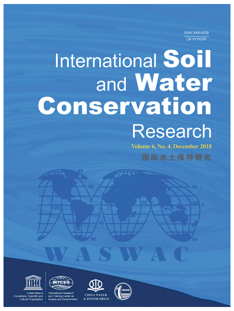
Journal Pre-proof



Gully erosion mapping susceptibility in a Mediterranean environment: A hybrid decision-making model

Sliman Hitouri, Mohajane Meriame, Ali Sk Ajim, Quevedo Renata Pacheco, Thong Nguyen-Huy, Pham Quoc Bao, Ismail ElKhrachy, Antonietta Varasano

PII: S2095-6339(23)00089-8

DOI: https://doi.org/10.1016/j.iswcr.2023.09.008

Reference: ISWCR 449

To appear in: *International Soil and Water Conservation Research*

Received Date: 1 March 2023

Revised Date: 25 September 2023

Accepted Date: 26 September 2023

Gully erosion mapping susceptibility in a Mediterranean environment: a hybrid decision-making model

Sliman Hitouri[1], Mohajane Meriame[2,*], Sk Ajim Ali[3], Renata Pacheco Quevedo[4], Thong Nguyen-Huy[5], Quoc Bao Pham[6], Ismail ElKhrachy[7], and Antonietta Varasano[8]

[1]Geosciences Laboratory, Department of Geology, Faculty of Sciences, University Ibn Tofail, Kenitra 14000, Morocco

[2]Construction Technologies Institute, National Research Council of Italy, Polo Tecnologico di San Giovanni a Teduccio, 80146 Napoli, Italy

[3]Department of Geography, Faculty of Science, Aligarh Muslim University (AMU), Aligarh, UP 202002, India

[4]Earth Observation and Geoinformatics Division, National Institute for Space Research (INPE), Sao Jose dos Campos, Sao Paulo, 12227010, Brazil

[5]Centre for Applied Climate Sciences, University of Southern Queensland, Toowoomba, 4350, QLD, Australia

[6]Faculty of Natural Sciences, Institute of Earth Sciences, University of Silesia in Katowice, Będzińska street 60, 41-200, Sosnowiec, Poland

[7]College of Engineering, Civil Engineering Department, Najran University, Najran 66291, Saudi Arabia

[8]ITC-CNR, Construction Technologies Institute, National Research Council, 70124 Bari, Italy

*Correspondence: mohajane@itc.cnr.it

Gully erosion mapping susceptibility in a Mediterranean environment: a hybrid decision-making model

Abstract

Gully erosion is one of the main natural hazards, especially in arid and semi-arid regions, destroying ecosystem service and human well-being. Thus, gully erosion susceptibility maps (GESM) are urgently needed for identifying priority areas on which appropriate measurements should be considered. Here, we proposed four new hybrid Machine learning models, namely weight of evidence -Multilayer Perceptron (MLP- WoE), weight of evidence –K Nearest neighbors (KNN- WoE), weight of evidence - Logistic regression (LR- WoE), and weight of evidence - Random Forest (RF- WoE), for mapping gully erosion exploring the opportunities of GIS tools and Remote sensing techniques in the El Ouaar watershed located in the Souss plain in Morocco. Inputs of the developed models are composed of the dependent (i.e., gully erosion points) and a set of independent variables. In this study, a total of 314 gully erosion points were randomly split into 70% for the training stage (220 gullies) and 30 % for the validation stage (94 gullies) sets were identified in the study area. 12 conditioning variables including elevation, slope, plane curvature, rainfall, distance to road, distance to stream, distance to fault, TWI, lithology, NDVI, and LU/LC were used based on their importance for gully erosion susceptibility mapping. We evaluate the performance of the above models based on the following statistical metrics: Accuracy, precision, and Area under curve (AUC) values of receiver operating characteristics (ROC). The results indicate the RF- WoE model showed good accuracy with (AUC = 0.8), followed by KNN-WoE (AUC = 0.796), then MLP-WoE (AUC = 0.729) and LR-WoE (AUC = 0.655), respectively. Gully erosion susceptibility maps provide information and valuable tool for decision-makers and planners to identify areas where urgent and appropriate interventions should be applied.

1. Introduction

Understanding how to use natural resources is essential to the existence of human communities. In addition to supporting basic human requirements like food, clean water, and air, soils are an important transporter for biodiversity. The depletion of natural resources, particularly soil, is one of the major issues of the modern era that has emerged in the past ten years (Turner et al., 2016; Wassie, 2020). Soil degradation is caused by population increase and resource extraction, which endangers human lives and property (Gomiero, 2016; Scherr, 2000). Soil erosion may impact soil productivity, surface water sources, their quality, ecological balance, and landscape (Bilotta et al., 2007; Issaka and Ashraf, 2017). Preventing land degradation proves to be challenging. Among the various types of soil erosion, gully erosion stands out as one of the most complex and hazardous forms, given its capacity to displace substantial amounts of soil. A gully is characterized as a deep, relatively permanent canal with vertical walls on either side that allow passing water currents for a short period. Gully erosion occurs when rushing surface water erodes a deep channel, removing and transporting the eroded surface soil (Ghorbanzadeh et al., 2020a). Over time, these gullies cause soil erosion, alter the surrounding environment, and accelerate the sedimentation of rivers and dams (Belayneh et al., 2020; Ghorbanzadeh et al., 2020a; Hancock and Evans, 2010). One of the most important techniques for managing this phenomenon is understanding the variables influencing the incidence of this form of erosion and its zoning. Gully erosion affects the environment in two ways: first, by eroding the surface and diminishing and reducing soil horizons, leading to high sediment production and bedding degradation; and second, by escalating surface discharge and decreasing groundwater nutrition.

In Morocco, gully erosion is one of the most significant environmental problems increasingly posing a threat to the country (Azedou et al., 2021; d’Oleire-Oltmanns et al., 2014; Meliho et al., 2018). There are several different types of soil erosion impacting half of Morocco's 20 million-hectare watersheds, such as sheet, rill, splash, and gully, which result in around 100 million tonnes of soil annually (Mosaid et al., 2022; Tairi et al., 2021). The gully is the most harmful type of erosion, causing several times more damage than other types of erosion, such as sedimentation of dams, destruction of energy and transportation transmission lines, loss of farmland productivity, land degradation, and long-term adverse economic effects (Belasri and Lakhouili, 2016; Bouslihim, 2020; Meliho et al., 2018). Even though there are anthropogenic causes for the generation of gullies, it is sped up by variables including climate change, geologic conditions, and soil characteristics (Nir et al., 2021; Poesen et al., 2003). In this sense, the framework of watershed management includes mapping and monitoring of regions susceptible to gully erosion.

Many conventional and numerical techniques were used for gully susceptibility mapping, by linking gully occurrence and conditioning factors (Jaafari et al., 2022; Rahmati et al., 2017). Field survey and data collection, although effective in mapping and evaluating gully erosion, are characterized by their time-consuming and labour-intensive nature, besides cannot forecast the spatial development of gully erosion (Jiang et al., 2021). As an alternative, the Water Erosion Prediction Project (Ghorbanzadeh et al., 2020b) and the European Soil Erosion Model (Quarteroni and Veneziani, 2003) applied physically-built models to estimate gully erosion. These models are less suitable for regional-scale study since they need extensive data and labor-intensive calibration procedures (Momm et al., 2012; Yuan et al., 2020). In addition, these models are adequate for numerically estimating the amount of gully erosion, however, they are less appropriate for gully erosion susceptibility mapping (Garosi et al., 2018; Rahmati et al., 2016).

The generation of gully erosion susceptibility maps currently uses a variety of probabilistic, knowledge-driven, and machine learning methods, including bivariate statistics (Meliho et al., 2018), weights-of-evidence (Shit et al., 2020), logistic regression (Conoscenti et al., 2014), information value (Paul and Saha, 2019), random forest (Avand et al., 2019), bivariate statistical models (Lana et al., 2022), maximum entropy (Azareh et al., 2019), frequency ratio (Amare et al., 2021), artificial neural network (Gafurov and Yermolayev, 2020), Functional tree (Tien Bui et al., 2019), Naïve Bayes tree (Hosseinalizadeh et al., 2019), support vector machine (Karami et al., 2015), and boosted regression trees (Arabameri et al., 2019).

