the lack of standard and specific procedures for landslide zonation (Polykretis et al., 2015), different approaches have been employed in the literature for this purpose. For instance, the statistical/probabilistic methods (Li et al., 2017; Liu and Duan, 2018; Kadavi et al., 2019) have been widely used. Using these methods has disadvantages, including (1) a large amount of data to run and (2) linear nature (Razavi Termeh et al., 2018). In some cases, these models have been combined to increase the output accuracy (Fan et al., 2017). In addition, recently, there has been a great interest in machine learning algorithms such as artificial neural networks (ANNs) (Aditian et al., 2018; Bragagnolo et al., 2020), adaptive neuro-fazzy inference system (ANFIS) (Chen et al., 2019; Paryani et al., 2020), Support vector machine (SVM) (Lee et al., 2017; Zhang et al., 2019), and random forest (RF) (Cao et al., 2019) due to their capability to solve nonlinear problems through learning. Although the mentioned models have been implemented successfully, many scholars have demonstrated that using ensemble models gives more accurate results (Xi et al., 2019; Panahi et al., 2020).

Peer review under responsibility of National Authority for Remote Sensing and Space Sciences.

* Corresponding author.
*E-mail address:* neshat.aminreza@srbiau.ac.ir (A. Neshat).

https://doi.org/10.1016/j.ejrs.2021.08.006
1110-9823/© 2021 National Authority for Remote Sensing and Space Sciences. Production and hosting by Elsevier B.V.

For example, Xi et al. (2019) used the ANN method and the hybrid ANN-PSO (particle swarm optimization) model. Despite the good performance of the ANN in calculating the LSI, the hybrid ANN-PSO model improved the performance and convergence in both training and validation phases. In another study, Balogun et al. (2021) used the SVR and three metaheuristic algorithms for spatial modeling of landslides. They reported that ensemble models provide more accurate results than stand-alone applications of SVR. However, due to the complexity of the landslide problem and the fact that each region has its own unique characteristics, using alternative methods is of great importance (Bui et al., 2019). The review of recent studies shows that although machine learning models have been widely used, their combination with meta-heuristic algorithms has received less attention. Therefore, the main purpose of this research is to fill this study gap using a new combination of ANFIS and SVR machine learning models with Harris hawk optimizer and compare them with ANFIS-BA and SVR-BA.

Iran has always been prone to landslide. According to Nadim et al. (2006), both the Alborz and Zagros Mountain Ranges fall within the moderate or high susceptible to landslides, but the Zagros Range is more affected by landslide patterns than the Alborz Range (Aghdam et al., 2017). The study region is located in the Zagros Mountain Range in Lorestan Province, Iran. Therefore, LSM can be employed as a useful tool for more efficient planning and management (Piacentini et al., 2015).

In this paper, to spatially predict the landslides, a combination of ANFIS and SVR machine learning methods with HHO and BA algorithms was used. The difference between the present study and previous research is in using two new hybrid models of SVR-HHO and ANFIS-HHO as well as their comparison with SVR-BA and ANFIS-BA. The SWARA model was first used to determine the correlation between dependent and independent variables. After calculating the LSIs, the final maps were generated and compared to select the most accurate model.

## 2. Studied region

The study region is located in the Middle Zagros Mountain Range in the east of Lorestan Province (Fig. 1). With an area of 3553 $km^2$, this region extends from the latitude $48^{o}\ 22'\ 20''$ to $49^{o}\ 29'\ 30''$ and the longitude $32^{o}\ 53'\ 27''$ to $33^{o}\ 28'\ 34''$. Its lowest and highest elevations are 486 and 4032 m, respectively. This region is affected by Mediterranean systems, and its annual rainfall varies from 263 mm to 600 mm (http://www.lorestanmet.ir/index.php/fa/). This region is affected by landslides every year due to its complex geology and climatic conditions. In addition, in recent years, with the increase of settlements and land use change without considering engineering standards, we have seen an increase in the risk of the occurrence of this natural phenomenon. Therefore, it is essential to generate LSM for risk management as well as better decision making for development of human activities.

## 3. Spatial database

### 3.1. Landslide inventory data

In the first step, the distribution of landslides was obtained from the Forest, Range, and Watershed Management Organization of Iran. The Google Earth Pro and field observations were then used to check and update the data (Fig. 2). According to Fig. 1, 70% of 255 detected landslides were used to train models. The remaining 30% points were used for the validation. In this regard, the same number of training and testing data were generated in the landslide-free areas using the create random points tool in the ArcMap 10.5 software. Most landslides occurred are of the slides (including transitional and rotational) with a number of rock falls.

### 3.2. Thematic layers preparation

According to the literature review (Aghdam et al., 2017; Chen et al., 2019) and available data, 14 conditioning factors were considered (Fig. 3). Table A.2 shows the source of data used. In the first step, ASTER global images (with 30 m resolution) were used to create the digital elevation model (DEM). The DEM of the study area was used to extract the slope angle, slope aspect, plan curvature, profile curvature, and topographic wetness index (TWI) (Fig. 3a–e and m). According to Fig. 4 the lithology of the studied region includes 17 classes. The land use map including agriculture, dry-farming, low forest, garden, poor range, range, and urban classes was obtained and then converted to raster format (Fig. 3f). Three GIS layers of distance to roads, distance to rivers, and distance to faults were prepared and classified into 9 classes using Euclidian distance (Fig. 3g, i, k). In the following, road and river density were

