Aeolian Research 67–69 (2024) 100924

Contents lists available at ScienceDirect

Aeolian Research

journal homepage: www.elsevier.com/locate/aeolia

Full Length Article

An explainable integrated machine learning model for mapping soil erosion by wind and water in a catchment with three desiccated lakes

Hamid Gholami [a,*], Mehdi Jalali [a], Marzieh Rezaei [a], Aliakbar Mohamadifar [a], Yougui Song [b,*], Yue Li [b], Yanping Wang [b,c], Baicheng Niu [d,e,f], Ebrahim Omidvar [g], Dimitris G. Kaskaoutis [h]

[a] Department of Natural Resources Engineering, University of Hormozgan, Bandar-Abbas, Hormozgan, Iran
[b] State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 710061, China
[c] University of Chinese Academy of Sciences, Beijing 100049, China
[d] Academy of Plateau Science and Sustainability, Xining 810008, China
[e] College of Geography, Qinghai Normal University, Xining 810008, China
[f] Dunhuang Gobi Desert Research Station, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China
[g] Department of Watershed Management Engineering, University of Kashan, Kashan, Iran
[h] Department of Chemical Engineering, University of Western Macedonia, Kozani, 50100, Greece

ARTICLE INFO

Keywords:
Soil erosion
Integrated spatial map
Machine learning
Interpretation techniques
Local dust sources
Bakhtegan catchment

ABSTRACT

Soil erosion by water and wind is a critical challenge for sustainable management of catchments in drylands and accurate spatial information can help in mitigation of its destructive consequences. Here, seven machine learning (ML) models were applied to map simultaneously the water and wind erosions in the Bakhtegan catchment, south Iran, with three dried lakes in its southern part and three dams established in upstream parts of the lakes. The analysis identified 10 and 11 effective variables controlling water and wind erosions, among 20 and 17 potential variables, respectively, via the MARS feature selection algorithm. According to the most accurate ML models (artificial neural network for water erosion, and Cubist for wind erosion), an integrated model was developed to map soil erosion by water and wind simultaneously. Permutation feature importance (PFI) and Shapley additive exPlanation (SHAP) interpretation techniques were employed to explain the model outputs, revealing that 19.7 % of the total area belonged to high and very high susceptibility classes to soil erosion by water and wind. The PFI plot revealed that the slope and wind speed were the most influencing factors for water and wind erosion, respectively. According to SHAP decision plot, slope had the highest contribution on the predictive water erosion model's output, whereas vegetation cover exhibited the highest contribution on the predictive wind erosion model's output. Due to climate change and intensified drought during the recent years, as well as due to construction of dams upstream of the lakes, the southern part of the study area was converted to a source of sand and dust storms. The hotspots with severe water erosion are distributed all over the study area, rendering it quite vulnerable to adverse climate change projections.

1. Introduction

Soil erosion by wind and water is one of the critical processes of soil degradation in the catchment areas, while the 56 % and 28 % of the total global degraded land was affected by water and wind erosion, respectively (Oldeman, 1994). At the global scale, soil erosion by water is the most severe type of erosion, while wind erosion is a widespread event, especially in the drylands. Soil erosion has many detritus consequences and the most important erosion-induced soil degradation processes can be classified as physical (mitigation of soil depth, water infiltration), chemical (increasing salinization, acidification, and water pollution), and biological (increasing eutrophication and hypoxia) (Blanco-Canqui and Lal, 2008; Shirani et al., 2020). Traditionally, water erosion and wind erosion have been analyzed and modelled separately (Tanner et al., 2023), but nowadays due to climate change and intensifying anthropogenic interference, these two types of soil erosion are both active in catchment areas located in the drylands (Vaezi et al., 2019; Boroughani et al., 2021). Therefore, simultaneous spatial mapping of land susceptibility to soil erosion due to water and wind in the catchments can be helpful to mitigate their negative consequences and for

* Corresponding authors.
E-mail addresses: hgholami@hormozgan.ac.ir (H. Gholami), ygsong@loess.llqg.ac.cn (Y. Song).

https://doi.org/10.1016/j.aeolia.2024.100924
Received 21 December 2023; Received in revised form 6 April 2024; Accepted 14 April 2024
Available online 27 April 2024
1875-9637/© 2024 Elsevier B.V. All rights reserved.

sustainable management of catchments (Qaderi Nasab and Rahnama, 2019; Gholami et al., 2020).

Machine learning (ML), as a field of data mining, has been extensively applied to map soil erosion hazard (Mohammadifar et al., 2021; Liu et al., 2023; Lana et al., 2022; Gholami et al., 2020b; Bag et al., 2022; Gholami and Mohammadifar, 2022). Application of ML models for mapping soil erosion due to water and wind, simultaneously has not been reported yet, and this study addresses this gap for the first time. On the other hand, one of the lesser-known aspects associated with ML models is their interpretability. ML models are a type of black box models, and their interpretation can help us to better understand the model's outputs and to analyze the contribution and impact of the input variables (influencing factors) on the model's outputs (Gholami et al., 2023a). Interpretation can also help us to understand the cause of a decision (Miller, 2019). Overall, interpretation techniques can be divided into global model-agnostics methods and local model-agnostic methods. More recently, application of the interpretation techniques has been attracted the scientific interest to explain ML model's output, especially in the spatial modelling of earthquake, flood, gully erosion, soil salinity, wind erosion and dust emissions from the land surfaces (e.g., Jena et al., 2023a,b; Pradhan et al., 2023a,b; Gholami et al., 2023a, b).

The current study aimed to develop an explainable ML model for the simultaneous spatial modelling of soil erosion in an arid environment in south Iran. The study area (Bakhtegan watershed basin) is a climate-change hotspot and constitutes an environmental crisis due to enhanced atmospheric warming in the Middle East and the human interference (Kafilzadeh et al., 2007; Gharechaei et al., 2015; Mozafari et al., 2022). Therefore, the main goals of this study are summarized to (i) identify the effective variables on water erosion and wind erosion using MARS algorithm, (ii) apply tolerance coefficient (TC) and variance inflation factor (VIF) tests to explore multi-collinearity among all variables controlling water erosion and wind erosion, (iii) mapping soil erosion by water and wind simultaneously using an integrated ML model, and (iv) due to black box nature of ML models, permutation feature importance (PFI) and Shapley additive exPlanation (SHAP) interpretation techniques were applied for better understanding, analysis and explanation of the output of ML models. To the best of our knowledge, there is not any other study indicating the application of the interpretation techniques mentioned above to explain the outputs of ML models for mapping the land susceptibility to soil erosion by water and wind simultaneously. Overall, integrated spatial maps of various hazards that are developed in this study, are especially important for soil erosion in the catchment area and can help the scientists, local authorities and policy makers in their management and in establishing the appropriate strategies to mitigate their negative consequences.

2. Materials and methods

2.1. Study area

Bakhtegan watershed, with an area of 31490 km^2 (51° 42' to 54° 31' E and 29° 00 to 31° 14' N), is located in the northern-central part of Fars Province, in southern Iran (Fig. 1). The study area has three lakes (Bakhtegan, Tashk and Maharloo). The Bakhtegan and Tashk Lakes are fed by Kor River, while the Maharloo Lake, located in southern Shiraz county, collects the surficial water and runoff from the upstream (Sajedipour et al., 2017). The main river in the study area is Kor, which emanates from the northwestern Fars Province and the Zagros Mountains, and its length is about 280 km. Sivand is the second river that feeds this lake. The mean elevation in the examined watershed basin is 2031 m, surrounded by high, arid and rocky mountains, while the mean annual temperature is about 17 °C. The region is characterized by a hot and dry climate, with an annual evaporation ranging between 1300 mm and 3500 mm and an annual rainfall amount of 365 mm. The construction of dams in the main rivers discharging in the lakes decreased the water surface in Bakhtegan Lake, while climate change and prolonged drought during the last decades had an important role in the desiccation of the lake, resulting in a substantial increase of dust emissions, in heavy metals and in total dissolved solids in lake waters (Karimzadeh and Taghizadeh, 2019; Vahidipour et al., 2022; Hamzeh et al., 2023).

2.2. Variables used for spatial mapping of soil erosion by water and wind

A wide range of effective factors consisting of land surface characteristics and roughness, topographic conditions, vegetation cover (with the annual mean values), lithology, parent materials, land use and land cover, along with climatic parameters (with the annual mean values) and soil characteristics have been found to control soil erosion by water and wind (Morgan, 2009; Gholami et al., 2020a; Mohammadifar et al., 2021; Renard and Freimund, 1994). Therefore, in this study, we applied 20 and 17 variables, as potential effective factors for water and wind erosion, respectively in the Bakhtegan catchment area, as shown in Table 1.

2.3. Soil erosion inventory map (SEIM)

SEIM indicates the points with the signs of soil erosion in the study area. This map is a necessary step for the spatial mapping of soil erosion. In order to map the inventory of soil erosion, 155 and 162 points with the signs of water erosion (e.g., rill erosion, sheet erosion, etc.) and wind erosion (e.g., soil detachment, deflation, abrasion, etc.), respectively were recorded by GPS during several field works and google earth images in the study area. For building the predictive models of water and wind erosion, 70 % and 30 % of the points were selected randomly as training and test datasets, respectively. In addition, 80 % and 20 % of the train dataset were selected for the training and validation of the predictive models, respectively. The inventory maps for the water and wind erosion are presented in Fig. 2.

2.4. Discrimination of important variables and collinearity tests

Following literature reviews (Gholami et al., 2020b, 2023a), 20 and 17 potential variables controlling water and wind erosion, respectively were mapped in the study region. The accuracy of the modelling process is depended on the input data. Furthermore, identification of the important and non-important variables is necessary before the spatial modelling phase of soil erosion by water and wind. In this respect, the MARS feature selection algorithm (Milborrow, 2014) was applied to discriminate the important factors, among all 20 and 17 variables that were initially considered as controlling factors. The *evimp* function in MARS was used to select the important variables and this technique utilized three measures (e.g., nsubsets, rss (residual sums of squares), and gcv (generalized cross-validation)) for identifying the important variables impacting on the soil erosion. The feature selection is a key step in the spatial modelling process of environmental hazards and this step can remarkably improve the accuracy of the predictions provided by ML models.