When there is insufficient data about the intensity and distribution of a phenomenon, such as gully erosion, GIS-based multi-criteria decision analysis (MCDA) models can be useful. The analytical hierarchy process (AHP) and analytical network process (ANP), two qualitative (knowledge-based) MCDA, have been applied to gully susceptibility mapping in various study areas (Arabameri et al., 2019, 2018a; Chakrabortty and Pal, 2023; Choubin et al., 2019; Nhu et al., 2020). Although these models appear to offer solutions for environmental susceptibility mapping, their major limitation is the uncertainty associated with the experts' assessments, which can occasionally result in inaccurate conclusions (Ghorbanzadeh et al., 2020b).

Machine Learning is a cutting-edge method for anticipating gully erosion as well as managing and minimising the harm this phenomenon causes (Chakrabortty and Pal, 2023). The use of Machine Learning algorithms in studies of natural hazards, such as floods, wildfires, sinkholes, droughts, earthquakes, land subsidence, groundwater, landslides, and gullies, has significantly advanced (Abu El-Magd et al., 2021; Ali et al., 2022, 2021, 2020; Ghorbanzadeh et al., 2019; Hitouri et al., 2022; Pham et al., 2021). There are some advantages of using machine learning algorithms for gully susceptibility mapping, such as being non-parametric. Researchers have applied tree-based machine learning techniques for gully erosion modelling, which outperformed traditional techniques in terms of performance and accuracy (Mohsin et al., 2022). The overfitting issue in these tree-based algorithms is quite minimal when compared to numerical models (Ajit, 2016). Moreover, backpropagation is a supervised learning method that is used by MLP during training. MLP differs from a linear perceptron due to its numerous layers and non-linear activation (Pham et al., 2022). One or more secret layers can be found in an MLP (apart from one input and one output layer). A multi-layer perceptron can learn non-linear functions in addition to linear functions, whereas a single-layer perceptron can only learn linear functions (Parvin et al., 2022). LR is more straightforward to use, comprehend, and train than other methods (Davis et al., 2016). Fitting the line values to the sigmoid curve is the goal of LR (Yin et al., 2020). The KNN method has the benefits of being flexible to different proximity calculations, being relatively intuitive, and using a memory-based approach (Merghadi et al., 2020).

In this sense, the main aim of this study was to present four new hybrid Machine Learning models for mapping gully erosion in the province of Taroudant, located in the Souss plain of Morocco, namely: i) weight of evidence - Multilayer Perceptron (MLP- WoE); ii) weight of evidence –K Nearest neighbors (KNN- WoE); iii) weight of evidence - Logistic regression (LR-WoE); iv) weight of evidence - Random Forest (RF- WoE). These ensemble models are a novel method that has not been used for gully erosion susceptibility in this area before. The RF, MLP, LR, and KNN algorithms were considered taking into account their advantages. In addition , the study integrated the WOE with the four machine learning algorithms due to its powerful ability to transform and select variables, and reveals the predictive ability of an independent variable in relation to the dependent variable (Elmoulat and Ait Brahim, 2018; Shafizadeh-Moghadam et al., 2017). For that, we followed these steps, (1) using multi-collinearity analysis to identify significant gully erosion conditioning factors, (2) creating the hybrid machine learning models to predict gully erosion susceptibility, (3) employing the k-fold cross-validation (CV) method to mitigate the negative effects of randomness on the results, and (4) assessing the capability and robustness of the four hybrid models by comparing their performance using the Receiver Operating characteristic Curve (ROC).

While there have been many recent advances and applications of Machine Learning techniques for gully erosion mapping studies in various study areas worldwide, their applicability in regions with limited ground-based data or inadequate data quality remains uncertain. To fill this gap, our study contributes to the literature by emphasizing the importance of using freely available data sources to identify and map gully erosion susceptibility in this watershed.

2. Study area

El Ouaar watershed is located in the province of Taroudant, Morocco. It is limited between longitudes (8°43’30’’ W - 8°56’30’’W), and latitudes (30°28’00’’N - 30°50′ 00’’N). El Ouaar watershed covers an area of 395.18 km^2 and is characterized by an arid and semi-arid climate. From a topographical point of view, the study area shows an altitude ranging between 214, located in the south of the basin, and an altitude of 3353 m, located in the north, with an average elevation of 1657.5 m and an average slope of 19° (Fig.1).The annual rainfall in the area varies significantly, ranging from 207 mm to 625 mm during the winter season, while temperatures tend to be cooler, averaging around 6.4 °C . In August, temperatures sometimes reach 45°C (Dijon, 1966).

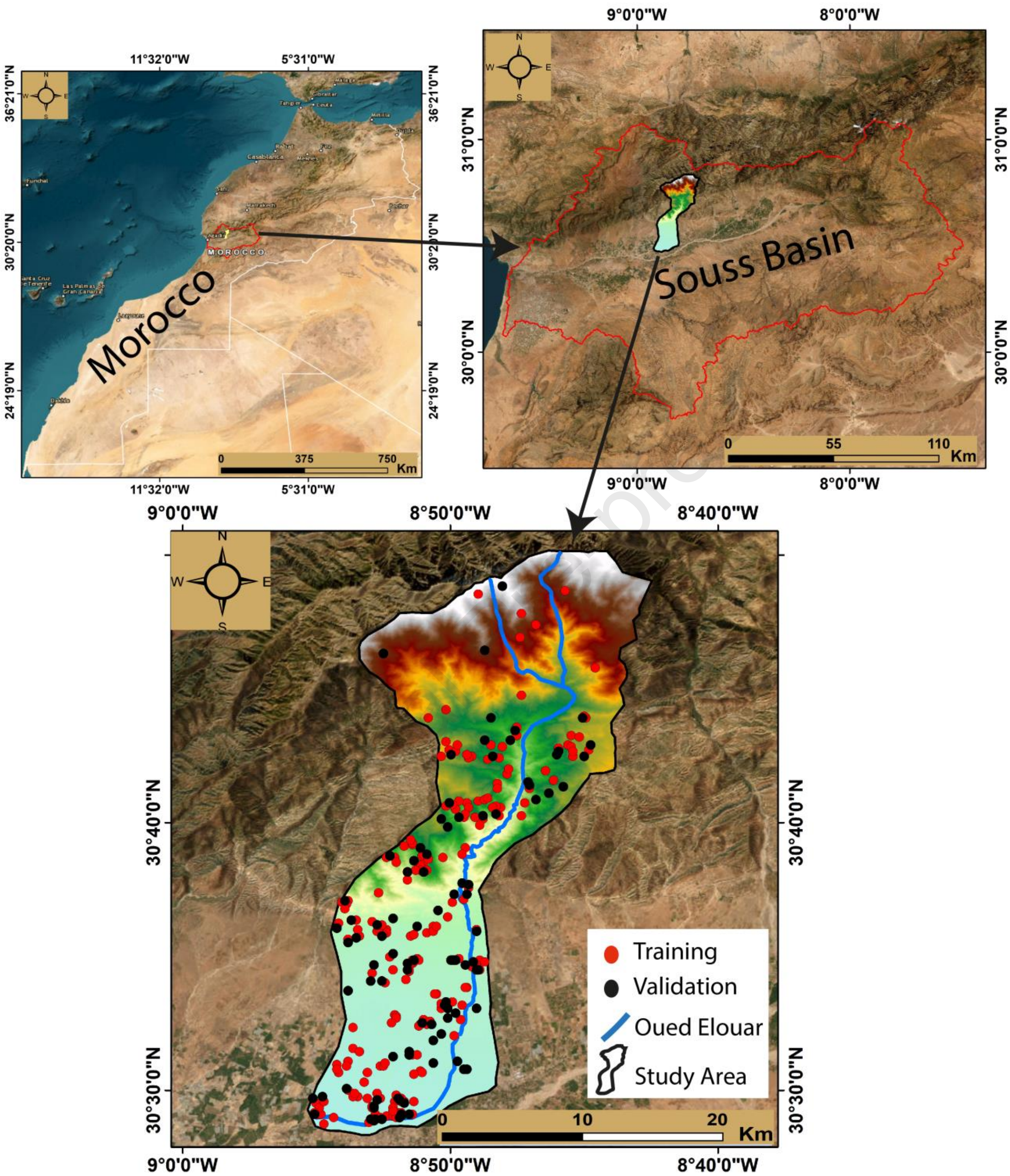


Fig. 1 Location of the study area.