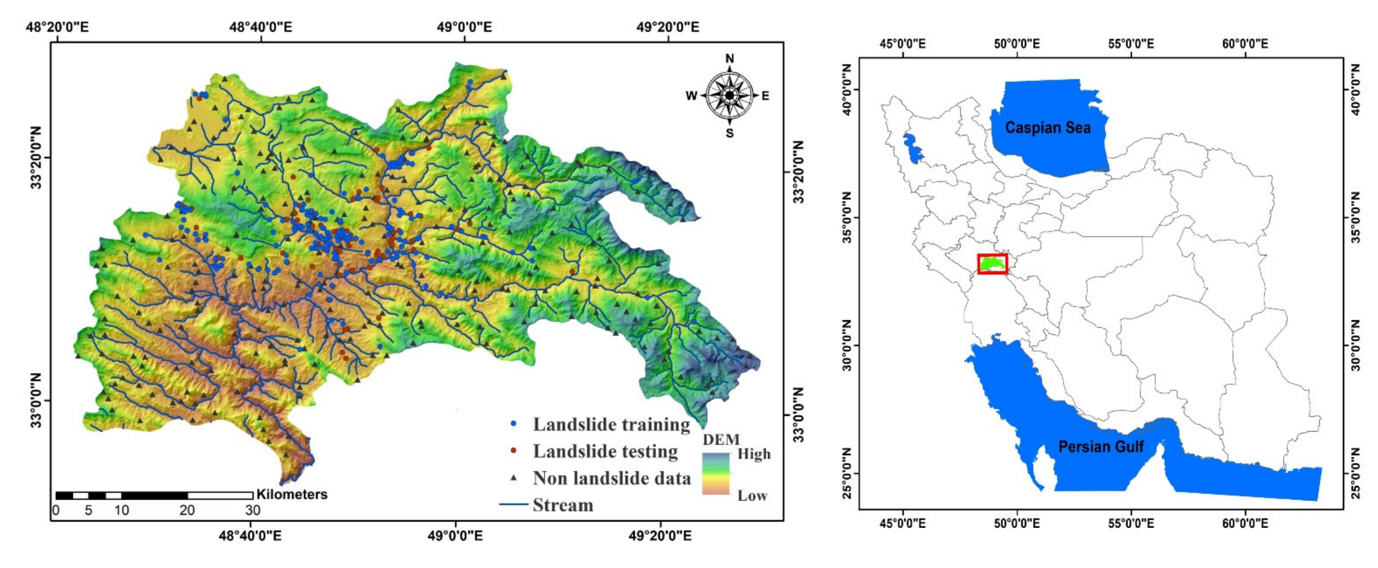


**Fig. 1.** Study area and inventory map.

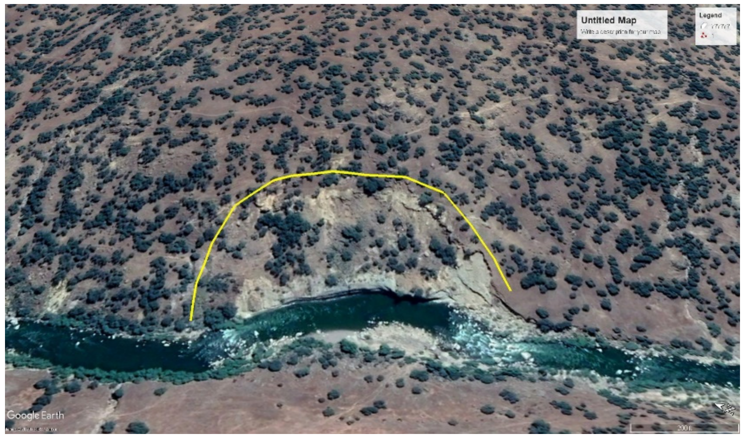


**Fig. 2.** Landslide locations using Google earth pro.

constructed using line density method and classified into 7 classes (Fig. 3h, j). The dataset obtained from the weather stations and inverse distance weighted (IDW) technique was also used to generate rainfall map (Fig. 3l).

Storage, analysis, and generation of layers were performed in ArcGIS 10.5 platform. Furthermore, in the current study, manual and natural break approaches were used to classify conditioning factors using literature review and expert's opinions. Fig. 5 illustrates all steps adopted in this research.

## 4. Methodology

### 4.1. Stepwise weight assessment ratio analysis

Stepwise weight assessment ratio analysis (SWARA) method is one of the multiple-criteria decision-making (MCDM) techniques designed to weight the study indices. Proposed by Keršuliene et al. (2010), in this method, expert opinions play a key role in selecting the significance of every criterion. It is implemented through the following steps (Torkashvand, 2020; Keršuliene et al., 2010):

- The first step is to identify the research criteria
- The second step is to rank the criteria based on their significance. A criterion is represented by the symbol $S_j$, therefore:

$$S_j = \frac{\sum_i^n A_i}{n} \tag{1}$$

where the subscript $j$ shows the criterion number, $n$ is the number of experts, and $A_i$ refers to the ranks recommended by experts for every criterion. The next step is to determine $K_j$ and $Q_j$.

$K_j$, which is a function of the relative importance of each criterion, is calculated as follows:

$$K_j = S_j + 1 \tag{2}$$

$Q_j$, which is the initial weight, is calculated as follows:

$$Q_j = \frac{X_{j-1}}{K_j} \tag{3}$$

where $X_j$ represents the weight of the class j.

- The last step is to determine the final normalized weight obtained from the following equation:

$$W_j = \frac{Q_j}{\sum_{j=1}^m Q_j} \tag{4}$$

where $j$ shows the criterion number and $m$ represents the number of criteria.

### 4.2. Adaptive Neuro-Fuzzy inference system

Integrating a fuzzy inference system with an artificial neural network (ANN), Jang (1993) proposed a powerful structure to solve nonlinear problems. Generally, the ANFIS consists of five layers, and its structure includes the interconnected nodes directly linked to each other. Considered a processing unit, every node has a function with tuning and constant parameters. Sugeno's system outperforms the other systems in terms of computational performance (Tien Bui et al., 2012). Therefore, the Sugeno's system was used in this study. According to Sugeno's rules:

$$\begin{aligned} &\text{Rule1} \;:\; \text{if } x \text{ is } A_1 \text{ and y is } B_1, \text{ then } f_1 = p_1x + q_1y + r_1 \\ &\text{Rule2} \;:\; \text{if } x \textit{ is } A_2 \textit{ and } \text{y is } B_2, \textit{ then } f_2 = p_2x + q_2y + r_2 \end{aligned}$$

where $x$ and $y$ are the non-fuzzy inputs, $f$ is the output, $A_1$, $A_2$, $B_1$, and $B_2$ are the fuzzy membership functions, and $p_i$, $q_i$, and $r_i$ $(\forall i = 1, 2)$ are the consequent parameters determined by the ANN (Jang, 1993). For further details, see the papers by Jang (1993) and Oh and Pradhan (2011).

### 4.3. Support vector regression

The support vector regression (SVR) is an adapted version of the Vapnik SVM developed to solve regression problems (Smola and Schölkopf, 2004). The SVR function can be either linear or nonlinear. The SVR model benefits from a series of linear functions $(f(x) = (w.x) + b)$ to make predictions, where $x$ and $w$ respectively show the input and weight vectors, and $b$ indicates the bias. In this process, a loss function is also employed to show the permissible deviation of the predicted values from the real values. Therefore, the following equations are utilized to minimize the optimization problem (Drucker et al., 1997):

$$Minimize: \frac{1}{2}||w||^2 + C\sum_{i=1}^{n}(\xi^* + \xi) \tag{6}$$

$$Subjectto: \begin{cases} y_i - (w.x_i + b) \le \varepsilon + \xi_i \\ (w.x_i + b) - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{7}$$

where $\epsilon$ shows the permissible error in a loss function, $\xi_i$, and $\xi_i^*$ are the slack variables, and $C$ is the penalty parameter. It is noteworthy that the SVR performance depends on the proper setting of some parameters, such as $C$, $\epsilon$, and the relevant kernel parameters. For more details, see Drucker et al. (1997).

### 4.4. Harris hawk optimization algorithm

Inspired by the collective behavior and pursuit style of Harris hawks, the population-based HHO algorithm was introduced by Heidari et al., in 2019. Generally, the mathematical simulation of this algorithm is based on three principles:

#### 4.4.1. Exploration phase

In the first step, they determine their own positions based on the positions of other family members and that of the prey (*e.g.* a rabbit) (q < 0.5). In the second step, the positions of hawks are totally random (q $\geq$ 0.5). From a mathematical point of view (Heidari et al., 2019):

$$X(iter+1)=\begin{cases} X_{rand}(iter)-r_1X_{rand}(iter)-2r_2X(iter) & if\ q\ \geq 0.5 \\ (X_{rabit}(iter)-X_m(iter))-r_3(LB+r_4(UB-LB)) & if\ q\ <0.5 \end{cases} \tag{8}$$

where iter represents the iteration, $X(iter+1)$ is a hawk's position vector in the next iteration, $X_{rabbit}(iter)$ and $X(iter)$ are a rabbit's and a hawk's positions, respectively, $X_{rand}(iter)$ represents the hawk selected randomly from the population, and LB and UB represent the lower and upper bounds of variables, respectively.

#### 4.4.2. Exploration-exploitation transmission

Obviously, the prey's energy decreases through escape. This energy can be modelled using the following equation:

$$E=2E_0\left(1-\frac{iter}{T}\right) \tag{9}$$

where $E_0$ is the initial prey's energy ranging from $-1$ to 1.