Then, in order to explore the multi-collinearity among the effective variables identified by MARS algorithm, two statistical tests consisting of TC and VIF was used. TC < 0.1 and VIF > 10 indicate the collinearity problem among the potential variables influencing the water erosion and wind erosion (Belsley, 1991; Salmerón-Gómez et al., 2020).

2.5. Spatial modelling of soil erosion

In the current research, seven ML models consisting of Cforest, support vector machine (SVM), Cubist, Elasticnet, stepwise regression, XGBoost, stepwise multivariate regression (SMR) and artificial neural network (ANN) were applied to map soil erosion. The use of seven of the most widely known and used ML models allows us to determine possible

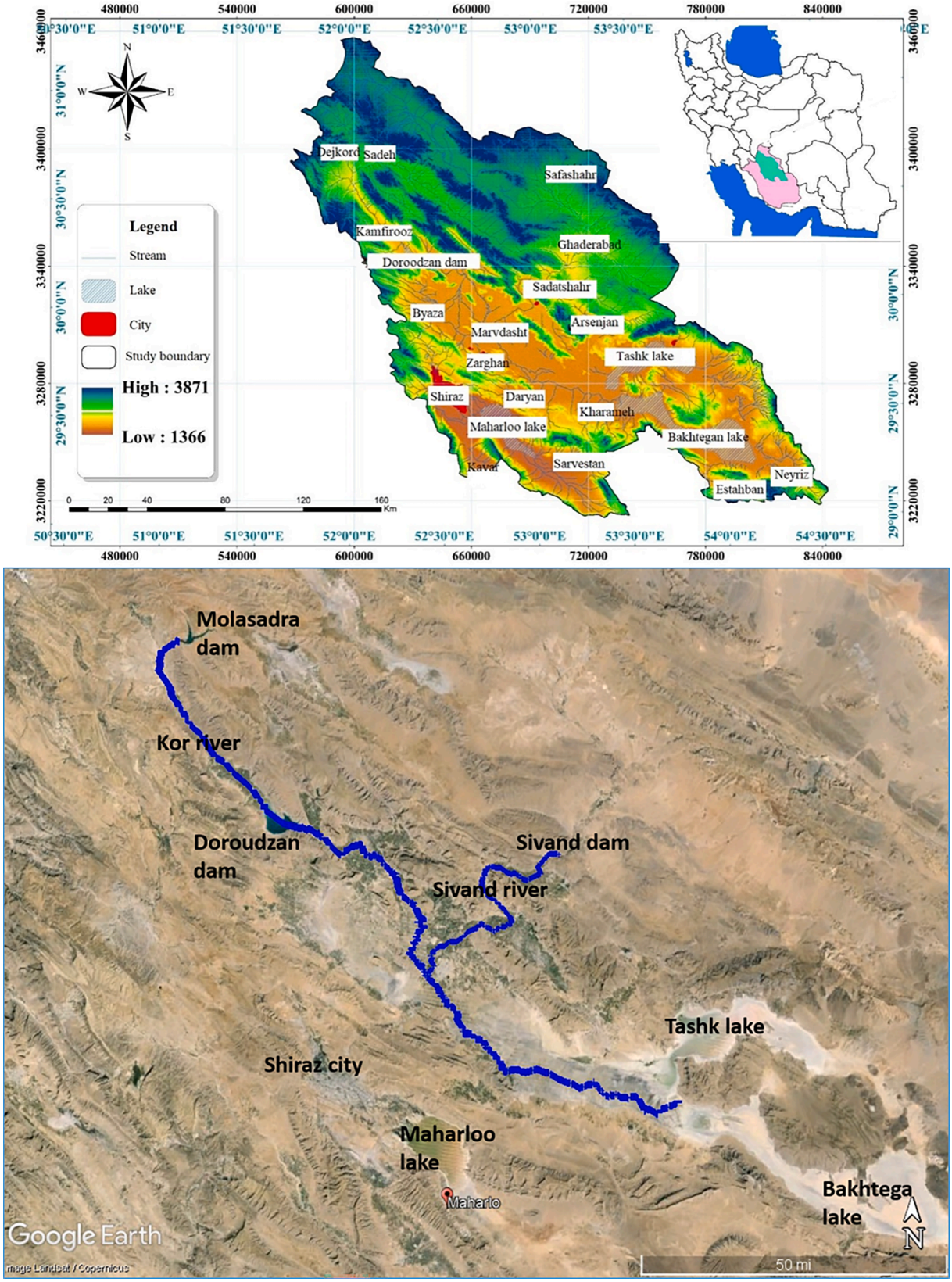

Fig. 1. Location of the Bakhtegan catchment area in Fars Province, southern Iran. Topography of the watershed basin with main toponyms of the lakes, cities, dams, etc.

Table 1
The description of variables used to map soil erosion by water and wind and their .

Variable	Water erosion	Wind erosion	Spatial resolution (m) and scale	Data source and software
DEM			30	Website: https://earthexplorer.usgs.gov/Software: ArcGIS 10.5
Slope			30	DEM and ArcGIS 10.5 software
Aspect			30	DEM and ArcGIS 10.5 software
Plan curvature		–	30	DEM and ArcGIS 10.5 software
Profile curvature		–	30	DEM and SagaGIS software
SPI		–	30	DEM and SagaGIS software
TWI		–	30	DEM and SagaGIS software
TRI			30	DEM and SagaGIS software
Clay content			250	Website: https://soilgrids.org/; Interpolation tools in ArcGIS10.5
Sand content			250	
Silt content			250	
Soil bulk density			250	
Coarse fragment			250	
Soil organic carbon			250	
CEC	–		250	
Soil nitrogen	–		250	
Soil moisture			250	
Wind speed	–		250	Website: https://cds.climate.copernicus.eu/; Interpolation tools in ArcGIS10.5
NDVI			90	Website: https://earthexplorer.usgs.gov/; ENVI software
Precipitation		–	250	Website: https:// worldclim.org/; Interpolation tools in ArcGIS10.5
Lithology			1:100,000	IWMRFO
Land use			1:100,000	
Distance from the river		–	1:100,000	

Sources. DEM, SPI, TWI, TRI, CEC and NDVI indicate digital elevation model, stream power index, topographic wetness index, terrain ruggedness index, cation exchange capacity, and normalized difference vegetation index, respectively

modelling biases, while the comparison between the outputs of the different models reveals the degree of consistency between the model predictions.

Cforest is an updated random forest (RF) model (Breiman, 2001). The accuracy of the predictive model depends on the optimal hyperparameters and this can improve the model's performance significantly. The high capability of the RF models has confirmed for spatial and temporal prediction purposes in different fields such as $PM_{2.5}$ concentrations (Tang et al., 2023; Zaman et al., 2021), COVID-19 health risk (Li et al., 2023), soil PH (Makungwe et al., 2021), gully erosion (Jiang et al., 2021), and fire occurrence (Oliveira et al., 2012), among many others.

SVM – a supervised algorithm introduced by Vapnik (1999) – is widely used to solve classification and regression problems. SVM algorithm constructs the hyperplane (Zheng et al., 2023) and the linear equation describing the hyperplane in the sample set space R^d can be written as follows (Dongliang et al., 2023):

$$w^T x + b = 0 \tag{1}$$

$$r = \frac{|w^T x + b|}{||w||}\pi r^2 \tag{2}$$

where, *w, b and r* indicate the hyperplane normal vector, the displacement constant and the hyperplane's distance, respectively.

A radial basis kernel function used in SVM model can be expressed as following:

$$K(x_i - x_j) = \exp\left(-\gamma|x_i - x_j|^2\right) \tag{3}$$

where γ indicates the kernel parameter and it determines the width of the radial basis kernel. The successful application of SVM has been reported for the prediction of water heat transfer coefficient (Dongliang et al., 2023), mineral prospectively mapping (Zheng et al., 2023), landslide spatial prediction (Hong et al., 2015), flood susceptibility mapping (Bera et al., 2022) and $PM_{2.5}$ predictions from space and ground observations (Zaman et al., 2021).

Cubist is a ML model based on the M5 algorithm (Quinlan, 1992; Quinlan, 1993) and in comparison with multivariate linear regression and neural networks, cubist models present better results and are easier to understand. This model has been employed for prediction of actual evapotranspiration (Zhang et al., 2023), prediction of the topsoil organic carbon (Szatmári et al., 2023), spatial and temporal prediction of ocean surface pCO_2 (Fu et al., 2020), soil sand content mapping (Qu et al., 2024), and other fields such as gully erosion, soil erosion by water and wind, soil salinity, and dust emissions.

Elastic net – a model based on the regularization mechanism of linear regression – is the extended LASSO (Friedman et al., 2010; Arabameri et al., 2021; Gholami et al., 2023b). This model is the combination of LASSO and ridge regression penalties (Ogutu et al., 2012), which can be written as follows:

$$\beta(enet) = \left(1 + \frac{\lambda_2}{n}\right)\left\{\beta^{argmin}||y - X\beta||_2^2 + \lambda_2||\beta||_2^2 + \lambda_1||\beta||_1\right\} \tag{4}$$

On the setting of $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, the elastic net estimator can be expressed as follows:

$$\beta(enet2) = \beta^{argmin}||y - X\beta||_2^2 \tag{5}$$

Subject to,

$$\rho_\alpha(\beta) = (1-\alpha)||\beta||_1 + \alpha||_2^2 \le sforsomes \tag{6}$$

where $\rho_\alpha(\beta)$ indicates the elastic net penalty (Zou and Hastie, 2005). The application of this model has been reported for different prediction purposes such as gully erosion, piping collapse, bio-concentration variables in fish and invertebrates, mapping wind erosion and dust sources (Gholami et al., 2023b), etc.