Geologically, El Ouaar watershed is characterized by a silty and clay lithological terrain of Quaternary age. The north part of this study area is presented by the Cretaceous basement containing carbonate lithology dominated by dolomite and limestone. The study area is a part of the Souss basin, which is limited to the north by the High Atlas ranges and to the south by the Anti-Atlasic ranges. It is characterized by quaternary lithological units, in the north of this basin where the limestone and dolomite deposits of Cretaceous are very dominant and the schistose Paleozoic deposits. In the center of this basin at the level of the plain of Souss there are only the quaternary deposits not very compact, especially the silts and the clays. The quaternary formations are constituted by four lithostratigraphic units (Figs. 2 and 3).

- The first 2 m of a basal unit (U1), whitish, massive, consisting of silt-sandstone encrusted and more consolidated than the overlying deposits. U1 is most likely thicker at depth;
- 2 m of a sandy-limono-conglomeratic unit (U2), made up of grano-decreasing sequences where the conglomerates are polygenic lenticular, with a more or less friable sandy matrix. 2.5 m of a unit (U3) made up of red-brown sandy-clayey silts with small pebbles and rare microconglomeratic pockets.
- 0.5 to 1 m of an alluvial unit (U4) consisting of polygenic conglomerate with a silty-sandy matrix. The elements are well blunted and rounded, where their size varies from a few millimeters to a few centimeters and are of varied nature: limestone, sandstone, and magmatic rocks. The north of the El Ouaar watershed, the Western High Atlas is characterized by a Paleozoic terrain rich in sandstone and shale, and the Cetacean, contains carbonate rocks especially limestone and dolomite (Ambroggi, 1963).

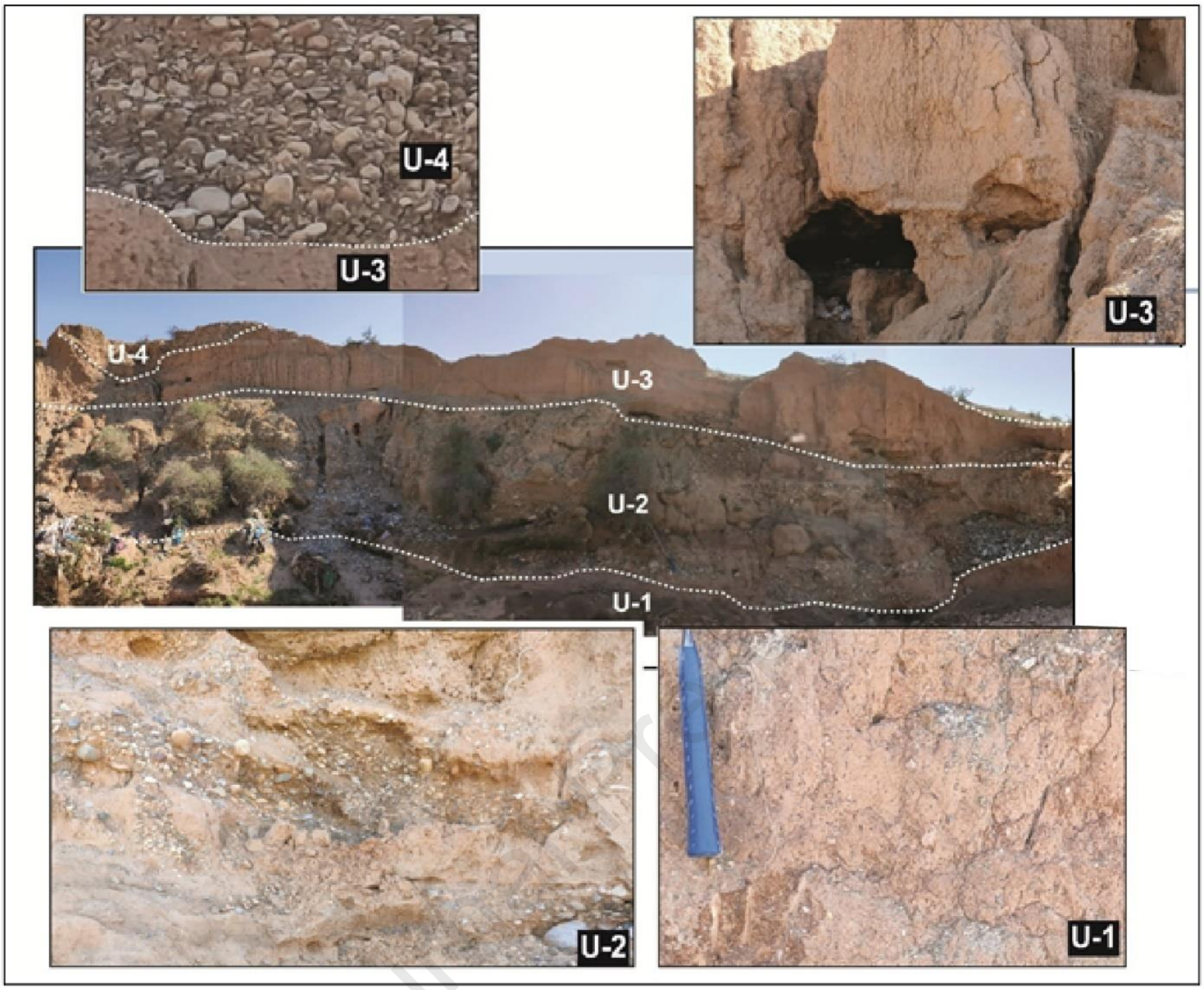


Fig. 2 : Quaternary lithostratigraphic deposits in the study area.