#### 4.4.3. Exploitation phase

In this case, Harris hawks employ soft besiege (|E|$\geq$0.5) and hard besiege (|E|<0.5) strategies to exhaust and easily hunt the prey. A variable showing the prey's chance of a successful escape (r) and the prey's energy (E) are employed to model this phase. For further details and equations, see the paper by Heidari et al. (2019).

### 4.5. Bat algorithm

Inspired by the collective behavior of bats in nature, Yang (2010) suggested a metaheuristic algorithm based on the swarm

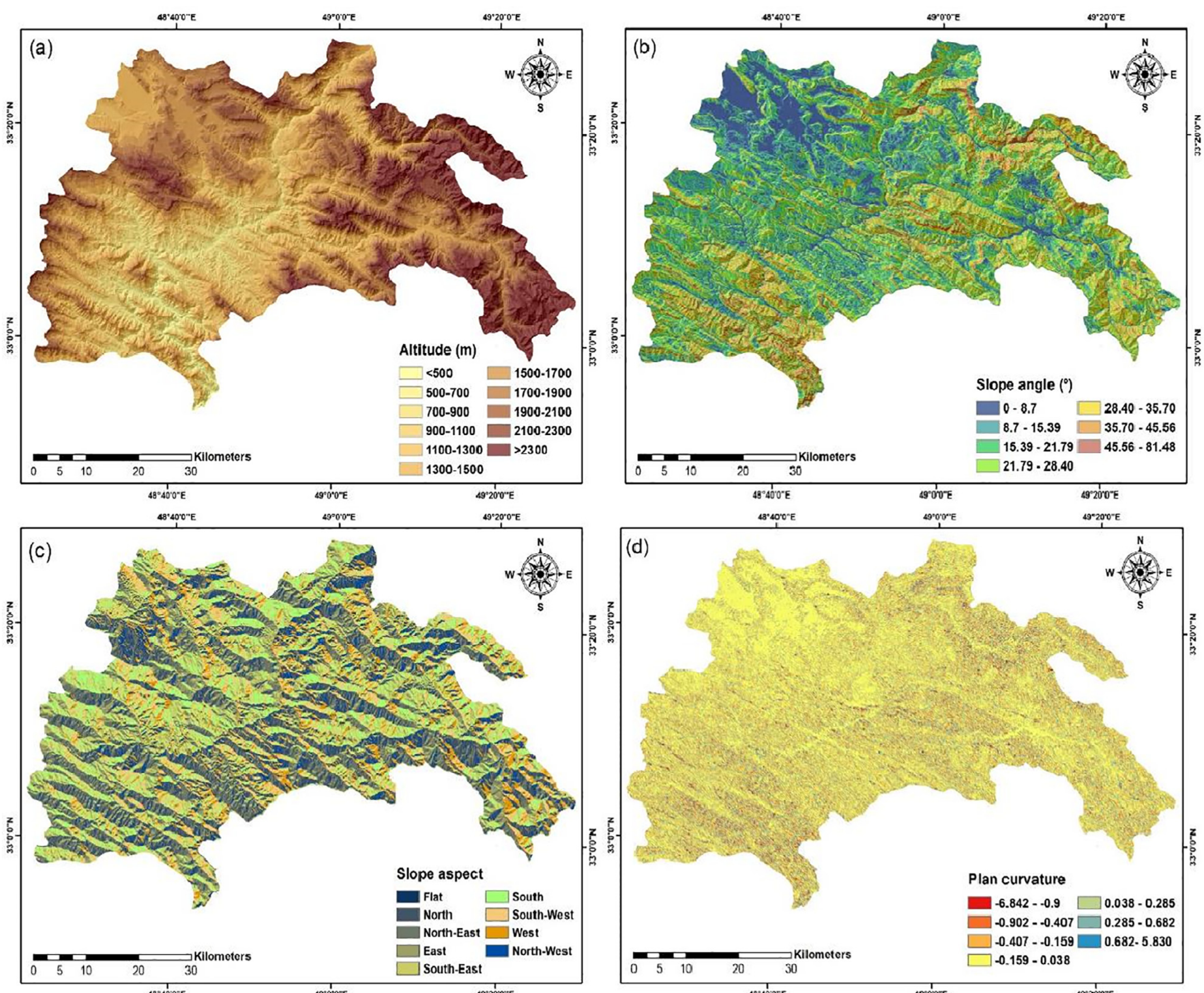


**Fig. 3.** Thematic layers: (a) Altitude, (b) Slope angle, (c) Slope aspect, (d) Plan curvature, (e) Profile curvature, (f) Land use, (g) Distance to roads, (h) road density, (i) Distance to rivers, (j) River density, (k) Distance to faults, (l) Rainfall, (m) TWI.