XGBoost is an optimized distributed gradient boosting library (Song et al., 2023) and offers various techniques to prevent overfitting and can also capture non-linear relationships in the model, suitable for handling large datasets. It has a wide range of hyperparameters that can be tuned to optimize the model's performance. Due to high speed and precision, this is a superior model (Parsa et al., 2020) and frequently used the CART model designed to reduce the computational time and scale up tree boosting algorithms (Chen and Guestrin, 2016).

SMR model simulates the linear relationship between dependent and explanatory variables. In this model, variables were selected automatically by stepwise regression, and simplify the initial model that used all variables (Liu et al.,2022). The successful application of this model has been reported in fields such as water quality assessment (Varol et al., 2022), crop yield variations analysis (Liu et al., 2021), nitrogen use efficiency analysis (Liu et al., 2022), soil moisture spatial modelling (Qiu et al., 2010), fine particles and nitrogen dioxide (Chen et al., 2019), among others.

ANN – an advanced model (Avdic et al., 2023) and an artificial

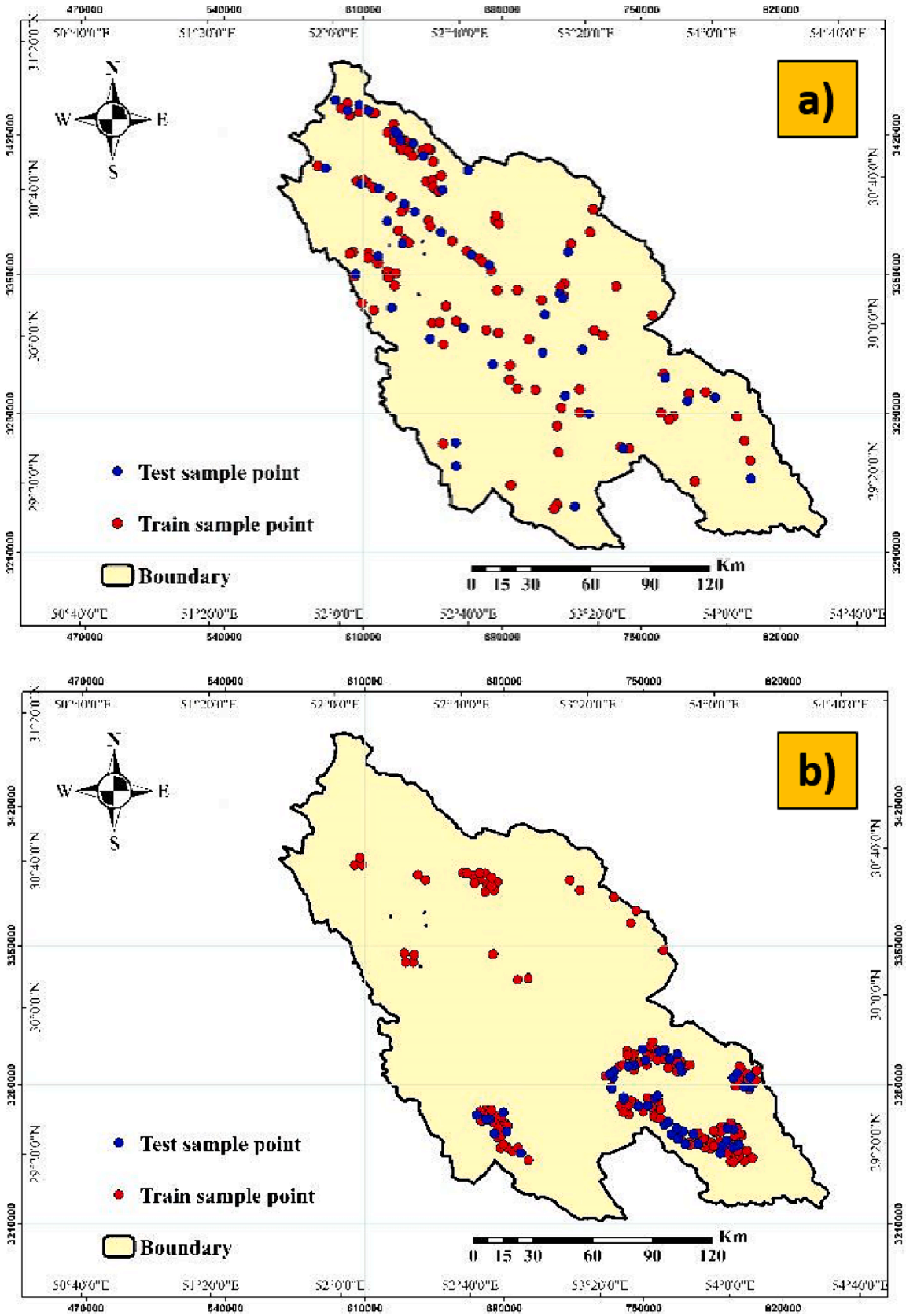

Fig. 2. The inventory map of soil erosion by a) water and b) wind in Bakhtegan catchment.

intelligence mechanism – is a good model to detect and simulate nonlinear relationships (Cander et al., 1998; Leandro and Santos, 2004). The sigmoid activation function used in the model can be written as follows:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{7}$$

2.6.8. Model's evaluation, integrated spatial map of soil erosion, and model's explanation

Area under curve (AUC) (Yesilnacar, 2005) was used to evaluate the predictive model performance. Based on the AUC values, the model's accuracy can be divided into five classes. Finally, the most accurate models identified for mapping water and wind erosion were used to generate an integrated spatial map of soil erosion over the study area.

Here, we applied the PFI measure and SHAP plot to explain the predictive model used to map soil erosion. The PFI measure was initially introduced by Breiman, (2001) and is computed by calculating the increase in the model's prediction error after permuting the feature.

SHAP – a technique based on game theory (Lundberg and Lee, 2017) and optimal Shapley values – was applied to address the nature of the black box of predictive ML model used to mapping soil erosion. More details about PFI and SHAP can find elsewhere (Mohammadifar et al., 2021; Gholami et al., 2023a). All the main steps and the whole procedure followed for generating integrated maps of soil erosion by water and wind in the Bakhtegan watershed are presented in the flow chart of Fig. 3.

3. Results and discussion

3.1. Effective variables, their collinearity behavior and importance

The important (effective) variables controlling soil erosion, as they identified by MARS, are presented in Table 2. The results show that out of 20 variables potentially controlling water erosion, ten variables consisting of land use type, slope, lithology, terrain ruggedness index (TRI), coarse fragment, clay content, vegetation cover (NDVI), aspect, soil bulk density and topographic wetness index (TWI) were computationally selected as important factors affecting water erosion in the study area, while the others were identified as non-important variables and were excluded from the rest of the analysis.

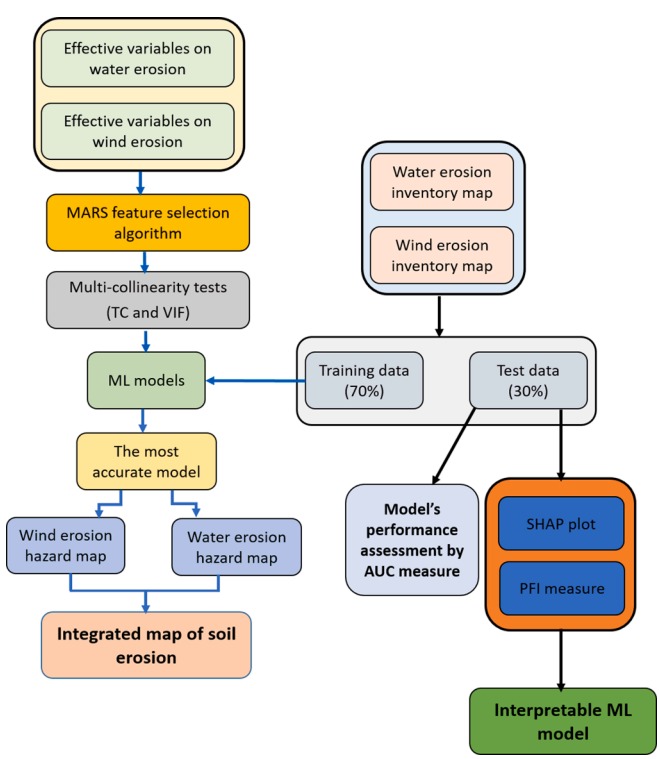

Fig. 3. Research diagram indicating all the main steps used to generate integrated maps of soil erosion by water and wind and its explanation.

Table 2
Effective (important) variables controlling water erosion and wind erosion as identified by MARS feature selection algorithm. Gcv, Rss, TRI, NDVI, TWI, DEM, and CEC indicate generalized cross-validation, residual sums of squares, terrain ruggedness index, normalized difference vegetation index, topographic wetness index, digital elevation model, and cation exchange capacity, respectively.

Soil erosion by water			
Variable	*Nsubsets*	*Gcv*	*Rss*
Land use type	15	100	100
Slope	15	60.2	64.8
Lithology	12	35	42.5
TRI	11	24.7	35.1
Coarse fragment	10	26	34.4
Clay content	7	19.1	26.9
NDVI	7	15.8	25.1
Aspect	6	13.4	22.6
Soil bulk density	5	14.3	21.4
TWI	4	9.3	17.6
Soil erosion by wind			
Variable			
DEM	16	100	100
Land use type	15	74.7	75.8
Wind speed	14	50.3	55.3
Slope	14	49	53.2
NDVI	11	33.6	39.2
Soil nitrogen	10	31.9	36.5
CEC	9	30.3	34.4
Lithology	8	25.1	29.9
Aspect	6	15	21.7
Clay content	4	10.8	16.9
Soil bulk density	1	2.2	7.3

On the other hand, the results of MARS feature selection algorithm revealed that DEM, land use type, wind speed, slope, NDVI, soil nitrogen, cation exchange capacity (CEC), lithology, aspect, clay content and soil bulk density were identified as important variables controlling wind erosion. Overall, seven variables, namely land use type, slope, lithology, clay content, NDVI (vegetation cover), aspect and soil bulk density were found as important factors for affecting soil erosion by both water and wind.