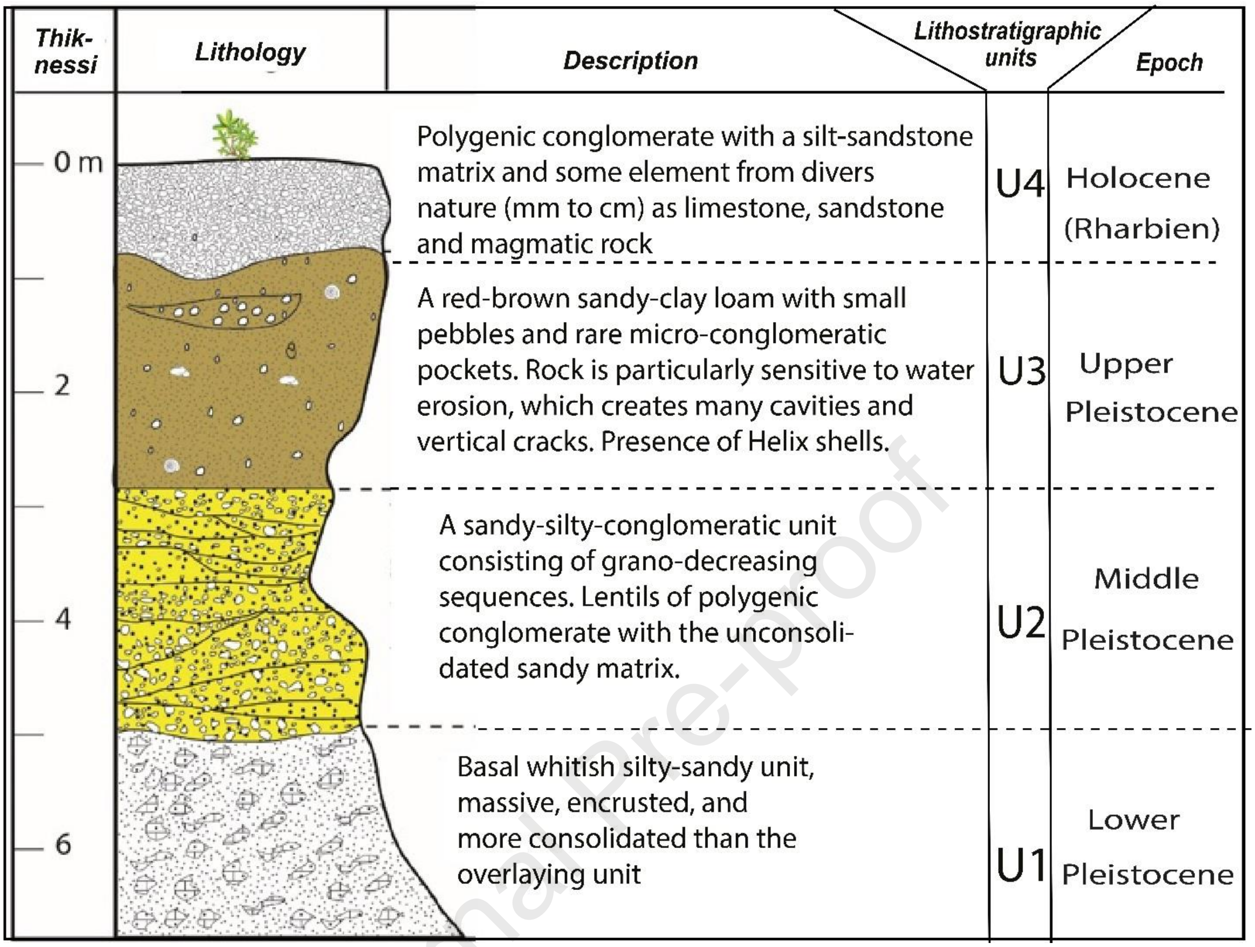


Fig. 3: Quaternary lithostratigraphic deposits in the study area.

3. Data and Methodology

The procedure followed in this is divided into three main parts: data collection and processing, ML implementation, and model performance assessment. All ML models elaborated in this study were developed in an R programming environment and GESM was reclassified into five classes: “very low”, “low”, “moderate”, “high” and “very high” susceptibility, using the natural breaks method in ArcGIS software. The methodology of the present work is presented in Fig. 4.

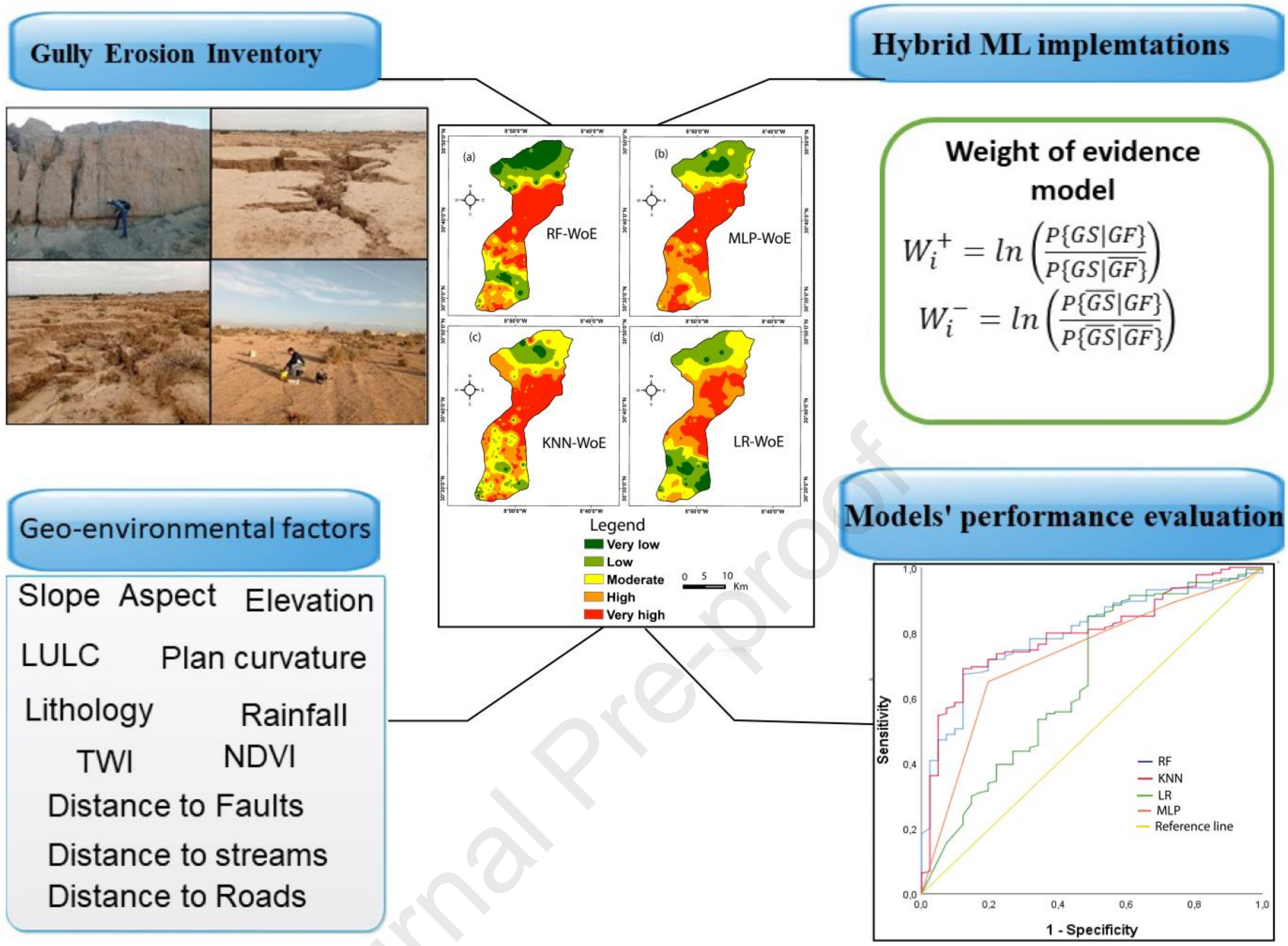


Fig. 4: The methodology adopted in this study.

3.1. Inventory of gully erosion locations

Gully erosion inventory is a primary and crucial step of gully erosion mapping. In this study, gully erosion points were collected from a variety of sources including field data and high-resolution aerial images in Google Earth. During the field survey, gully points were collected with their geographical coordinates using Global Positioning System (GPS) tools. The gully points in the study area showed that the width of erosion can reach 4 m, with a depth varying from 0.5 to 2m, and sometimes can reach 3m, especially in areas near the Wadi El Ouaar. Infrastructures and natural resources (e.g., roads, schools, and agricultures areas) are strongly influenced by water erosion in this region (Fig. 5). Gully erosion points were randomly divided into training and validation datasets in the ratio of 70 % (220 points) and 30 % (94 points) for models implementations. Also, a total of 314 non-gully erosion points were collected and randomly divided into 70% (220 points) and 30% (94 points) for training and testing, respectively. The gully erosion locations were assigned the value “1” and the non-gully erosion locations were assigned the value “0”.

Fig. 5: Gully erosion in El Ouaar watershed from a field survey.

3.2. Gully erosion conditioning factors

Determination of environmental factors is the first step in gully erosion susceptibility modelling for identifying important factors that contribute to gully occurrence in a given terrain (Rahmati et al., 2017). In this study, 12 geo-environmental factors were selected, these include: Elevation, Plan curvature, Aspect, Slope, Rainfall, Lithology, Land use/land cover, NDVI, Distance to roads, Distance to streams, Distance to faults, Topographic Wetness Index (TWI) (See Table 1 for details).

Table 1 Details of thematic data layers and data sources used in this study

Data	Data types in GIS	Scale	Source
Erosion inventory	Polygon	_	Google earth and field data
Elevation	Grid	30×30 m	DEM 30 m, from https://earthexplorer.usgs.gov/ (accessed on 20 August 2021)
Aspect	Grid	30×30 m	DEM 30 m, from https://earthexplorer.usgs.gov/ (accessed on 20 August 2021)
Slope	Grid	30×30 m	DEM 30 m, from https://earthexplorer.usgs.gov/ (accessed on 20 August 2021)
Plan curvature	Grid	30×30 m	DEM 30 m, from https://earthexplorer.usgs.gov/ (accessed on 20 August 2021)
TWI	Grid	30×30 m	DEM 30 m, from https://earthexplorer.usgs.gov/ (accessed on 20 August 2021)
Rainfall	Grid	30×30 m	ERA-Interim, from https://apps.ecmwf.int/datasets(accessed on 18 July 2021)
NDVI	Grid	30×30 m	Landsat-8-OLI image, from https://earthexplorer.usgs.gov/ (accessed on 12 July 2021)
Lithology	Polygon	-	Geological map of Morocco at a scale of 1:000 000
Faults	Polygon	-	Geological map of Morocco at a scale of 1:000 000
Roads	Polygon	-	https://www.geojamal.com
Streams	Polygon	-	https://geossc.ma
Land use land cover	Polygon	-	Landsat-8-OLI image, from https://earthexplorer.usgs.gov/

3.2.1. Elevation

Elevation is an important factor in the evolution of susceptibility to gully erosion, based on the occurrence and development of the gully erosion (Zabihi et al., 2018), because it affects vegetation, precipitation, and gully erosion (Golestani et al., 2014). This factor was reclassified into three classes: 1801-3353 m; 792-1801 m and 214 -792 m (Fig.6a).