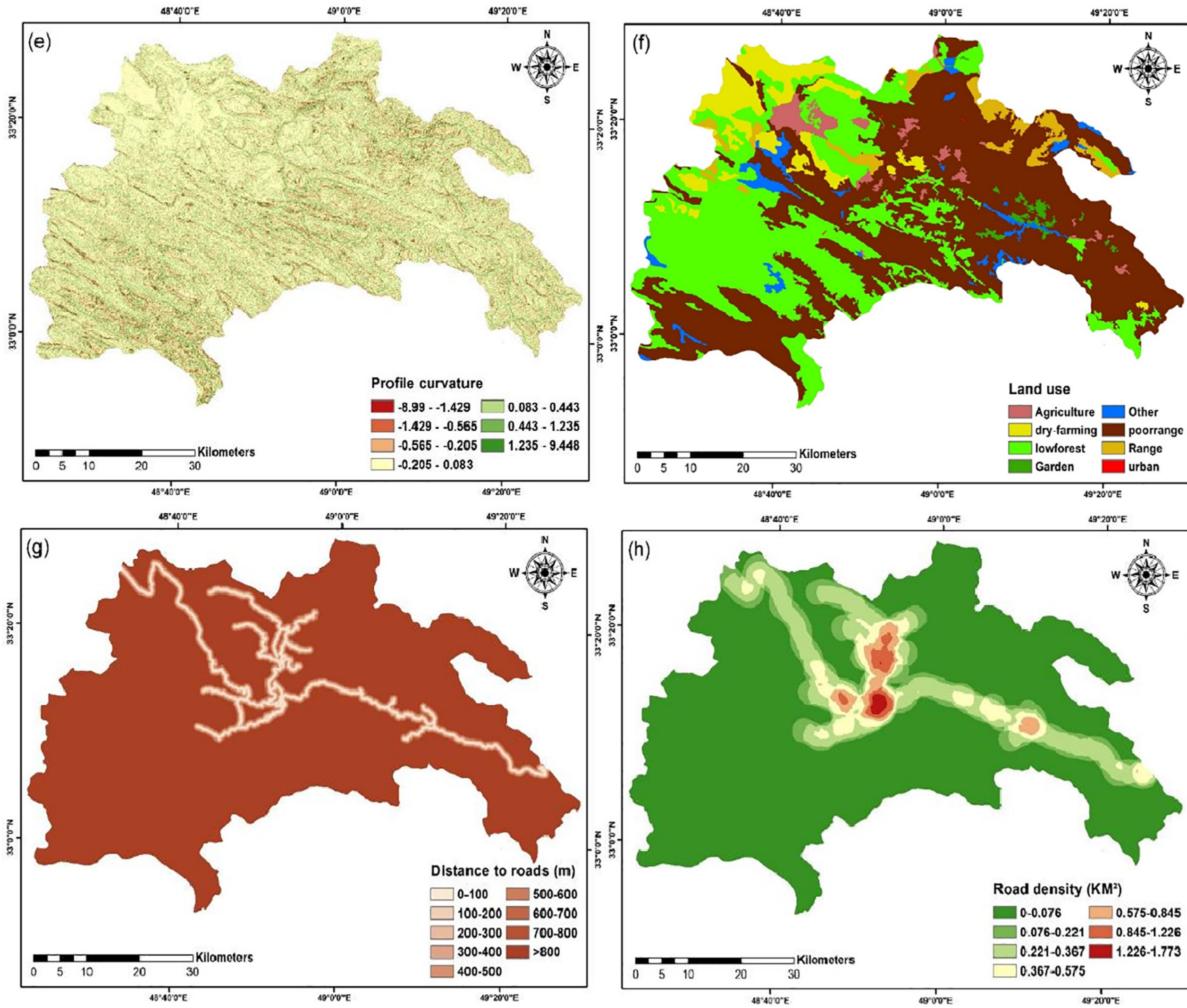


**Fig. 3** (*continued*)

intelligence called the bat algorithm. This strategy enables them to hunt their preys in the dark of the night. This behavior can be simulated mathematically to optimize nonlinear problems (Yang, 2010).

The velocity and position of each bat ($i$) in a $d$-dimensional search space in every iteration ($t$) are respectively represented by $V_i^t$ and $X_i^t$. These values are obtained from the following equations:

$$f_i = f_{min} + (f_{max} - f_{min})\beta \tag{10}$$

$$V_i^t = V_i^{t+1} + (X_i^t - X_*)f_i \tag{11}$$

$$X_i^t = X_i^{t-1} + V_i^t \tag{12}$$

where $\beta$ is a random vector from 0 to 1, and $X_*$ is the best global position (solution) selected by comparing the positions of $n$ bats. Depending on the problem size, $f_{min}$ and $f_{max}$ were considered as 0 and 100, respectively. After each iteration, a new position (solution) is generated for every bat, as follows (Yang, 2010):

$$X_{new} = X_{old} + \epsilon A^t \tag{13}$$

where $\epsilon$ is a random value within the $[-1, 1]$ interval. Moreover, $A^t$ shows the mean amplitude of all bats at $t$. $A_{min}$ and $A_{max}$ can be assigned to the amplitude. This value is changed and updated through the following equations:

$$A_i^{t+1} = \alpha A_i^t \tag{14}$$

$$r_i^{t+1} = r_i^0[1 - e(-\gamma t)] \tag{15}$$

where $\alpha$ and $\gamma$ are constants ($0 < \alpha < 1$; $\gamma > 0$).

For more details, see Yang (2010).

### 4.6. Model and map assessment

#### 4.6.1. Statistical methods

The model error assessment is a major step in the training and testing phase when using machine learning methods and/or their hybrids (Kia et al., 2012). The assessment results indicate which model was trained better and yielded more accurate outputs. The MSE and RMSE indices were used in this study:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_i\right)^2 \tag{16}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_i\right)^2} \tag{17}$$

where n shows the total number of samples, Xi is the target values, and $\overline{X}_i$ shows the output values.

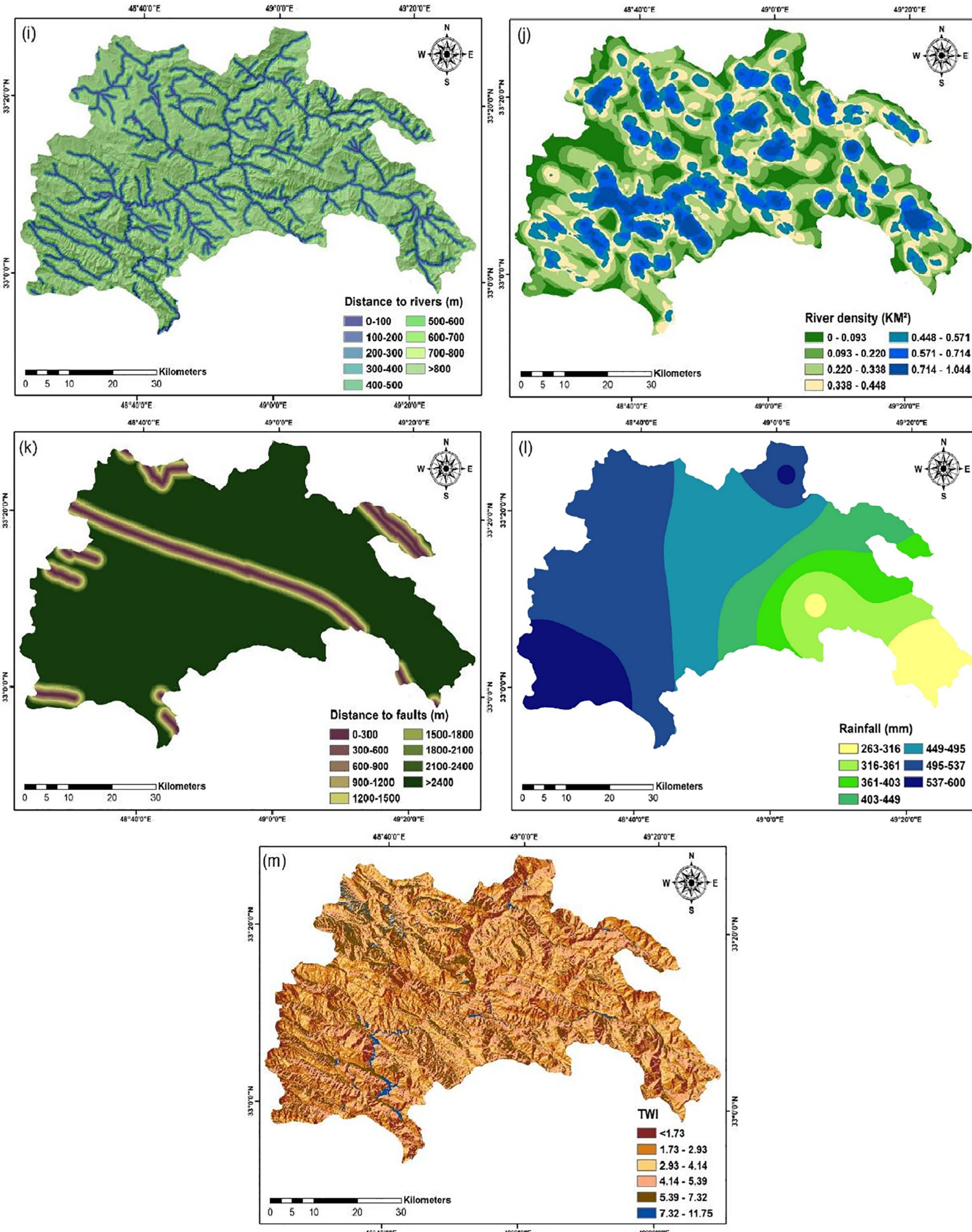


**Fig. 3** (*continued*)

#### 4.6.2. ROC curve

The ROC curve is a graphical representation of the balance between negative and positive error values (Razavi Termeh et al., 2018). This curve was employed to estimate the model accuracy using success and prediction rates. The training and testing datasets were respectively used for the success and prediction rates.