The PFI plots indicating the relative importance of variables impacting on water erosion and wind erosion are presented in Fig. 4. According to Fig. 4**a**, among the ten identified effective variables impacting on the water erosion, five of them, namely slope, TWI, coarse fragment, land use and TRI were considered as the most important ones with a high relative importance. The severity of water erosion and wind erosion depends on various variables (e.g., soil characteristics, topography conditions, land use type, precipitation regime type, etc. (Renard and Freimund, 1994; Morgan, 2009; Parajuli et al., 2014; Padarian et al., 2020; Gholami et al., 2020a).

The importance of variables controlling soil erosion is highly site or area-specific, since they depend on the specific characteristics of each area, soil conditions, anthropogenic interference, and the prevailing synoptic meteorology. For example, Mohammadifar et al. (2021) reported that variables extracted from DEM (e.g., TWI, TRI, terrain surface texture, slope and SPI) were the most effective factors controlling water erosion in Hormozgan Province, southern Iran. In the work by Khosravi et al. (2023), among 14 variables explored as affective ones for controlling water erosion, elevation or DEM, rain erosivity, vegetation cover, TWI and plan curvature were identified as the most important factors. Although the importance of the topographic variables is different between catchment areas of different climatic, soil and

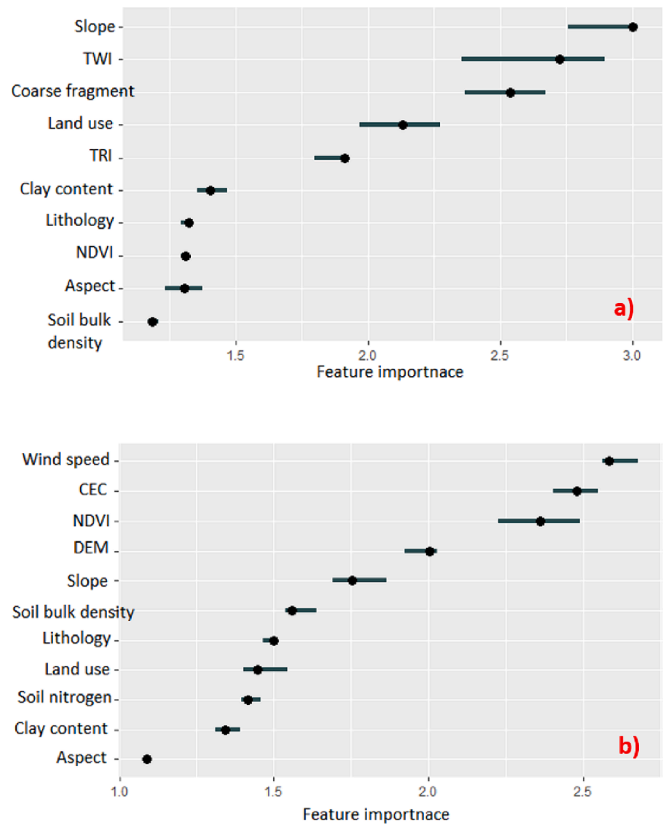

Fig. 4. PFI measure plots indicating the importance of the variables controlling a) water erosion, and b) wind erosion.

topographic characteristics, slope length variable and channel network distance were identified as the most important factors controlling gully erosion in the black-soil region of China (Huang et al., 2022). According to Roy and Saha (2022), land use – land cover, length of overland flow, elevation, slope, and plan curvature were introduced as the five most important factors controlling gully headcut in India. Altitude, distance from the stream network, distance from the road, land use type and vegetation cover were selected as the key variables controlling gully erosion in Bandar Lengeh, Hormozgan Province, south Iran (Wang et al., 2021). Soil properties, DEM, land use type, and vegetation cover were recognized as the critical variables impacting on the gully erosion in Shamil-Mianb plain, southern Iran (Gholami et al., 2023a). This shows that evaluation of the effective variables and mapping of soil erosion should be emphasized on local to regional scales in order to achieve the most reliable results, due to large spatial heterogeneity, human interference and different climatic conditions of each area.

The PFI plot regarding the relative importance of variables influencing wind erosion is presented in Fig. 4**b**. The PFI results revealed that, among the 11 selected variables impacting wind erosion, those of wind speed, CEC, NDVI, DEM and slope are the five ones with the highest relative importance. Overall, soil erosion by wind is impacted by different variables and its severity varied from one point to another, exhibiting high association with topographic and meteorological conditions. For example, Gholami et al., (2020a) revealed that among thirteen variables evaluated as controlling factors for wind erosion in central Asia, three of them consisting of DEM or elevation, precipitation and vegetation cover exhibited the highest relative importance. In another work by Gholami et al., (2020b), three variables (e.g., wind speed, land use type and elevation) were the most important controlling factors on dust source regions in southwestern Iran. According to Boroughani et al. (2021), soil orders and geomorphology units were the most important variables controlling dust emissions from the land surfaces in the arid Sistan watershed. Pourhashemi et al. (2023) revealed that land use type exhibited the highest impact on the dust source susceptibility in a critical region located at the border of Iran and Iraq. Overall, depending on many variables such as human activities, environmental, climatic, topographic and land surface conditions, soil erosion by water and wind may significantly change between different catchment areas.

The results of TC and VIF tests, which were used to explore the collinearity among the variables impacting on water erosion and wind erosion, are presented in Tables 3 and 4. The results revealed that there is no significant collinearity among the effective variables controlling the water and wind erosions, because the values of TC and VIF were calculated > 0.1 and < 10, respectively for all effective variables. Therefore, all the implemented factors are considered as independent, and without strong relationships between them, and therefore, could be used to generate maps of wind erosion by water and wind in the Bakhtegan watershed basin.

3.2. Water erosion and wind erosion spatial maps

Seven ML models were applied to map the soil erosion in the Bakhtegan watershed basin, and then, their performance was assessed by AUC technique. The results of the assessment of the model's performance were presented in Fig. 5. These results revealed that, among seven models used to map water erosion and wind erosion, ANN and Cubist are the most accurate models or the models with the highest reliability to map soil erosion by water and wind, respectively, since the area under the curve maximizes in these cases (Fig. 5). According to ANN model (Fig. 6**a**), 13.5 %, 39.4 %, 18.2 % and 28.9 % of the total land area in the examined watershed were classified as low to very high susceptible to water erosion, while, Cubist model (Fig. 6**b**) revealed that 9.2 %, 24.3 %, 28.6 % and 37.7 % of the study area was characterized by low to very high susceptibility to soil erosion by wind.

The water erosion map revealed that the central lands of the study area exhibited high susceptibility to soil erosion by water, while several hotspots of water erosion were also identified in the northern parts of the catchment area.

The southern and northern lands of the Bakhtegan watershed basin seem to be highly susceptible to wind erosion, because three desiccated lakes consisting of Tashk, Bakhtegan and Mharloo are located in these regions. After their dryness, fine-grained particles (e.g., very fine sand, silt and clay) remained in the topsoil, which are highly susceptible to wind and can be emitted in the atmosphere under strong near surface winds causing intense dust storms in the Bakhtegan Basin (Karimzadeh and Taghizadeh, 2019; Hamzeh et al., 2023).

Bakhtegan Lake was the second largest lake in Iran, and the largest lake in the central Iranian Plateau, with a length of 120 km. Unfortunately, due to intensified climate change, drought and limited precipitation over the region during the recent years, combined with anthropogenic activities (i.e. excessive consumption of lake's water, withdrawal of underground aquifers, drilling of agricultural wells), the

Table 3
TC and VIF values for the effective variables controlling soil erosion by water.

Variable	Collinearity tests	
	TC	VIF
TWI	0.371	2.698
TRI	0.125	8.542
Slope	0.146	8.082
Clay content	0.473	2.115
Coarse fragment	0.115	8.684
Aspect	0.983	1.017
Soil bulk density	0.273	3.663
Land use type	0.304	3.287
Lithology	0.566	1.767
NDVI	0.670	1.493

Table 4
TC and VIF values for the effective variables controlling soil erosion by wind.

Variable	Collinearity tests	
	TC	VIF
Aspect	0.937	1.068
Soil bulk density	0.546	1.83
CEC	0.416	2.405
DEM	0.182	5.504
Land use type	0.626	1.597
Lithology	0.678	1.475
NDVI	0.459	2.177
Slope	0.528	1.895
Clay content	0.721	1.388
Soil nitrogen	0.649	1.541
Wind speed	0.229	4.428

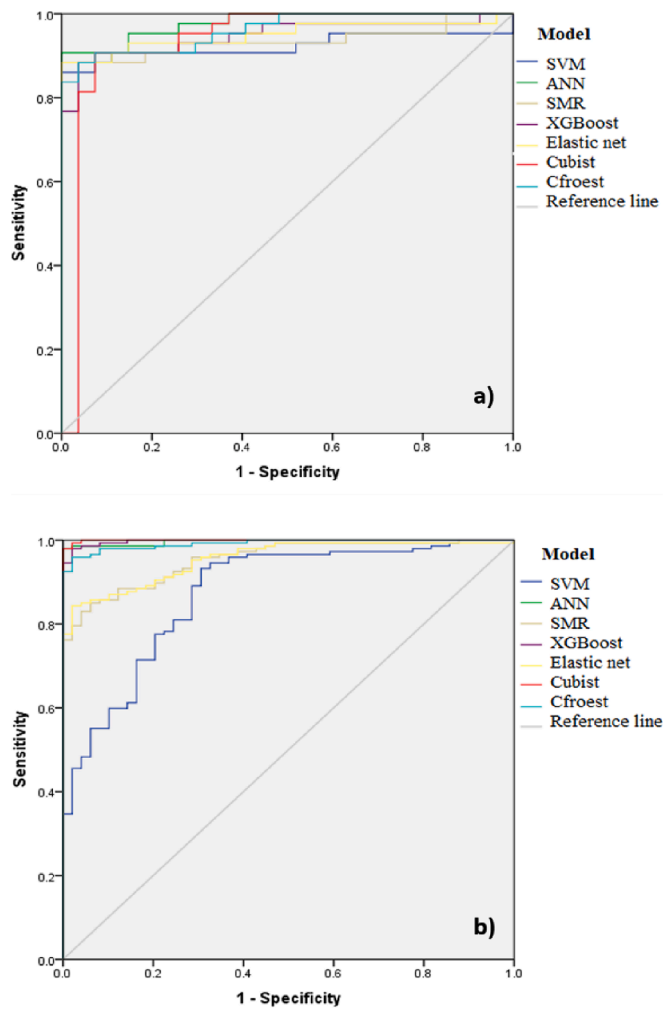

Fig. 5. The AUC plots for assessment of the performance of ML models used to map a) water erosion, and b) wind erosion.