3.2.2. Slope angle

Slope runoff and surface drainage contribute to erosion (Ghorbani Nejad et al., 2017). It is considered to be an important predictor of gully erosion processes (Conforti et al., 2011; Lucà et al., 2011). This factor was reclassified into five classes: 0-7.23%, 7.23-15.75%, 15.75-24.79%, 24.79 -35.13% and 35.13-65.86% (Fig. 6b).

3.2.3. Aspect

Aspect is an important conditioning factor in gully erosion mapping, it determines the direction of the slope in the basin. In this study, this factor is extracted from the DEM of Morocco. It is defined by the following equation (Zhou and Liu, 2004).

$$Aspect = 270° + arctan\left\{\frac{\mathcal{F}_y}{\mathcal{F}_x}\right\} - 90°\frac{\mathcal{F}_x}{\mathcal{F}_y} \quad (1)$$

$$\mathcal{F}_x = \frac{Z_8 - Z_2}{2\omega} \quad (2)$$

$$\mathcal{F}_y = \frac{Z_6 - Z_4}{2\omega} \quad (3)$$

Where Z_1 to Z_9 are cells of the 30 × 30 moving window and W is the grid resolution. The aspect factor map shows nine classes: Flat (F), North (N), Northeast (NE), East (E), South (S), Southwest (SW), West (W), and Northwest (NW) (Fig. 6c).

3.2.4. Plan curvature

Plan curvature contributes to the divergence or convergence of the water distribution and is generally defined as the curvature of the contour line that is formed by the intersection of a horizontal plan and the surface (Hitouri et al., 2022; Rahmati et al., 2022). The negative value represents the concave area, the positive value refers to the convex area and the zero value indicates the flat area (Fig. 6d).

3.2.5. Distance to road

Roads facilitate transportation and removal of eroded upland matter (Conoscenti et al., 2014). The road distance map was extracted from the road network map of Morocco, using the Euclidean distance tool available in ArcGIS software (version 10.8). It was subdivided into five classes: 1 – 1,308m; 1,308- 2,956m; 2,956- 4,894m; 4,894 -7,462 and 7,462 -12,357 m (Fig. 7a).

3.2.6. Distance to stream

This factor allows us to study the influence of the watercourses on gully erosion. It has an impact on erosion activities and also influences the wetting capacity of the surface. The values of this factor are classified into five categories: 0–1,229m; 1,229–2,683m; 2,683–4,322m; 4,322–6,297m and 6,297–9,502m (Fig. 7b).

3.2.7. Distance to faults

This factor is based on the geological structure of the study area. It was extracted from the geological map of Morocco with a scale of 1,000,000 and from the faults detected during the mission fields and from the interpretation of geophysical data. It is characterized by values classified into five categories: 0–2,184 m; 2,184–4,817m; 4,817–7,561m; 7,561–10,586m and 10,586–14,283m (Fig. 7c).

3.2.8. Rainfall

Rainfall determines the probability of gully occurrence in a given area. It represents the climate conditions of a study area (Roy and Saha, 2019).The annual average rainfall of the study area is 416 mm. We used the Inverse Distance-Weighted (IDW) interpolation method for preparing the rainfall map of El Ouaar watershed. Rainfall data used in this study were downloaded from (https://apps.ecmwf.int/datasets, accessed on 18 July 2021) and classified into five classes: 207–273 mm, 273–363 mm, 363–443 mm, 443–525 mm, and 525–625 mm (Fig. 7d).

3.2.9. TWI

The topographic wetness index calculates the quantity of water in the study area, which contributes to gully erosion (Moore and Wilson, 1992). It is defined by applying the following equation (Moore et al., 1991):

$$TWI = ln(As/tan\beta) \tag{4}$$

Where AS is the basin area and β is the slope gradient in degrees.

In this study, TWI was classified into five classes: -6.50 – (-5.53); -5.53 –1.55; 1.55 – 4.76, and 4.76 –12.51 (Fig. 8a).

3.2.10. Lithology

The lithology plays an important role in erosion; it is considered a fundamental variable for mapping the susceptibility of dust sources and terrain. It allows us determining the source areas with low hardness compared to other resistant units as well as the nature and types of soil (Sissakian et al., 2013). The lithology layer of this area study was prepared by digitizing the geological map of Morocco at a 1:1000, 000 scale, and field data.

The study area is composed of nine geological formations: A) Upper Pleistocene and Holocene, B) Middle and Upper Miocene, C) Phosphate Eocene, D) High Cretaceous, E) Upper Cretaceous phosphate facies, F) Middle Cretaceous, G) Granites and granodiorites (Tichka and Jbilet), H) Ordovician, I) Cambrien (Fig. 8b).

3.2.11. Land use / Land cover

This factor controls the occurrence of gullies, depending on the type of land use/ Land cover (Band et al., 2020). It indicates a negative correlation between erosion rate and vegetation density (Hughes et al., 2001). In this study, one Landsat 8 Operational Land Imager (OLI) satellite image acquired on 12 July 2021, downloaded from the United States Geological Survey website (USGS) was used for land cover / land use information. Therefore, the radiometric and atmospheric corrections are performed based on the Dark Object Subtraction (DOS) algorithm in ENVI 5.2 software. After, the land cover classification process was applied using Support Vector Machine (SVM) supervised classifier and five land cover / land use (LULC) classes were identified, namely, water, greenhouses, barelands, construction/buildings, and agricultural land (Fig. 8c).

A total of 150 training sample points, i.e., 30 samples per LULC class, were done by visual and manual on-screen digitizing based on our expert knowledge of the study area and high-resolution imagery from Google Earth. The generated land cover output achieved an overall accuracy of 95.32%.

3.2.12. NDVI

The Normalized Difference Vegetation Index (NDVI) represents a good indicator of photosynthetic activity (Pourghasemi et al., 2014). In our study, NDVI was calculated using one Landsat 8 OLI satellite image acquired on 12 July 2021 downloaded from the United States Geological Survey website (USGS) website following equation 5. NDVI was calculated and reclassified into 5 classes using ArcGIS 10.8.

$$\mathrm{NDVI} = \frac{\mathrm{NIR}-\mathrm{Red}}{\mathrm{NIR}+\mathrm{Red}} \quad (5)$$

Where NIR and Red values represent the infrared and red portion of the electromagnetic spectrum respectively. For Landsat 8 OLI image, the NIR and RED bands are band 5 (0.85–0.88 µm) and band 4 (0.64–0.67 µm), respectively. After, NDVI values were reclassified into 5 classes: (-0.30-0.12), (0.12-0.17), (0.17-0.25), (0.25-0.36), and (0.36-0.71) (Fig. 8d).