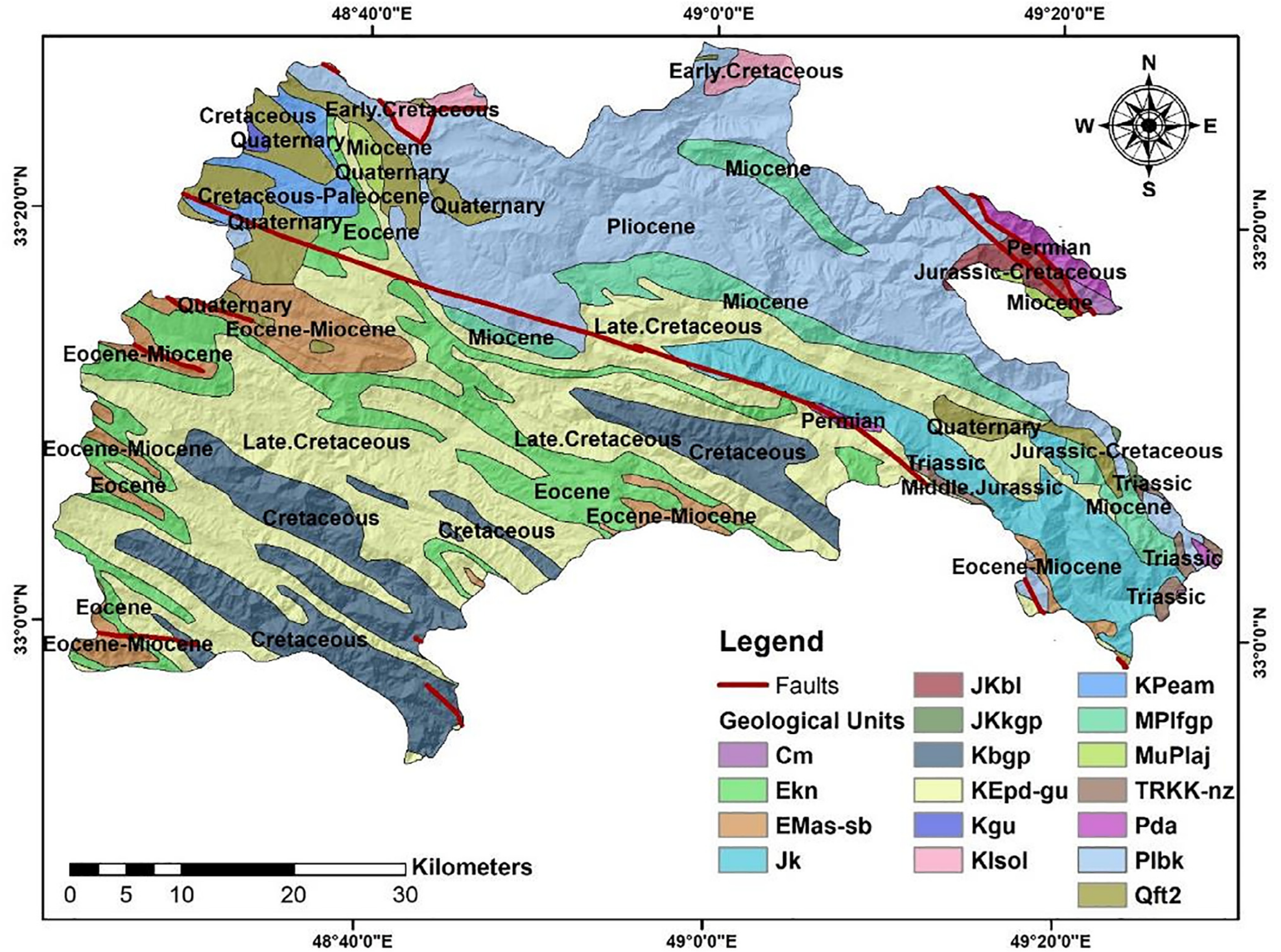


**Fig. 4.** Geological map.

## 5. Results

### 5.1. Application of SWARA to indicate correlation between landslides and conditioning factors

The SWARA model was used in this study to show a latent relationship between factors and landslides (Table A.3). The closer the values to 0.5, the more significant relationships; however, the closer the values to 0, the weaker the relationships. In the case of altitude, the class of 900–1100 showed the strongest correlation, with a weight of 0.337. The highest number of landslides has also occurred at the altitude of 900–1100 m. Regarding the slope degree, the class of 15.39$^{o}$-21.79$^{o}$ (with a SWARA value of 0.442) has the highest number of landslide in addition to having most coverage in the study area. Moreover, the analysis of different slope directions indicated that the east and southwest directions had the highest degrees of susceptibility. Regarding the plan and profile curvature, the classes (−0.407) – (−0.159) and 0.443–1.235 illustrated the highest landslide probability (0.349 and 0.373, respectively). The lithology of MPlfgp and EMas-sb units showed the highest SWARA rates of 0.328 and 0.2, respectively (Table A.3 of Appendix A). For land use, dry-farming class (with the weight of 0.330) was highly correlated with landslides in the study region. Investigation of land use factor showed that although the dry-farming class covers only 7% of the study area, it has the highest impact on landslide occurrence. In the case of distance to roads and rivers, the class of 0–100 m had the highest values (0.291 and 0.271, respectively). Concerning road density, the 0.575–0.845 class (0.265) had the highest importance, whereas the 0.714–1.044 class (0.383) showed the highest landslide potential for the density of the river. Related to the distance to faults, the 600–900 m class showed the highest probability of landslide occurrence (0.393). Regarding the rainfall factor, the ranges of 449–495 mm (0.204) illustrated the highest correlation. For TWI, the class of <1.73 has the highest probability of landslide occurrence (0.405).

### 5.2. Generate landslide susceptibility maps using ANFIS and SVR ensemble models

The metaheuristic HHO and BA algorithms were employed in this study to optimize the ANFIS and SVR methods. To this end, the MATLAB 2015b was applied. Furthermore, the training datasets (including 179 landslide/non-landslide points 1/0) and testing datasets (including 76 landslide/non-landslide points 1/0) were entered into the MATLAB environment in order to implement hybrid models. After the algorithms were executed, the outputs were entered into the ArcMap software to produce LSM.

Despite various methods for the LSM classification, there is no specific rule as to which method should be used (Ayalew et al., 2004). In this study, the natural breaks method was used. According to Fig. 6, the study area was classified into five classes namely, very low, low, moderate, high and very high. In addition, Fig. 7 illustrates the percentage of susceptible classes for each map. As