Kor River was dried, thus causing the complete dryness of the lake (Sajedipour et al., 2017; Karimzadeh and Taghizadeh, 2019; Vahidipour et al., 2022). Construction of three dams namely Doroudzan dam – constructed in the Kor River and upstream of the study area in 1971 –, Sivand dam on the Sivand River, and Molasadra dam decreased dramatically the incoming water flow into Bakhtegan lake (Sajedipour et al., 2017). Therefore, the dryness of the lakes was intensified in recent years, converting the basin into an active dust source, with frequent emissions of saline dust storms due to alluvial material and halite content remained in the topsoil after the lake's dryness (Karimzadeh and

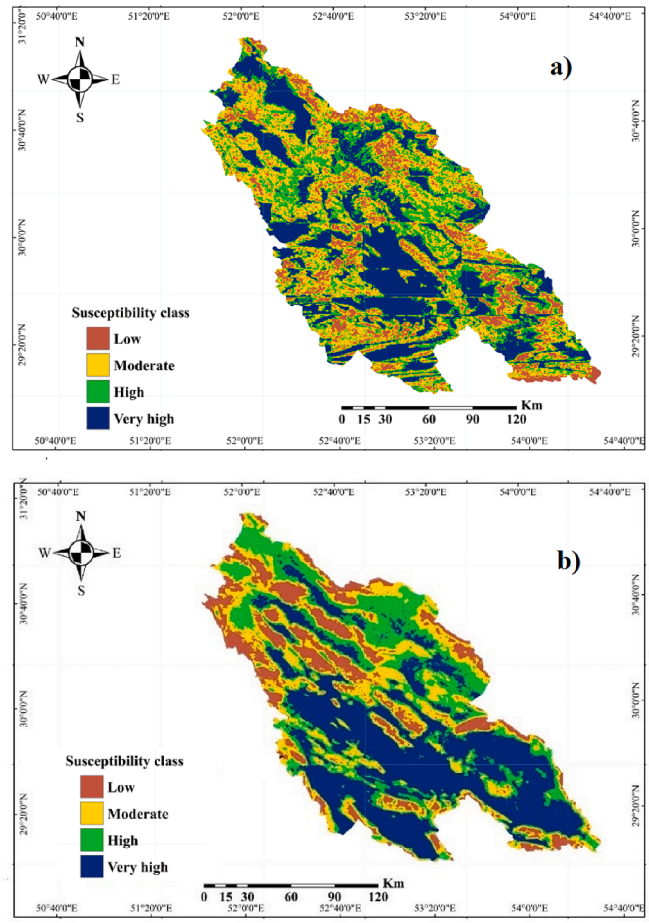

Fig. 6. The land susceptibility classes due to a) water erosion predicted by ANN model, and b) wind erosion predicted by Cubist model in the Bakhtegan watershed basin in south Iran.

Taghizadeh, 2019; Hamzeh et al., 2023). Consequently, this caused a degradation of air quality, accompanied with an increase in PM concentrations and potentially toxic metals, with a serious threat for human health and ecosystem services (Vahidipour et al., 2022).

Similar phenomena of intense dust storms originated from dried lake beds or desiccated lakes have been observed in the Urmia lake, NW Iran (Motaghi et al., 2020; Hamzehpour et al., 2022; Hamzeh et al., 2023; Mobarak Hassan et al., 2023), in the Sistan Basin, SE Iran (Rashki et al., 2013, 2021), in the Aral Sea (Karami et al., 2021; Banks et al., 2022; Ge et al., 2022), as well as in cases in Africa (Etosha Pans, Chad Lake), Australia (Lake Eyre) and north America. Overall, the desiccated lake beds present high susceptibility to wind erosion and dust emissions and may play a critical role to intensifying air pollution in the neighboring affected areas (Zucca et al., 2021; Wang et al., 2022; Wu et al., 2022; Tehrani et al., 2023). For example, Behrooz et al. (2019) and Rashki et al. (2013) reported that dried beds of the Hamoun lakes act as dust source in the Sistan Basin in southeastern Iran, while dryness of these lakes can increase dust emissions by an average of 40 %, reaching even up to 80 % (Boloorani et al., 2022). The current results revealed that on watershed basins in the interior of arid lands, apart from the wind erosion, water erosion is also especially important for land degradation and negative impacts on environmental processes. Detailed (under high spatial resolution) and accurate mapping of land susceptibility to wind and water erosion is especially important for sustainable resource management in the arid environments and for mitigating the effects of dust emissions.

3.3. Contribution of variables (features) controlling the soil erosion

Due to nature of the black-box of ML models, understanding the importance and contribution of inputs or effective variables on the predictive model's outputs is necessary. Therefore, interpretation techniques are effective tools to explain the model's output. Here, SHAP plots (Fig. 7), as interpretation techniques, were applied to assess the importance and contribution of the effective variables controlling water erosion and wind erosion. According to the results, slope, land use, TWI, TRI and soil bulk density exhibited the highest impact (contribution) on the predictive water erosion model's output (Fig. 7**a**), while SHAP plot revealed that NDVI, slope, land use, wind speed and soil clay content were the most important variables contributing to the wind erosion model's output (Fig. 7**b**). More recently, the application of interpretation techniques such as SHAP, PFI, PDP, is increasingly available to explain the model's output used to map soil erosion by water and wind. For example, Gholami et al., (2023b) applied a SHAP interpretation method to explain the output of a deep learning model used for mapping soil erosion by wind in central Iran, and they revealed that soil coarse fragment, vegetation cover and land use were the three most important variables contributing to the wind erosion model's output. According to Mohammadifar et al. (2021), DEM-based variables (e.g., TWI, TRI, slope length, sediment connectivity index, and TPI) exhibited the highest contribution to the water erosion model's output in the Kahorestan catchment, southern Iran. In another work, three variables consisted of clay content, silt content, and precipitation were recognized as the most important contributing factors on Middle East's land susceptibility to dust emissions (Gholami and Mohammadifar, 2022).

3.4. Integrated map of soil erosion by water and wind

Among all ML models used to map the water and wind erosions, ANN and Cubist exhibited the highest accuracy for mapping land susceptibility to water erosion and wind erosion, respectively and the generated maps were based on these two models (Fig. 6). In a step forward, these two models were combined to integrate water erosion and wind erosion concurrently in the studied catchment area (Fig. 8). It should be noted that it's the first time that a combined map for soil erosion is constructed concurrently for wind and water erosion in south Iran. The results revealed that 7.7 %, 9.6 % and 3.4 % of the total area are characterized by very high susceptibility (VHS), VHS + high susceptibility (HS), and HS classes to soil erosion by water and wind, respectively (Table 5). Climate change (increase in temperature, decrease in precipitation, prolonged droughts) in the Middle East, combined with land use changes on steep slopes in the study area, are expected to significantly impact soil erosion and may cause environmental challenges with negative atmospheric, climatic, social and economic consequences (Nunes et al., 2011; Vahidipour et al., 2022). Intensified droughts resulted from climate change had a remarkable influence on the water resources in Iran, especially on the lakes in Fars Province and in southeastern parts of the country (Shahraki et al., 2017; Hashemi Monfared et al., 2017; Farrokhzadeh et al., 2020). Due to global warming and climate change, as well as increasing water demands for agricultural, potable and industrial purposes, several dams were constructed in the upstream parts of the three lakes in Bakhtegan watershed, and the discharge into the wetlands was decreased remarkably. Nowadays, the dried lakebeds have been converted to local sources of dust and sand storms, thus modulating the climate, atmospheric conditions and air quality of the study area. Shiraz city, with a population of about 1.8 million, is located in the Bakhtegan Basin and is highly affected by local dust storms originated from the dried lakes. Consequently, PM concentrations have increased in the city, which also suffers from large amounts of potentially toxic heavy metals transferred or adhered in dust particles, with deleterious effects on human health, as several recent works underlined (Keshavarzi et al., 2015; Moghtaderi et al., 2019; Torghabeh et al., 2020; Abbasi et al., 2021). Therefore, the current constructed map combining the effects of water and wind erosion on the arid lands in the Bakhtegan Basin is especially important for scientists working on the hydrological, geological, atmospheric and health topics, as well as for local authorities and policy makers for establishing the appropriate mitigation strategies in a way to protect land from further degradation and soil from excessive wind and water erosion.