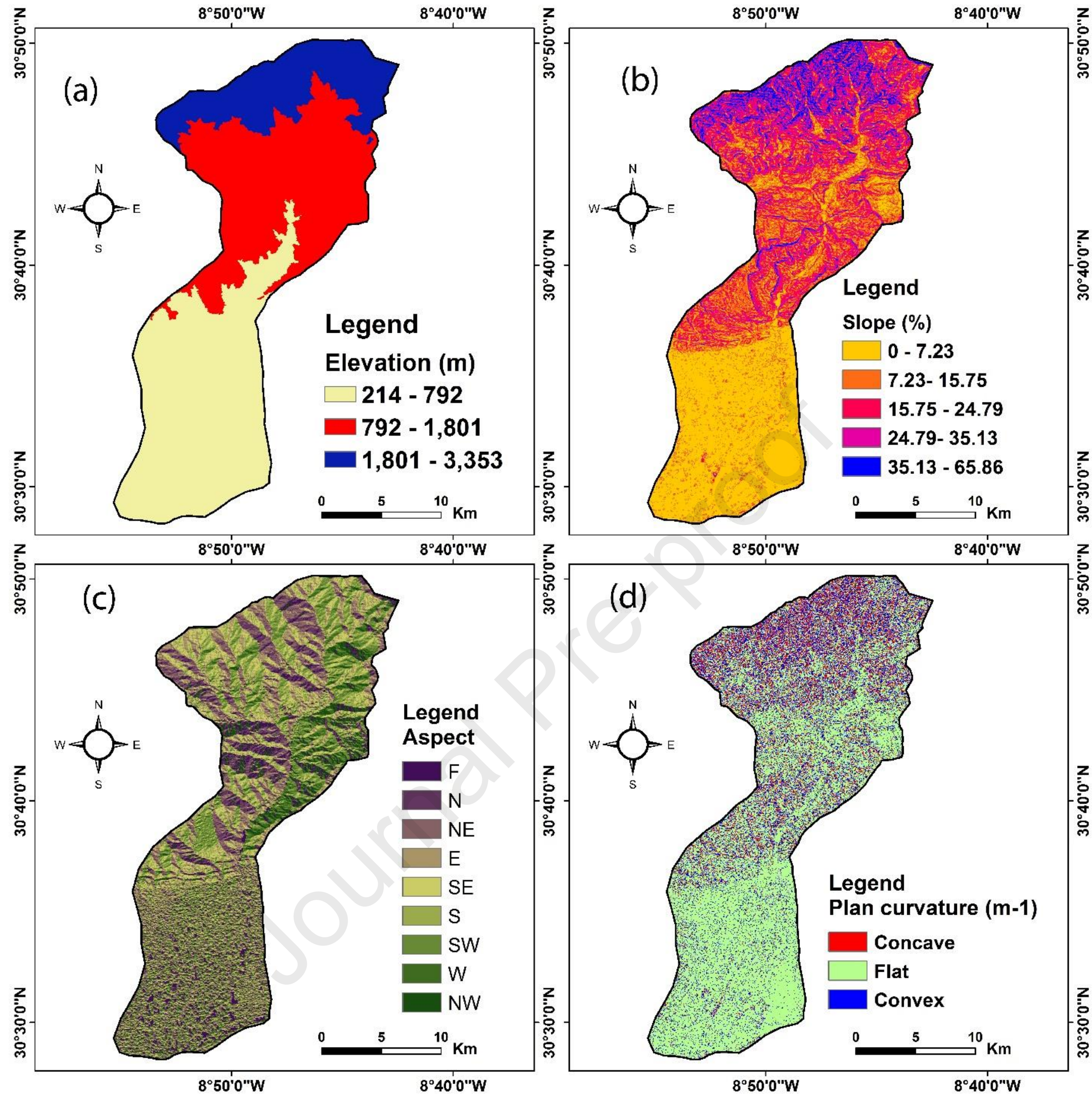


Fig. 6: Gully erosion conditioning factors: (a) Elevation, (b) Slope, (c) Aspect , and (d) Plan curvature.

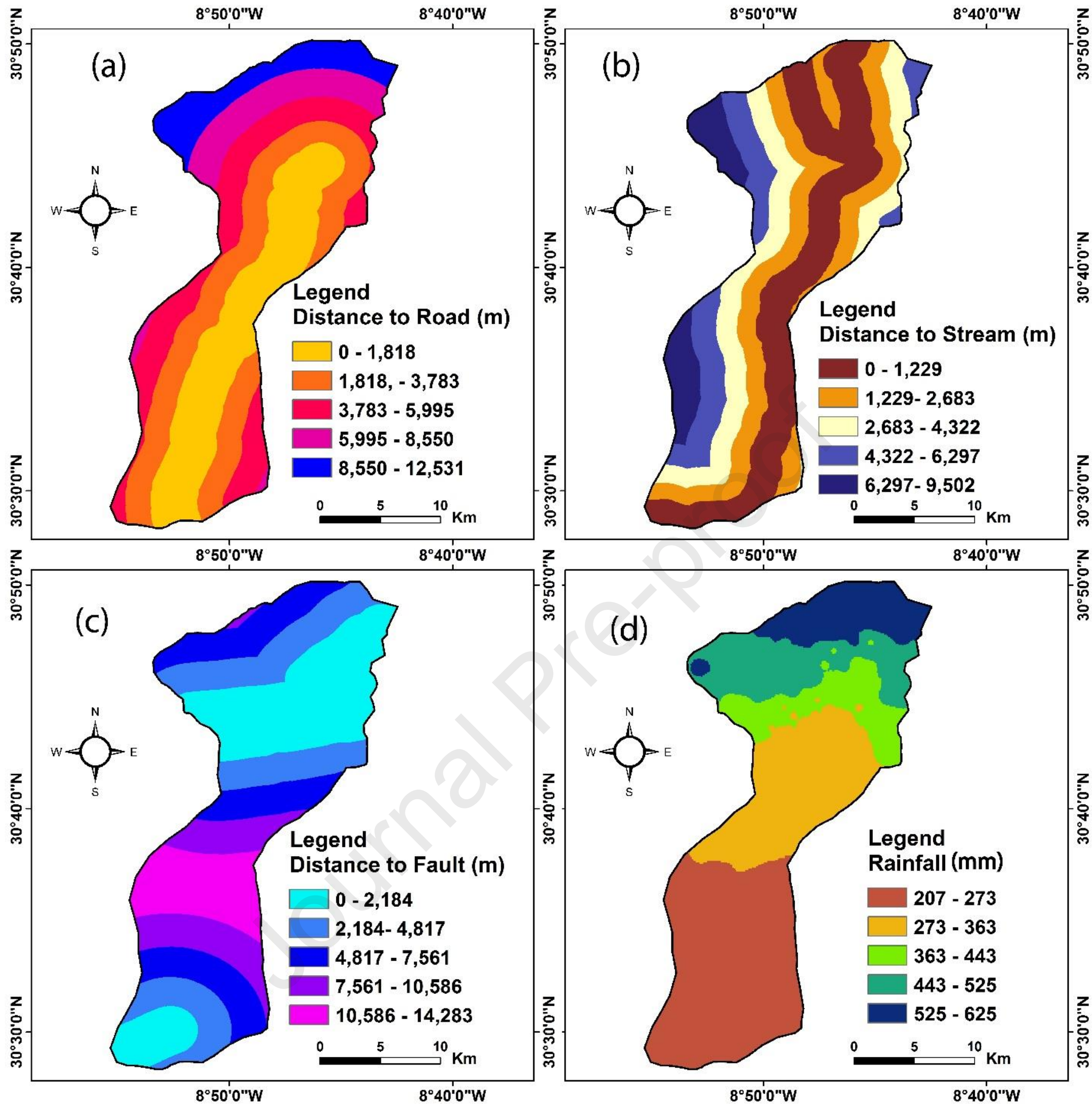


Fig. 7: Gully erosion conditioning factors: (a) Distance to road, (b) Distance to stream, (c) Distance to Faults and (d) Rainfall.