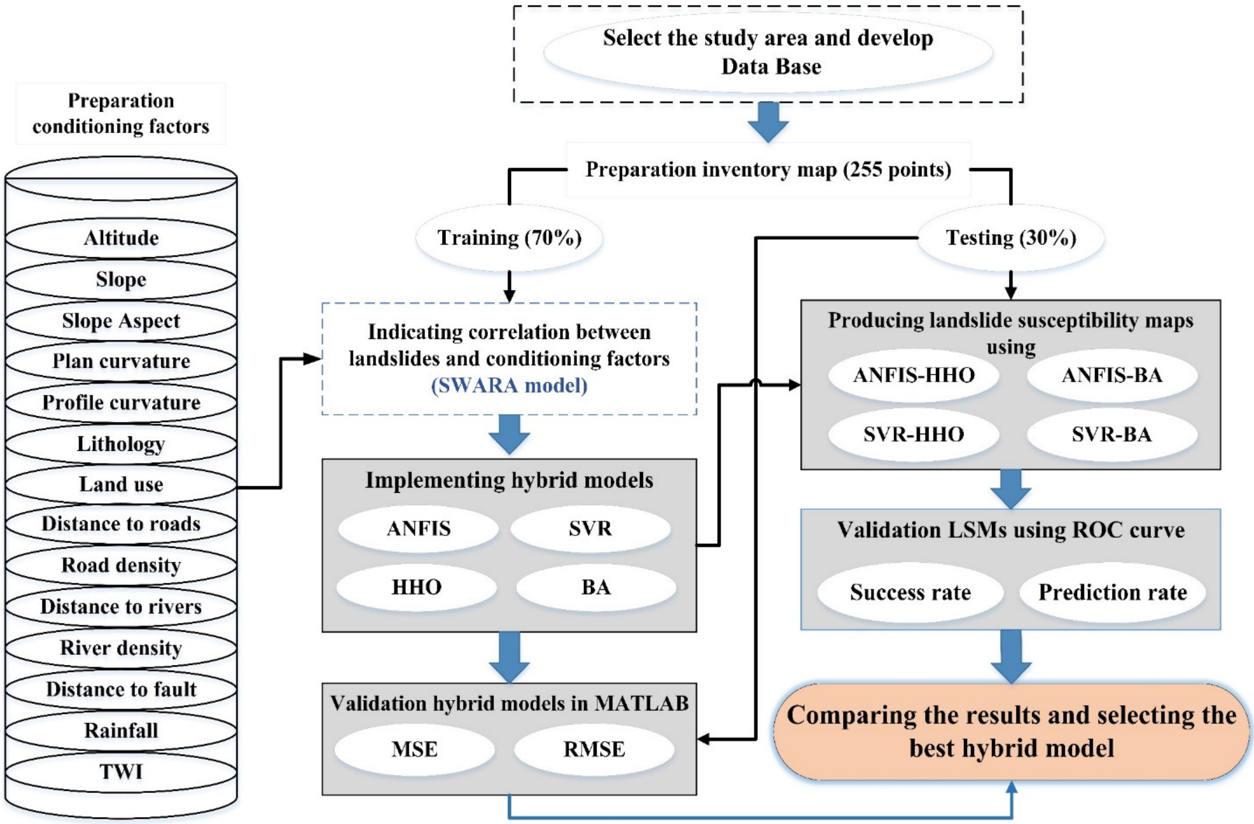


**Fig. 5.** Procedure of methodology.

clearly seen, the hybrid ANFIS-HHO, ANFIS-BA, SVR-HHO, and SVR-BA models showed the lowest coverage percentages of 3.06%, 2.97%, 8.35%, and 4.53%, respectively, in the very high class. To this end, the highest percentages of coverage were obtained for ANFIS-HHO in the low class and for three other models in very low class.

### 5.3. Validation and comparison of models

#### 5.3.1. Statistical index

According to Figs. A.1 and A.2 of Appendix A, the MSE values of the ANFIS-HHO, ANFIS-BA, SVR-HHO, and SVR-BA models in the testing phase were 0.178, 0.218, 0.147, and 0.158, respectively. Based on the obtained values, the lower error rate of the hybrid SVR-HHO model (0.147) leads to more accurate prediction. The SVR-BA model was ranked as second. Finally, the ANFIS-BA model showed the poor performance among all models.

#### 5.3.2. ROC curve analysis

Fig. 8 shows the values of AUROC curve for training and testing datasets. In line with the values of the success rate, the values of prediction rate for ANFIS-HHO, ANFIS-BA, SVR-HHO, and SVR-BA hybrid models were 0.849, 0.82, 0.895, and 0.865, respectively (Fig. 8b and Table 1). According to the comparison results, the SVR outperformed the ANFIS, whereas the HHO outperformed the BA (Fig. 9). Table 2 also shows the comparison between the MSE, RMSE and ROC values of the models.

## 6. Discussion

The hybrid ANFIS-HHO, ANFIS-BA, SVR-HHO, and SVR-BA models were employed for the spatial prediction of landslides in a region located in Lorestan Province, Iran. To this end, a comparison was made between the ANFIS and the SVR to estimate the performance of each machine learning method in combination with two metaheuristic algorithms, i.e. the HHO and the BA. Many scholars have acknowledged that using machine learning-metaheuristic ensemble models can improve the LSM's accuracy. Razavi Termeh et al. (2018), for example, used three hybrid ANFIS-ACO, ANFIS-GA, and ANFIS-PSO models to construct flood susceptibility map. They reported that, although all models applied have excellent results, ANFIS-PSO exhibited better performance. In line with the study conducted, we concluded that the ANFIS based models with more than 80% accuracy have very good performance. In another study, Panahi et al. (2020) used the SVR and ANFIS machine learning models and their combination with two Bee and GWO algorithms to spatially predict landslides. Their results indicated that, SVR ensemble model provides more accurate results compared to the ANFIS. Consistent with this study, we found that combination of SVR with metaheuristic algorithms provides better accuracy than ANFIS ensemble model. Balogun et al. (2021) integrated SVR model with three metaheuristic algorithms for modeling landslides in western Serbia. The model validation indicated that ensemble models applied yielded better results than only SVR. Our results also revealed that SVR-HHO and SVR-BA ensemble models with more than 80% accuracy provide very good performance.

In this study, the validation results indicated that the ANFIS-HHO, ANFIS-BA, SVR-HHO, and SVR-BA models had very good performance with AUCs above 80%. There are two important points regarding the results of this research. First, the SVR showed more accurate results than the ANFIS. In other words, the SVR model showed better adaptability in combination with meta-heuristic algorithms due to the advantages of this method, including 1) solv-