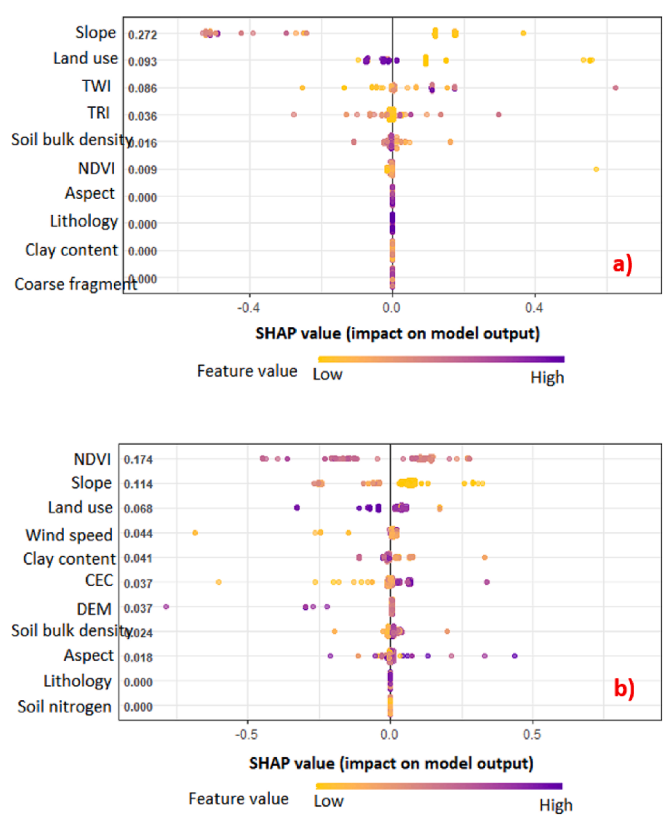

Fig. 7. SHAP plots indicating the contribution of the variables controlling a) water erosion, and b) wind erosion, on the model's output.

4. Conclusions

This study is the first application of explainable ML modeling based on interpretation techniques and ML models to map land susceptibility simultaneously to water erosion and wind erosion in an arid catchment area in southern Iran with three desiccated lakes.

The MARS feature selection algorithm was applied to identify the factors that influence water and wind erosion in the study area, among 20 and 17 potential variables, respectively. The results showed that 10 and 11 factors controlled the water and wind erosion, respectively, which are implemented in seven machine learning models in a way to provide the most reliable and accurate map for water and wind erosion in Bakhtegan watershed basin. The collinearity behavior among the effective variables was explored through the tolerance coefficient and variance inflation factor tests, which showed an insignificant correlation between the independent variables. The most accurate ML models for water and wind erosion in the Bakhtegan catchment area were found to be the artificial neural network (ANN) and Cubist, respectively, while an integrated model was developed for simultaneous mapping of soil erosion due to water and wind. The results indicated that 17.25 % of the total area is characterized by very high susceptibility to water erosion, while the respective fraction for the wind erosion was found to be 37.7 %.

Due to intensified droughts during the last two decades, construction of dams in the upstream parts of the three lakes, and dryness of the lakes,

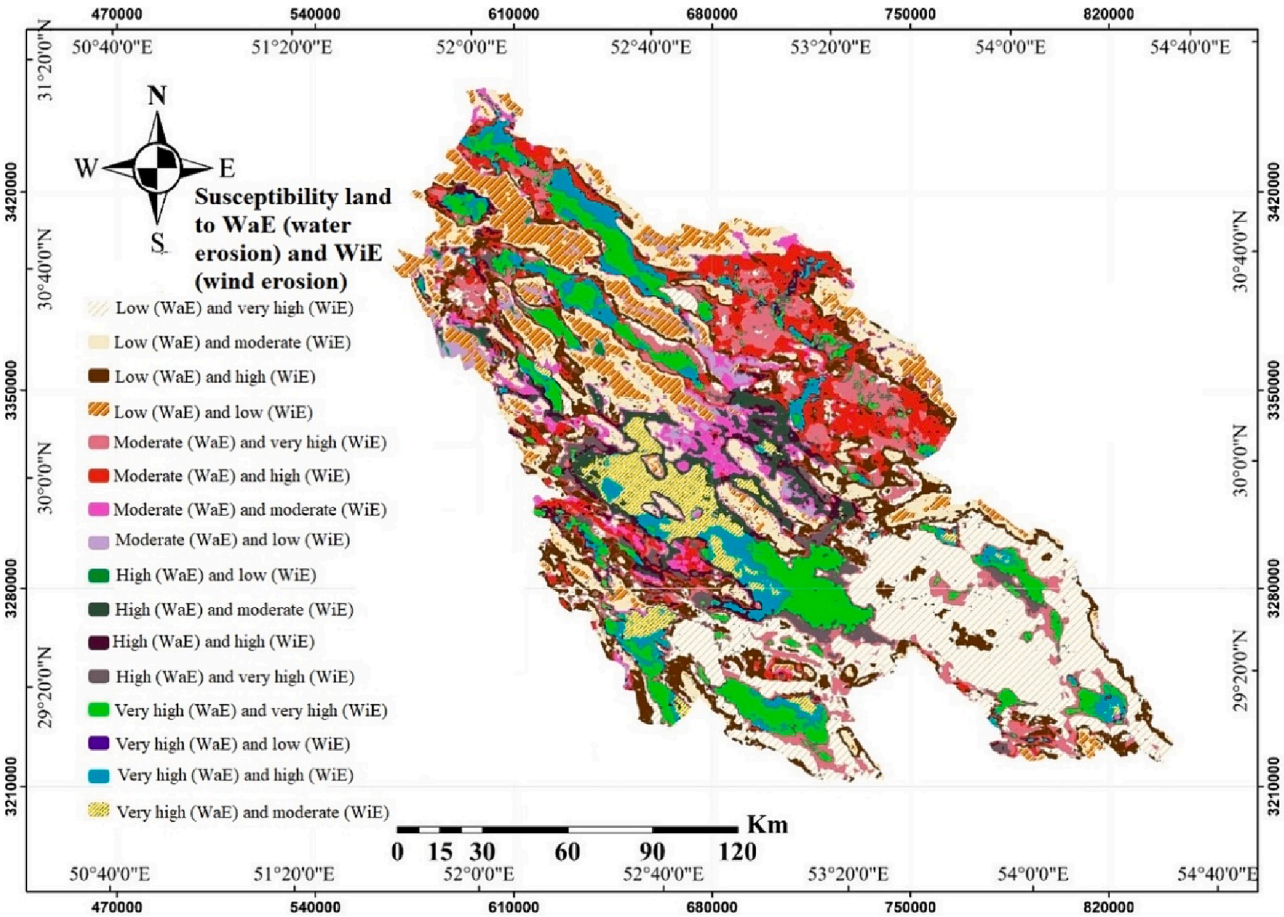

Fig. 8. Integrated map of soil erosion by water and wind in the Bakhtegan watershed area, south Iran.

Table 5
The area (in km^2 and %) of land susceptibility classes to soil erosion by water and wind. LS, MS, HS and VHS indicate the low susceptibility, moderate susceptibility, high susceptibility and very high susceptibility, respectively.

Water erosion	Wind erosion	Area (ha)	Area (%)
LS	LS	2396	7.62
LS	MS	3875	12.32
LS	HS	3431	10.9
LS	VHS	5661	18
MS	LS	452	1.44
MS	MS	1317	4.2
MS	HS	2724	8.638
MS	VHS	2542	8.1
HS	LS	71	0.23
HS	MS	1254	4
HS	HS	1055	3.4
HS	VHS	1234	3.9
VHS	LS	0.5	0.002
VHS	MS	1211	3.85
VHS	HS	1789	5.7
VHS	VHS	2426	7.7

the central-southern parts of the study area were converted to sources of dust and sand storms, also presenting high to very high susceptibility to water and wind erosion.

Overall, integrated spatial maps of different environmental hazards (e.g., soil erosion, landslide, land subsidence, soil salinity, etc.) can provide accurate information for sustainable management and restoration of different environments. On the other hand, interpretation of black-box model's output is necessary to detect relationship between input variables and model's output. Therefore, we recommend that future research should explore this lesser-known aspect (interpretation) of the spatial ML models through common techniques like PDP, LIME, ALE, SHAP, and other local and global interpretation techniques. Simultaneous mapping of land susceptibility to water and wind erosion over the arid and semi-arid areas is especially important for local authorities, policy makers, scientists, and local habitants in order to apply environmental restoration techniques and to prevent further land degradation under a negative climate change projection for the Middle East region.

CRediT authorship contribution statement

Hamid Gholami: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mehdi Jalali:** Methodology, Investigation, Formal analysis. **Marzieh Rezaei:** Investigation, Formal analysis. **Aliakbar Mohamadifar:** Methodology, Investigation, Formal analysis. **Yougui Song:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Investigation, Formal analysis, Conceptualization. **Yue Li:** Investigation, Formal analysis, Conceptualization. **Yanping Wang:** Investigation, Formal analysis, Conceptualization. **Baicheng Niu:** Investigation, Formal analysis, Conceptualization. **Ebrahim Omidvar:** Investigation, Formal analysis, Conceptualization. **Dimitris G. Kaskaoutis:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization.

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.

Data availability

Data will be made available on request.

Acknowledgment

This research is funded by the PIFI and Strategic Priority Research Program of Chinese Academy of Sciences (XDB40000000) and also Natural Science Foundation of China (No: 41977385).