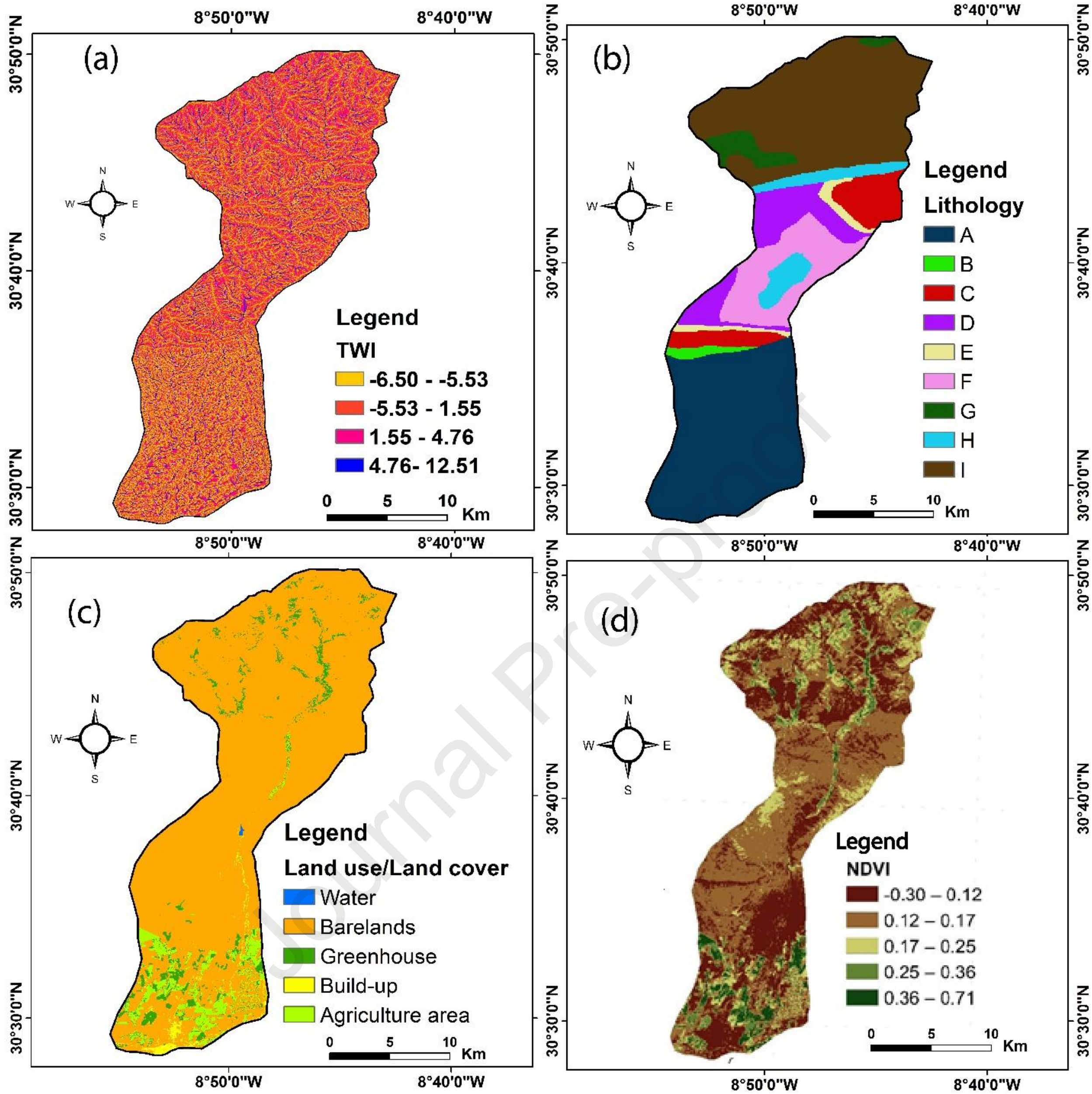


Fig. 8: Gully erosion conditioning factors: (a) TWI, (b) Lithology, (c) Land use/Land cover , and (d) NDVI.

4. Modelling process

4.1. Multi-Layer Perceptron Neural Network (MLP NN)

The MLP NN is an artificial neural network algorithm, widely used for classification approaches (Roy and Saha, 2021). It consists of an input layer, hidden layer, and output layer. The hidden layers process the data, while the output layers provide the classification results (Paola and Schowengerdt, 1995). Connection weights between neurons are updated (Oliveira et al., 2015). The main advantage of MLP is the non-dependency of prior assumptions of data distribution (Gardner and Dorling, 1998). In this study, we considered 30 neurons and 2 hidden layers, the Linear Unit Rectification (Relu) activation function, and the Adam optimizer (Adaptive Moment Optimization) developed by (Kingma and Ba, 2017).

4.2. Logistic Regression (LR)

LR is a multivariate statistical model, used for fitting Bernoulli distributions (Arabameri et al., 2018b). Unlike linear regression, logistic regression outcomes are binary or dichotomous (Hosmer et al., 2000). The model describes the relationship between dependent and independent variables, such as the presence or absence of gully erosion and the conditioning factors (Lucà et al., 2011).

4.3. K-Nearest Neighbours (KNN)

The KNN is a non-parametric model, considered one of the simplest machine learning algorithms (Zhang et al., 2018). The classification is based on the nearest neighbours. The number of neighbours (K) should be defined, which is used for the voting process (Abraham et al., 2021). The output consists of class membership likelihoods, according to Euclidean distance (Avand et al., 2019; Hussain et al., 2022). The most voted class is assigned to the analyzed data point. In this study, we select 8 nearest neighbours (K=8).

4.4. Random Forest (RF)

RF is a non-parametric ensemble learning algorithm that combines multiple decision tree models (Breiman, 2001). It randomly separated the input data into subsets for each internal decision tree (Quevedo et al., 2021). This study used the regression approach to generate numeric outcomes, for gully erosion susceptibility. The result is obtained by averaging the prediction of trees. RF also calculates the variable importance using mean decrease accuracy and mean decrease Gini index (Hitouri et al., 2022). In this study, 200 trees and 2 variables were selected for the main node split.

4.5. Weight of Evidence

The Weight of Evidence is a bivariate statistical test based on Bayesian probability (Bonham-Carter et al., 1988) that estimates the relative importance of each conditioning factor (Saha et al., 2020; Yang et al., 2021), using prior and posterior probability. The prior probability of gully erosion occurrence considers the number of pixels containing gully erosion and the total number of pixels in the study area (Pradhan et al., 2010). Then, positive and negative weights are calculated to identify the relationship between gully erosion conditioning factors and gully erosion occurrence. Finally, we calculate the standardized value of the difference to estimate the posterior probability relative certainty (Chen et al., 2018).

4.6. Modelling evaluation

The model performance assessment is used to determine and select the appropriate model for environmental hazards modelling (Chu et al., 2019; Lin and Chen, 2012; Pham et al., 2020). In this study, several statistical metrics widely used in previous studies were considered, including; AUC, specificity, sensitivity, and accuracy.

The ROC curve area (AUC) measures the performance of machine learning models. AUC values were classified into four precision categories, which are comprised between 0 and 1: poor (AUC = 0.6 to 0.7), fair (AUC = 0.7 to 0.8), good (AUC = 0.8–0.9), and excellent (AUC = 0.9–1) (Fressard et al., 2014). High values indicate a strong model, while low values mean a weak model (Hong et al., 2017). Overall accuracy (OA) represents the probability of occurrence of correctly classified pixels. It is calculated by the sum of true positive and true negative divided by all available singular tests (equation 7). Precision is used to measure the quality of the results. It is calculated by dividing the true positive by the sum of the true positive and false positive (equation 8). Sensitivity is calculated by dividing the true-negative values by the sum of true negatives and false positives (equation 9) (Huang et al., 2023). Specificity represents the proportion of gully erosion pixels correctly predicted as gully erosion (equation 10).

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (6)$$

$$Precision = TP/(TP + FP) \quad (7)$$

$$Sensitivity = TP/(TP + TN) \quad (8)$$

$$Specificity = TN/(TN + FP) \quad (9)$$

$$AUC = X=1\text{-}specificity=1\text{-}(\frac{TP}{TN+FP}) \quad (10)$$

$$Y = sensitivity = \frac{TP}{TP+FN}$$

Where TP represents true positive, TN represents true negative, FP represents false positive and FN represents false negative. The receiver operating characteristic (ROC) curve is represented through the AUC (Area Under the ROC curve), plotting sensitivity on the y-axis, and specificity on the x-axis.

4.7. Multicollinearity analysis

The multicollinearity test determines the relationship among the gully erosion conditioning factors and values the level of non-independence among them (Ghosh and Maiti, 2021). The presence of collinearity may generate bias in the modelling process and decrease the predictive performance (Arabameri et al., 2021). In this study, two indices were used: Tolerance (TOL) and Variance Inflation Factors (VIF), calculated as follows:

$$TOL = 1 - R_i^2 \qquad (11)$$

$$VIF_i = \frac{1}{TOL} \qquad (12)$$

Where R indicates the coefficient of determination of each conditioning factor i (O'brien, 2007). If the value of TOL is less than 0.1 and the value of VIF is greater than 10, collinearity exists amongst the variables. Table 2 represents the multicollinearity analysis of the gully erosion factors used.