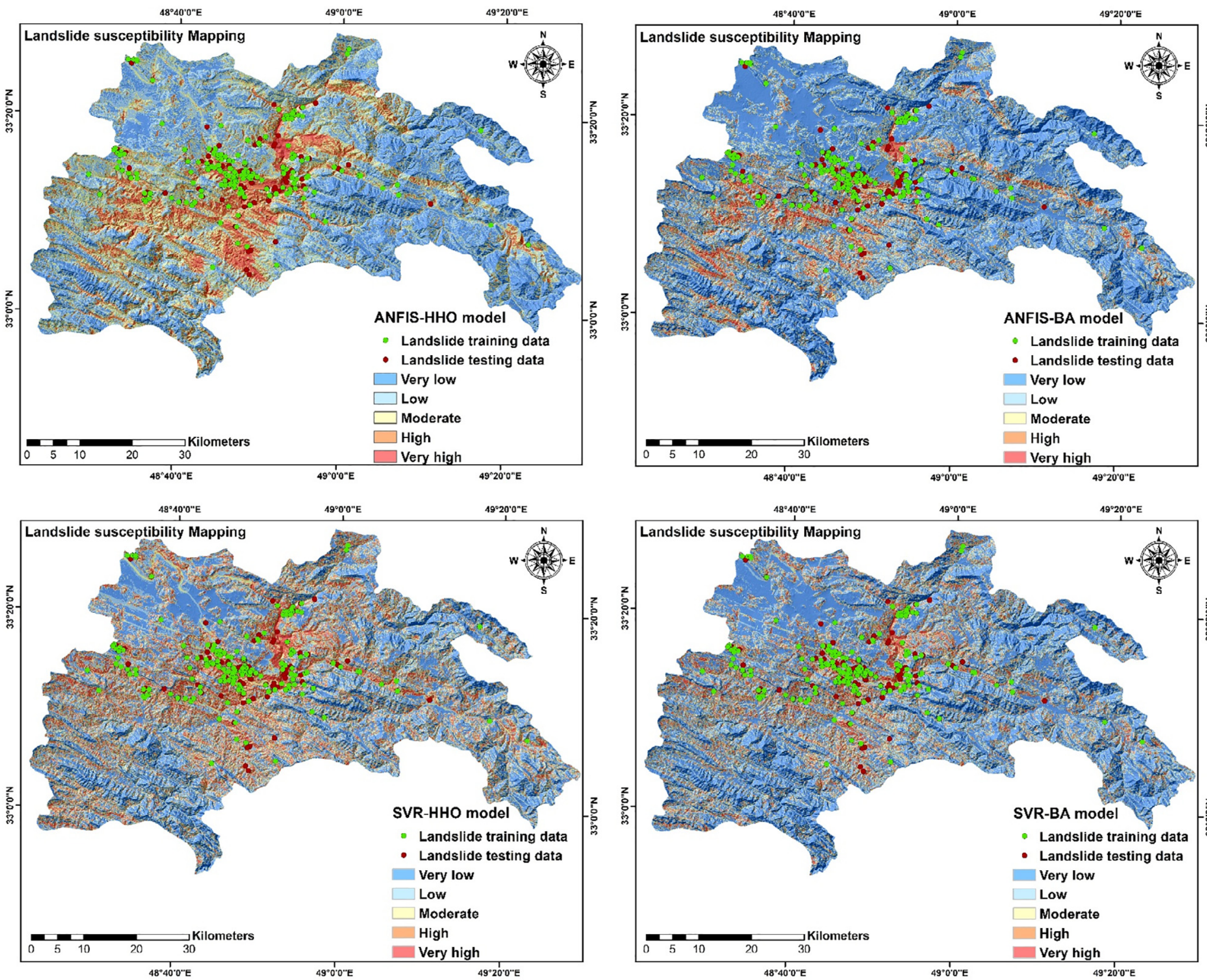


**Fig. 6.** Landslide susceptibility maps.

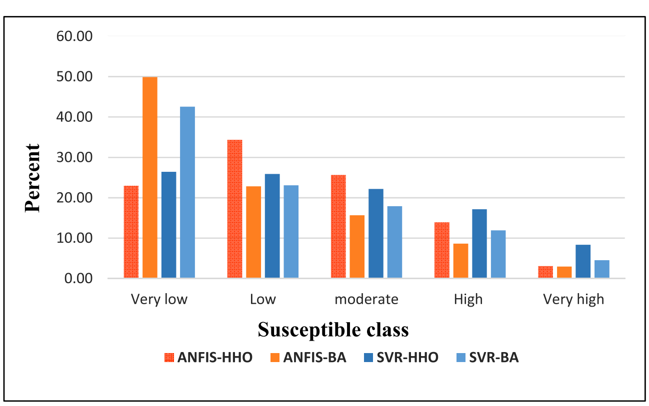


**Fig. 7.** Bar diagram showing the percentage of each class.

ing nonlinear and multi-dimensional problems effectively and 2) its flexibility in integration with other optimization models (Panahi et al., 2020). In addition, Balogun et al. (2021) stated that the SVR model has the ability to solve nonlinear problems with a small number of samples and high dimensions. Secondly, although both HHO and BA succeeded in optimization, the HHO was more accurate than the BA. Heidari et al. (2019) assessed the HHO performance through 29 unconstraint benchmark problems and showed that the HHO algorithm outperformed the other intelligent algorithms in finding more accurate solutions and not converting on the local optimum. This finding is backed by the superiority of the HHO algorithm to the BA in this study. It is worth noting that the weights of SWARA can show the correlation between landslides and conditioning factors well. The results obtained from the reviewed studies (Chen et al., 2019) confirm this claim.

The proposed hybrid models can also be used for the spatial modelling of other phenomena. It is also suggested that other intelligent algorithms be used in future studies in combination with machine learning methods.

## 7. Conclusion

Prevention has always been the main strategy in facing potential natural hazards such as landslides. In this study, a combination of SVR and ANFIS machine learning methods with HHO and BA metaheuristic algorithms was used to prepare a landslide susceptibility map of the study area. Moreover, three indices of MSE, RMSE and AUROC were also used to determine the accuracy and compare the performance of each model. The results revealed that the SVR-HHO model with the highest AUROC of 0.895 and lowest MSE of 0.147 exhibited better performance than other models. Comparison between hybrid models indicated that, the SVR method outperformed the ANFIS. Moreover, the HHO algorithm illustrated more

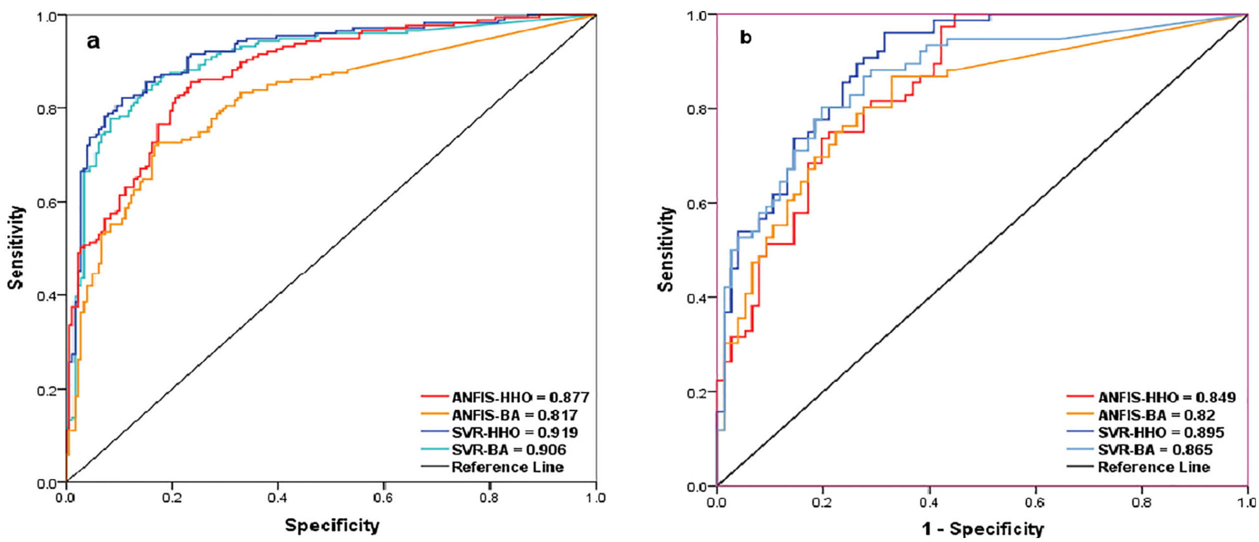


**Fig. 8.** ROC curve for success rate (a) and prediction rate (b).

**Table 1**
Details of AUROC for prediction rate.

| Test Result Variable(s) | Area | Std. Error | Asymptotic Sig. | Asymptotic 95% ConfidenceInterval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| ANFIS-HHO | 0.849 | 0.030 | 0.000 | 0.790 | 0.909 |
| ANFIS-BA | 0.820 | 0.034 | 0.000 | 0.753 | 0.888 |
| SVR-HHO | 0.895 | 0.025 | 0.000 | 0.846 | 0.944 |
| SVR-BA | 0.865 | 0.030 | 0.000 | 0.807 | 0.924 |


**Fig. 9.** Comparison between success and prediction rates.

**Table 2**
Comparison between MSE, RMSE and AUROC.

| Ensemble model | MSE | RMSE | Prediction rate | Success rate |
|---|---|---|---|---|
| ANFIS-HHO | 0.178 | 0.422 | 0.849 | 0.877 |
| ANFIS-BA | 0.218 | 0.467 | 0.820 | 0.817 |
| SVR-HHO | 0.147 | 0.384 | 0.895 | 0.919 |
| SVR-BA | 0.158 | 0.398 | 0.865 | 0.906 |

suitable than the BA in combination with the ANFIS and SVR methods. Based on the study's outcome, the proposed hybrid methods can act as an effective approach for the zonation of other landslide-prone regions and also help decision-makers to manage disaster better.