References

Abbasi, S., Rezaei, M., Keshavarzi, B., Mina, M., Ritsema, C., Geissen, V., 2021. Investigation of the 2018 Shiraz dust event: Potential sources of metals, rare earth elements, and radionuclides; health assessment. Chemosphere 279, 130533.
Arabameri, A., Rezaie, F., Pal, S.C., Cerda, A., Saha, A., Chakrabortty, R., Lee, S., 2021. Modelling of piping collapses and gully headcut landforms: Evaluating topographic variables from different types of DEM. Geosci. Front. 12 (6), 101230.
Avdic, S., Dykin, V., Croft, S., Pázsit, I., 2023. Item identification with a space-dependent model of neutron multiplicities and artificial neural networks. In: Nuclear Instruments and Methods in Physics Research Section a: Accelerators, Spectrometers, Detectors and Associated Equipment, p. 168800.
Bag, R., Mondal, I., Dehbozorgi, M., Bank, S.P., Das, D.N., Bandyopadhyay, J., Nguyen, X.C., 2022. Modelling and mapping of soil erosion susceptibility using machine learning in a tropical hot sub-humid environment. J. Clean. Prod. 364, 132428.
Banks, J.R., Heinold, B., Schepanski, K., 2022. Impacts of the Desiccation of the Aral Sea on the Central Asian Dust Life-Cycle. J. Geophys. Res. Atmos. 127 (21), e2022JD036618.
Behrooz, R.D., Gholami, H., Telfer, M.W., Jansen, J.D., Fathabadi, A., 2019. Using GLUE to pull apart the provenance of atmospheric dust. Aeolian Res. 37, 1–13.
Belsley, D.A., 1991. A guide to using the collinearity diagnostics. Computer Science in Economics and Management 4 (1), 33–50.
Bera, S., Das, A., Mazumder, T., 2022. Evaluation of machine learning, information theory and multi-criteria decision analysis methods for flood susceptibility mapping under varying spatial scale of analyses. Remote Sens. Appl.: Soc. Environ. 25, 100686.
Blanco-Canqui, H., Lal, R., 2008. Principles of soil conservation and management. Springer Science & Business Media.
Boloorani, A.D., Najafi, M.S., Soleimani, M., Papi, R., Torabi, O., 2022. Influence of Hamoun Lakes' dry conditions on dust emission and radiative forcing over Sistan plain. Iran. Atmospheric Research 272, 106152.
Boroughani, M., Pourhashemi, S., Gholami, H., Kaskaoutis, D.G., 2021. Predicting of dust storm source by combining remote sensing, statistic-based predictive models and game theory in the Sistan watershed, southwestern Asia. J. Arid. Land 13 (11), 1103–1121.
Breiman, L., 2001. Random forests. Machine Learning 45, 5–32.
Cander, L.R., Milosavljevic, M.M., Stankovic, S.S., Tomasevic, S., 1998. Ionospheric forecasting technique by artificial neural network. Electron. Lett 34 (16), 1573–1574.
Chen, J., de Hoogh, K., Gulliver, J., Hoffmann, B., Hertel, O., Ketzel, M., Hoek, G., 2019. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. Environ. Int. 130, 104934.
Chen, T., Guestrin, C., 2016. August). Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794.
Dongliang, M., Yi, L., Tao, Z., Yanping, H., 2023. Research on prediction and analysis of supercritical water heat transfer coefficient based on support vector machine. Nucl. Eng. Technol. 55 (11), 4102–4111.
Farrokhzadeh, S., Hashemi Monfared, S.A., Azizyan, G., Sardar Shahraki, A., Ertsen, M. W., Abraham, E., 2020. Sustainable water resources management in an arid area using a coupled optimization-simulation modeling. Water 12 (3), 885.
Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1.
Fu, Z., Hu, L., Chen, Z., Zhang, F., Shi, Z., Hu, B., Liu, R., 2020. Estimating spatial and temporal variation in ocean surface pCO2 in the Gulf of Mexico using remote sensing and machine learning techniques. Sci. Total Environ. 745, 140965.
Ge, Y., Wu, N., Abuduwaili, J., Kulmatov, R., Issanova, G., Saparov, G., 2022. Identifying Seasonal and Diurnal Variations and the Most Frequently Impacted Zone of Aerosols in the Aral Sea Region. Int. J. Environ. Res. Public Health 19 (21), 14144.
Gharechaei, H., Moghaddam Nia, A., Malekian, A., Ahmadi, A., 2015. Separation of the effects of climate variability and human activities on runoff of Bakhtegan Basin. Iranian J. Ecohydrol. 2 (4), 445–454.
Gholami, H., Mohamadifar, A., Sorooshian, A., Jansen, J.D., 2020a. Machine-learning algorithms for predicting land susceptibility to dust emissions: The case of the Jazmurian Basin, Iran. Atmos. Pollut. Res. 11 (8), 1303–1315.
Gholami, H., Mohammadifar, A., Bui, D.T., Collins, A.L., 2020b. Mapping wind erosion hazard with regression-based machine learning algorithms. Sci. Rep. 10 (1), 20494.
Gholami, H., Mohammadifar, A., 2022. Novel deep learning hybrid models (CNN-GRU and DLDL-RF) for the susceptibility classification of dust sources in the Middle East: a global source. Sci. Rep. 12 (1), 19342.
Gholami, H., Mohamadifar, A., Collins, A.L., 2020c. Spatial mapping of the provenance of storm dust: Application of data mining and ensemble modelling. Atmos. Res. 233, 104716.
Gholami, H., Mohammadifar, A., Golzari, S., Song, Y., Pradhan, B., 2023a. Interpretability of simple RNN and GRU deep learning models used to map land susceptibility to gully erosion. Sci. Total Environ. 904, 166960.
Gholami, H., Mohammadifar, A., Fitzsimmons, K.E., Li, Y., Kaskaoutis, D.G., 2023b. Modeling land susceptibility to wind erosion hazards using LASSO regression and graph convolutional networks. Front. Environ. Sci. 11, 1187658.
Hamzeh, N.H., Shukurov, K., Mohammadpour, K., Kaskaoutis, D.G., Saadatabadi, A.R., Shahabi, H., 2023. A comprehensive investigation of the causes of drying and increasing saline dust in the Urmia Lake, northwest Iran, via ground and satellite observations, synoptic analysis and machine learning models. Eco. Inform. 78, 102355.
Hamzehpour, N., Marcolli, C., Pashai, S., Klumpp, K., Peter, T., 2022. Measurement report: The Urmia playa as a source of airborne dust and ice-nucleating particles–Part 1: Correlation between soils and airborne samples. Atmos. Chem. Phys. 22 (22), 14905–14930.
Hashemi Monfared, S.A., Zamani, Y., Azhdary Moghaddam, M., 2017. Water resource simulation in an arid watershed considering an allocation for dust stabilization and different scenarios of water supply and demand. J. Hydrosci. Environ. 1 (1), 33–41.
Hong, H., Pradhan, B., Xu, C., Bui, D.T., 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 133, 266–281.
Huang, D., Su, L., Fan, H., Zhou, L., Tian, Y., 2022. Identification of topographic factors for gully erosion susceptibility and their spatial modelling using machine learning in the black soil region of Northeast China. Ecol. Ind. 143, 109376.
Jena, R., Shanableh, A., Al-Ruzouq, R., Pradhan, B., Gibril, M.B.A., Khalil, M.A., Ghorbanzadeh, O., Ghamisi, P., 2023a. Earthquake spatial probability and hazard estimation using various explainable AI (XAI) models at the Arabian Peninsula. Soc. Environ. Remote Sens. Appl., 101004
Jena, R., Shanableh, A., Al-Ruzouq, R., Pradhan, B., Gibril, M.B.A., Khalil, M.A., Ghamisi, P., 2023b. Explainable artificial intelligence (XAI) model for earthquake spatial probability assessment in Arabian peninsula. Remote Sens. (Basel) 15 (9), 2248.
Jiang, C., Fan, W., Yu, N., Liu, E., 2021. Spatial modeling of gully head erosion on the Loess Plateau using a certainty factor and random forest model. Sci. Total Environ. 783, 147040.
Kafilzadeh, F., Javid, H., Kargar, M., 2007. Isolation of halophilic and halotolerant microorganisms from the Bakhtegan lake and the effect of physicochemical factors on their frequency. Water and Wastewater 18 (3), 81–87.
Karimzadeh, S., Taghizadeh, M.M., 2019. Potential of dust emission resources using small wind tunnel and GIS: case study of Bakhtegan playa, Iran. Appl. Water Sci. 9 (8), 174.
Keshavarzi, B., Tazarvi, Z., Rajabzadeh, M.A., Najmeddin, A., 2015. Chemical speciation, human health risk assessment and pollution level of selected heavy metals in urban street dust of Shiraz, Iran. Atmos. Environ. 119, 1–10.
Khosravi, K., Rezaie, F., Cooper, J.R., Kalantari, Z., Abolfathi, S., Hatamiafkoueieh, J., 2023. Soil water erosion susceptibility assessment using deep learning algorithms. J. Hydrol. 618, 129229.
Lana, J.C., Castro, P.D.T.A., Lana, C.E., 2022. Assessing gully erosion susceptibility and its conditioning factors in southeastern Brazil using machine learning algorithms and bivariate statistical methods: A regional approach. Geomorphology 402, 108159.
Leandro, R.F., Santos, M.C., 2004. Comparison between autoregressive model and neural network for forecasting space environment parameters. Bollettino Di Geodesia e Scienze Affini 63 (3), 197–212.
Li, W., Dai, F., Diehl, J.A., Chen, M., Bai, J., 2023. Exploring the spatial pattern of community urban green spaces and COVID-19 risk in Wuhan based on a random forest model. Heliyon 9 (9).
Liu, C., Fan, H., Jiang, Y., Ma, R., Song, S., 2023. Gully erosion susceptibility assessment based on machine learning-A case study of watersheds in Tuquan County in the black soil region of Northeast China. Catena 222, 106798.
Liu, Y., Heuvelink, G.B., Bai, Z., He, P., Xu, X., Ding, W., Huang, S., 2021. Analysis of spatio-temporal variation of crop yield in China using stepwise multiple linear regression. Field Crop Res 264, 108098.
Liu, Y.X., Heuvelink, G.B., Zhanguo, B.A.I., Ping, H.E., Jiang, R., Huang, s. h., & xu, x. p., 2022. Statistical analysis of nitrogen use efficiency in Northeast China using multiple linear regression and Random Forest. J. Integr. Agric. 21 (12), 3637–3657.
Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30.
Makungwe, M., Chabala, L.M., Chishala, B.H., Lark, R.M., 2021. Performance of linear mixed models and random forests for spatial prediction of soil pH. Geoderma 397, 115079.
Milborrow, S. (2014). Notes on the earth package. Retrieved October, 31, 2017.
Miller, T., 2019. Explanation in artificial intelligence: Insights from the social sciences. Artif. Intell. 267, 1–38.
Mobarak Hassan, E., Fattahi, E., Habibi, M., 2023. Temporal and Spatial Variability of Dust in the Urmia Basin, 1990–2019. Atmos. 14 (12), 1761.
Moghtaderi, T., Aminiyan, M.M., Alamdar, R., Moghtaderi, M., 2019. Index-based evaluation of pollution characteristics and health risk of potentially toxic metals in schools dust of Shiraz megacity, SW Iran. Human Ecol. Risk Assess.: an International Journal 25 (1–2), 410–437.