## Declaration of Competing 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.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejrs.2021.08.006.

## References


Aditian, A., Kubota, T., Shinohara, Y., 2018. Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. Geomorphology 318, 101–111.

Aghdam, I.N., Pradhan, B., Panahi, M., 2017. Landslide susceptibility assessment using a novel hybrid model of statistical bivariate methods (FR and WOE) and adaptive neuro-fuzzy inference system (ANFIS) at southern Zagros Mountains in Iran. Environ. Earth Sci. 76 (6), 237.

Arabameri, A., Pradhan, B., Rezaei, K., Lee, S., Sohrabi, M., 2019. An ensemble model for landslide susceptibility mapping in a forested area. Geocarto Int 35 (15), 1680–1705.

Ayalew, L., Yamagishi, H., Ugawa, N., 2004. Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. Landslides 1 (1), 73–81.

Balogun, A.-L., Rezaie, F., Pham, Q.B., Gigović, L., Drobnjak, S., Aina, Y.A., Panahi, M., Yekeen, S.T., Lee, S., 2021. Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression (SVR) with GWO, BAT and COA algorithms. Geosci. Front. 12 (3), 101104. https://doi.org/10.1016/j.gsf.2020.10.009.

Tien Bui, D., Pradhan, B., Lofman, O., Revhaug, I., Dick, O.B., 2012. Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. Comput. Geosci. 45, 199–211.

Bui, D.T., Hoang, N.D., Nguyen, H., Tran, X.L., 2019. Spatial prediction of shallow landslide using Bat algorithm optimized machine learning approach: A case study in Lang Son Province, Vietnam. Adv. Eng. Inf. 42, 100978.

Bragagnolo, L., da Silva, R.V., Grzybowski, J.M.V., 2020. Landslide susceptibility mapping with r. landslide: a free open-source GIS-integrated tool based on Artificial Neural Networks. Environ. Modell. Software 123, 104565. https://doi.org/10.1016/j.envsoft.2019.104565.

Cao, J., Zhang, Z., Wang, C., Liu, J., Zhang, L., 2019. Susceptibility assessment of landslides triggered by earthquakes in the Western Sichuan Plateau. Catena 175, 63–76.

Chen, W., Panahi, M., Tsangaratos, P., Shahabi, H., Ilia, I., Panahi, S., Li, S., Jaafari, A., Ahmad, B.B., 2019. Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility. Catena 172, 212–231.

Drucker, H., Burges, C.J., Kaufman, L., Smola, A.J., Vapnik, V., 1997. Support vector regression machines. In Advances in neural information processing systems pp. 155–161.

Fan, W., Wei, X.-S., Cao, Y.-b., Zheng, B., 2017. Landslide susceptibility assessment using the certainty factor and analytic hierarchy process. J. Mt. Sci.-Engl. 14 (5), 906–925.
Heidari, A.A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., Chen, H., 2019. Harris hawks optimization: Algorithm and applications. Future Gener. Comp. Sy. 97, 849–872.
Highland, L., Bobrowsky, P.T., 2008. The Landslide Handbook: A Guide to Understanding Landslides. US Geological Survey, Reston, p. 129.
Jang, J.S., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst., Man, Cybernetics 23 (3), 665–685.
Kadavi, P.R., Lee, C.W., Lee, S., 2019. Landslide-susceptibility mapping in Gangwon-do, South Korea, using logistic regression and decision tree models. Environ. Earth Sci. 78 (4), 116.
Keršuliene, V., Zavadskas, E.K., Turskis, Z., 2010. Selection of rational dispute resolution method by applying new step-wise weight assessment ratio analysis (SWARA). J. Business Econ. Manage. 11 (2), 243–258.
Kia, M.B., Pirasteh, S., Pradhan, B., Mahmud, A.R., Sulaiman, W.N.A., Moradi, A., 2012. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Environ. Earth Sci. 67 (1), 251–264.
Lee, S., Hong, S.M., Jung, H.S., 2017. A support vector machine for landslide susceptibility mapping in Gangwon Province, Korea. Sustainability 9 (1), 48.
Li, L., Lan, H., Guo, C., Zhang, Y., Li, Q., Wu, Y., 2017. A modified frequency ratio method for landslide susceptibility assessment. Landslides 14 (2), 727–741.
Liu, J., Duan, Z., 2018. Quantitative assessment of landslide susceptibility comparing statistical index, index of entropy, and weights of evidence in the Shangnan area, China. Entropy 20 (11), 868.
Oh, H.-J., Pradhan, B., 2011. Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. Comput. Geosci. 37 (9), 1264–1276.
Nadim, F., Kjekstad, O., Peduzzi, P., Herold, C., Jaedicke, C., 2006. Global landslide and avalanche hotspots. Landslides 3 (2), 159–173.
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. Sci. Total Environ. 741, 139937. https://doi.org/10.1016/j.scitotenv.2020.139937.
Paryani, S., Neshat, A., Javadi, S., Pradhan, B., 2020. Comparative performance of new hybrid ANFIS models in landslide susceptibility mapping. Nat. Hazards 103 (2), 1961–1988.
Piacentini, D., Devoto, S., Mantovani, M., Pasuto, A., Prampolini, M., Soldati, M., 2015. Landslide susceptibility modeling assisted by Persistent Scatterers Interferometry (PSI): an example from the northwestern coast of Malta. Nat. Hazards 78 (1), 681–697.
Polykretis, C., Ferentinou, M., Chalkias, C., 2015. A comparative study of landslide susceptibility mapping using landslide susceptibility index and artificial neural networks in the Krios River and Krathis River catchments (northern Peloponnesus, Greece). B Eng. Geol. Environ. 74 (1), 27–45.
Pourghasemi, H.R., Mohammady, M., Pradhan, B., 2012. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. Catena 97, 71–84.
Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Statistics Computi. 14 (3), 199–222.
Razavi Termeh, S.V., Kornejady, A., Pourghasemi, H.R., Keesstra, S., 2018. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. Sci. Total Environ. 615, 438–451.
Torkashvand, Maryam et al.Neshat, A., Javadi, S., & Yousefi, H., 2020. DRASTIC framework improvement using stepwise weight assessment ratio analysis (SWARA) and combination of genetic algorithm and entropy. Environ. Sci. Pollut. Res., 1–21 https://doi.org/10.1007/s11356-020-11406-7.
Xi, W., Li, G., Moayedi, H., Nguyen, H., 2019. A particle-based optimization of artificial neural network for earthquake-induced landslide assessment in Ludian county, China, Geomatics. Nat. Hazards Risk 10 (1), 1750–1771.
Yang, X.S., 2010. A new metaheuristic bat-inspired algorithm. In Nature inspired cooperative strategies for optimization (NICSO 2010) (pp. 65-74). Springer, Berlin, Heidelberg.
Yilmaz, I., 2009. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). Comput. Geosci. 35 (6), 1125–1138.
Zhang, T.-y., Han, L., Zhang, H., Zhao, Y.-H., Li, X.-a., Zhao, L., 2019. GIS-based landslide susceptibility mapping using hybrid integration approaches of fractal dimension with index of entropy and support vector machine. J. Mountain Sci. 16 (6), 1275–1288.