Mohammadifar, A., Gholami, H., Comino, J.R., Collins, A.L., 2021. Assessment of the interpretability of data mining for the spatial modelling of water erosion using game theory. Catena 200, 105178.
Morgan, R.P.C., 2009. Soil erosion and conservation. John Wiley & Sons.
Motaghi, F.A., Hamzehpour, N., Abasiyan, S.M.A., Rahmati, M., 2020. The wind erodibility in the newly emerged surfaces of Urmia Playa Lake and adjacent agricultural lands and its determining factors. Catena 194, 104675.
Mozafari, M., Hosseini, Z., Fijani, E., Eskandari, R., Siahpoush, S., Ghader, F., 2022. Effects of climate change and human activity on lake drying in Bakhtegan Basin, southwest Iran. Sustain. Water Resour. Manage. 8 (4), 109.
Nunes, A.N., De Almeida, A.C., Coelho, C.O., 2011. Impacts of land use and cover type on runoff and soil erosion in a marginal area of Portugal. Appl. Geogr. 31 (2), 687–699.
Ogutu, J. O., Schulz-Streeck, T., & Piepho, H. P. (2012, December). Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. In *BMC proceedings* (Vol. 6, pp. 1-6). BioMed Central.
Oldeman, L.R., 1994. The global extent of land degradation. In: Greenland, D.J., Szabolcs, I. (Eds.), Land Resilience and Sustainable Land Use. CAB International, Wallingford UK, pp. 99–118.
Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., Pereira, J.M., 2012. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. For. Ecol. Manage. 275, 117–129.
Padarian, J., McBratney, A.B., Minasny, B., 2020. Game theory interpretation of digital soil mapping convolutional neural networks. Soil 6 (2), 389–397.
Parajuli, S.P., Yang, Z.L., Kocurek, G., 2014. Mapping erodibility in dust source regions based on geomorphology, meteorology, and remote sensing. J. Geophys. Res.: Earth Surface 119 (9), 1977–1994.
Parsa, A.B., Movahedi, A., Taghipour, H., Derrible, S., Mohammadian, A.K., 2020. Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. Accid. Anal. Prev. 136, 105405.
Pourhashemi, S., Asadi, M.A.Z., Boroughani, M., Azadi, H., 2023. Mapping of dust source susceptibility by remote sensing and machine learning techniques (case study: Iran-Iraq border). Environ. Sci. Pollut. Res. 30 (10), 27965–27979.
Pradhan, B., Lee, S., Dikshit, A., Kim, H., 2023a. Spatial flood susceptibility mapping using an explainable artificial intelligence (XAI) model. Geosci. Front. 14 (6), 101625.
Pradhan, B., Dikshit, A., Lee, S., Kim, H., 2023b. An explainable AI (XAI) model for landslide susceptibility modeling. Appl. Soft Comput. 142, 110324.
Qiu, Y., Fu, B.J., Wang, J., Chen, L., Meng, Q., Zhang, Y., 2010. Spatial prediction of soil moisture content using multiple-linear regressions in a gully catchment of the Loess Plateau, China. J. Arid Environ. 74 (2), 208–220.
Qu, L., Lu, H., Tian, Z., Schoorl, J.M., Huang, B., Liang, Y., Liang, Y., 2024. Spatial prediction of soil sand content at various sampling density based on geostatistical and machine learning algorithms in plain areas. Catena 234, 107572.
Quinlan, J. R. (1992). Learning with continuous classes. In proceedings AI, 92 (Adams & Sterling, Eds), 343-348.
Quinlan, J. R. (1993, June). Combining instance-based and model-based learning. In *Proceedings of the tenth international conference on machine learning* (pp. 236-243).
Rashki, A., Kaskaoutis, D.G., Goudie, A.S., Kahn, R.A., 2013. Dryness of ephemeral lakes and consequences for dust activity: the case of the Hamoun drainage basin, southeastern Iran. Sci. Total Environment 463, 552–564.
Rashki, A., Middleton, N.J., Goudie, A.S., 2021. Dust storms in Iran-Distribution, causes, frequencies and impacts. Aeolian Res. 48, 100655.
Renard, K.G., Freimund, J.R., 1994. Using monthly precipitation data to estimate the R-factor in the revised USLE. J. Hydrol. 157 (1–4), 287–306.
Roy, J., Saha, S., 2022. Ensemble hybrid machine learning methods for gully erosion susceptibility mapping: K-fold cross validation approach. Artificial Intelligence in Geosciences 3, 28–45.
Sajedipour, S., Zarei, H., Oryan, S., 2017. Estimation of environmental water requirements via an ecological approach: A case study of Bakhtegan Lake, Iran. Ecol. Eng. 100, 246–255.
Salmerón-Gómez, R., Rodríguez-Sánchez, A., García-García, C., 2020. Diagnosis and quantification of the non-essential collinearity. Comput. Stat. 35 (2), 647–666.
Shahraki, A.S., Shahraki, J., Monfared, S.A.H., 2017. Water resources management by simulation under virtual water scenario in agricultural sector, case Study: Hirmand catchment, Iran. Int. J. Agric. Manage. Dev. (IJAMAD) 9 (1047–2019-3463), 25–35.
Shirani, M., Afzali, K.N., Jahan, S., Strezov, V., Soleimani-Sardo, M., 2020. Pollution and contamination assessment of heavy metals in the sediments of Jazmurian playa in southeast Iran. Sci. Rep. 10 (1), 4775.
Song, Y., Zhang, C., Jin, X., Zhao, X., Huang, W., Sun, X., Wang, S., 2023. Spatial prediction of PM2. 5 concentration using hyper-parameter optimization XGBoost model in China. Environ. Technol. Innov. 32, 103272.
Szatmári, G., Pásztor, L., Laborczi, A., Illés, G., Bakacsi, Z., Zacháry, D., Jakab, G., 2023. Countrywide mapping and assessment of organic carbon saturation in the topsoil using machine learning-based pedotransfer function with uncertainty propagation. Catena 227, 107086.
Tang, Y., Deng, R., Liang, Y., Zhang, R., Cao, B., Liu, Y., Yu, J., 2023. Estimating high-spatial-resolution daily PM2. 5 mass concentration from satellite top-of-atmosphere reflectance based on an improved random forest model. Atmos. Environ. 302, 119724.
Tanner, S., Ben-Hur, M., Argaman, E., Katra, I., 2023. The effects of soil properties and aggregation on sensitivity to erosion by water and wind in two Mediterranean soils. Catena 221, 106787.
Tehrani, N.A., Farhanj, F., Janalipour, M., 2023. Introducing a novel dust source identification method based on edge points and paths extracted from integration of time-series MODIS products. Remote Sens. Appl.: Soc. Environ. 32, 101054.
Torghabeh, A.K., Pradhan, B., Jahandari, A., 2020. Assessment of geochemical and sedimentological characteristics of atmospheric dust in Shiraz, southwest Iran. Geosci. Front. 11 (3), 783–792.
Vaezi, A., Ghazban, F., Tavakoli, V., Routh, J., Beni, A.N., Bianchi, T.S., Kylin, H., 2019. A Late Pleistocene-Holocene multi-proxy record of climate variability in the Jazmurian playa, southeastern Iran. Palaeogeogr. Palaeoclimatol. Palaeoecol. 514, 754–767.
Vahidipour, M., Raeisi, E., van der Zee, S.E., 2022. Potentially toxic metals in sediments, lake water and groundwater of the Ramsar wetlands Bakhtegan-Tashk, South Iran: Distribution and source assessment. Environ. Technol. Innov. 28, 102789.
Vapnik, V., 1999. The nature of statistical learning theory. Springer science & business media.
Varol, M., Karakaya, G., Alpaslan, K., 2022. Water quality assessment of the Karasu River (Turkey) using various indices, multivariate statistics and APCS-MLR model. Chemosphere 308, 136415.
Wang, F., Sahana, M., Pahlevanzadeh, B., Pal, S.C., Shit, P.K., Piran, M.J., Mosavi, A., 2021. Applying different resampling strategies in machine learning models to predict head-cut gully erosion susceptibility. Alex. Eng. J. 60 (6), 5813–5829.
Wang, W., Samat, A., Abuduwaili, J., Ge, Y., De Maeyer, P., Van de Voorde, T., 2022. Temporal characterization of sand and dust storm activity and its climatic and terrestrial drivers in the Aral Sea region. Atmos. Res. 275, 106242.
Wu, N., Ge, Y., Abuduwaili, J., Issanova, G., Saparov, G., 2022. Insights into variations and potential long-range transport of atmospheric aerosols from the Aral Sea basin in Central Asia. Remote Sens. (Basel) 14 (13), 3201.
Yesilnacar, E.K., 2005. The application of computational intelligence to landslide susceptibility mapping in Turkey. University of Melbourne Department, 200.
Zaman, N.A.F.K., Kanniah, K.D., Kaskaoutis, D.G., Latif, M.T., 2021. Evaluation of machine learning models for estimating pm2. 5 concentrations across malaysia. Appl. Sci. 11 (16), 7326.
Zhang, Y., He, T., Liang, S., Zhao, Z., 2023. A framework for estimating actual evapotranspiration through spatial heterogeneity-based machine learning approaches. Agric Water Manag 289, 108499.
Zheng, C., Yuan, F., Luo, X., Li, X., Liu, P., Wen, M., Albanese, S., 2023. Mineral prospectivity mapping based on Support vector machine and Random Forest algorithm-A case study from Ashele copper-zinc deposit, Xinjiang, NW China. Ore Geol. Rev., 105567
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. r. Stat. Soc. Ser. B Stat Methodol. 67 (2), 301–320.
Zucca, C., Middleton, N., Kang, U., Liniger, H., 2021. Shrinking water bodies as hotspots of sand and dust storms: The role of land degradation and sustainable soil and water management. Catena 207, 105669.