Table 2 Multi-collinearity among conditioning factors.

Factors	Collinearity Statistics	
	Tolerance	VIF
Elevation	0.427	2.341
Slope	0.881	1.135
Aspect	0.968	1.033
Lithology	0.718	1.393
Plan curvature	0.903	1.108
Distance to fault	0.910	1.098
Rainfall	0.414	2.413
Distance to stream	0.856	1.168

Distance to Road	0.920	1.087
TWI	0.967	1.034
NDVI	0.812	1.232
LULC	0.957	1.045

5. Results

5.1. Weight of evidence

Table 3 represents the results of W+, W- and the Cw, calculated for the 12 factors used in this study. It is shown that: the greatest sensitivity to erosion has a slope in the range of 7.23° to 15.75º (Cw=1.857), this confirms the interpretation that the phenomenon of erosion is less visible due to the fact that some steeply sloping areas are made of very hard rocks (dolomites and limestones). For the elevation, the most important class for erosion is the class; 792 m to 1.801, which is characterized by Cw= 2.201. The greatest susceptibility to erosion is in the southwestern aspect class (Cw=0.316). For the curvature plane, the strongest erosion factor is represented by the concave land; the most erosion-sensitive class is between 792 m and 1801 m (Cw= 1.849). The highest erosion sensitivity is observed when the distance to roads parameter is between 1818 m and 3783 m (Cw= 2.694) and the distance to faults is between 4817 m and 7561 m (Cw= 2.081). The highest sensitivity to erosion is observed when the TWI parameter is between 10.54 and 17.88 (Cw= 0.8) and the precipitation is between 273 mm and 363 mm (Cw= 2.635). For the lithology, the maximum sensitivity to erosion was observed in the loose formations of Quaternary age in class A (Cw= 2.152). The non-agricultural areas in Elouar watershed, represent the highest vulnerability areas to erosion, due to the absence of vegetation. The LC/LU factor shows that the building and construction areas in the study area represent the most erodible areas (Cw= 4.513). The NDVI class most susceptible to erosion is characterized by index values between (-0.3) and 0.12 (Cw= 2.833). The high and very high frequency erodible areas were observed in the middle and south of the El Ouaar watershed. The interpretation of the table shows that the higher the value of Cw, the more sensitive the class is to erosion, in consideration with the measurements of W+ and W-.

Table 3 WoE (C) values and factors affecting gully erosion.

Factors	Class/type	W+	W-	Cw
Slope	0-7.23	-1.411	0.465	-1.876
	7.23-15.75	1.297	-0.560	1.857
	15.75-24.79	0.435	-0.122	0.557
	24.79-35.13	-0.764	0.101	-0.865
	35.13-65.86	-0.764	0.047	-0.811
Elevation	214-792	-1.004	0.550	-1.554
	792-1,801	0.948	-1.254	2.201
	1,801-3,353	0.000	0.267	-0.267
Aspect	F	0.022	-0.003	0.025
	N	-0.878	0.078	-0.956
	NE	-0.360	0.053	-0.413
	E	-0.378	0.061	-0.439
	SE	-0.551	0.091	-0.642
	S	0.099	-0.020	0.119
	SW	0.264	-0.052	0.316
	W	-0.063	0.007	-0.070
	NW	-0.573	0.039	-0.612
Plan curvature	Concave	1.456	-0.393	1.849
	Flat	-1.227	1.457	-2.685
	Convex	0.651	-0.251	0.902
Distance to road	0-1,818	-1.554	0.346	-1.900
	1,818-3,783	1.652	-1.041	2.694
	3,783-5,995	-0.082	0.020	-0.102
	5,995-8,550	-1.042	0.091	-1.134
	8,550-12,531	0.000	0.148	-0.148
Distance to stream	0-1,229	-0.094	0.057	-0.152
	1,229-2,683	-0.204	0.081	-0.285
	2,683-4,322	-0.071	0.021	-0.092
	4,322-6,297	-0.490	0.078	-0.568
	6,297-9,502	-0.653	0.055	-0.708
Distance to fault	0-2,184	-0.309	0.144	-0.453
	2,184-4,817	-0.611	0.176	-0.786
	4,817-7,561	1.953	-0.128	2.081

	7,561-10,586	0.220	-0.046	0.266
	10,586-14,283	0.486	-0.114	0.601
TWI	10.54-17.88	0.714	-0.085	0.800
	17.88-19.63	0.455	-0.172	0.626
	19.63-21.30	0.436	-0.202	0.638
	21.30-29.07	0.592	-0.061	0.654
Rainfall	207-273	-1.218	0.451	-1.670
	273-363	1.904	-0.731	2.635
	363-443	1.032	-0.253	1.286
	443-525	0.000	0.267	-0.267
	525-625	0.000	0.167	-0.167
Lithology	A	2.071	-0.082	2.152
	B	0.000	-0.123	0.123
	C	0.000	-0.208	0.208
	D	0.000	-0.348	0.348
	E	0.000	-0.045	0.045
	F	0.000	-0.477	0.477
	G	0.000	0.004	-0.004
	H	0.000	-0.070	0.070
	I	0.000	0.041	-0.041
LC/LU	Water	0.000	0.000	0.000
	Greenhouse	0.196	-0.011	0.207
	Agriculture	0.689	-0.045	0.734
	Building/Construction	4.329	-0.184	4.513
	Soil	-0.262	1.083	-1.345
NDVI	(-0.3)-0.12	1.634	-1.199	2.833
	0.12-0.17	-1.333	0.474	-1.808
	0.17-0.25	-0.224	0.025	-0.249
	0.25-0.36	-0.199	0.009	-0.208
	0.36-0.71	-1.526	0.030	-1.556

5.2. Gully erosion susceptibility mapping and models performance

The gully erosion susceptibility of each model was classified into 5 classes: very low, low, moderate, high, and very high. The results were presented in Fig. 9 and Table 4.

The gully erosion susceptibility map developed using the RF model showed that 23.44% of the study area had very high erosion susceptibility, while 17.02%, 11.67%, 17.01%, and 30.85% of the area were classified as very low, low, moderate and high susceptibility, respectively.

For the MLP model, 28.39% of the area was classified as very high susceptibility, while 8.37%, 18.3%, 15.58%, and 29.37% had very low, low, moderate, and high susceptibilities, respectively. For the KNN model, 27.64% of the study area was classified as very high gully risk, while 17.72%, 30.96%, 18.47%, and 5.21% had very low, low, moderate, and high sensitivities, For the LR model 19.66% of the study, the area was classified as very high gully, while 6.4%, 24.02%, 24.23%, and 25.69% had very low, low, moderate and high susceptibilities, respectively.

The spatial distribution of gully erosion susceptibility within this study area was quite similar for all Machine Learning models developed here. These models show that the most eroded areas are located in the southern part of Elouar watershed. These areas are characterized by a very high intensity of erosion where the lithological formation is mainly dominated by poorly consolidated quaternary deposits (silts and clays). These areas are characterized by a variation in slope, which rapidly increases the transport of fine sediments in this area. In addition, inappropriate agricultural practices and overgrazing of the area may also act as driving forces of gully erosion in the study area. The northern part is characterized by less intense erosion than the southern part of this basin. Indeed, this area is constituted by highly consolidated geological deposits which are represented by limestones and dolomites (Ambroggi, 1963).

From a geomorphological point of view, the study area is part of the Souss plain which is located between the High Atlas to the north and the Anti Atlas to the south, their geomorphological position gives a great variation in the altitude, and for this reason the factors of erosion during rainy periods are very intense. In addition, the intersection between the High Atlas and the Souss plain shows a very strong water current and favours water erosion, and add to this, the sudden succession of floods in the region during the past years.

443 **Table 4** Percentages of gully erosion susceptibility classes.

Models	**% RF - WoE**	**% MLP- WoE**	**% KNN- WoE**	**% LR- WoE**
Very low	23.44	8.37	5.21	6.40
Low	17.02	18.30	18.47	24.02
Moderate	11.67	15.58	30.96	24.23
High	17.01	29.37	17.72	25.69
Very high	30.85	28.39	27.64	19.66

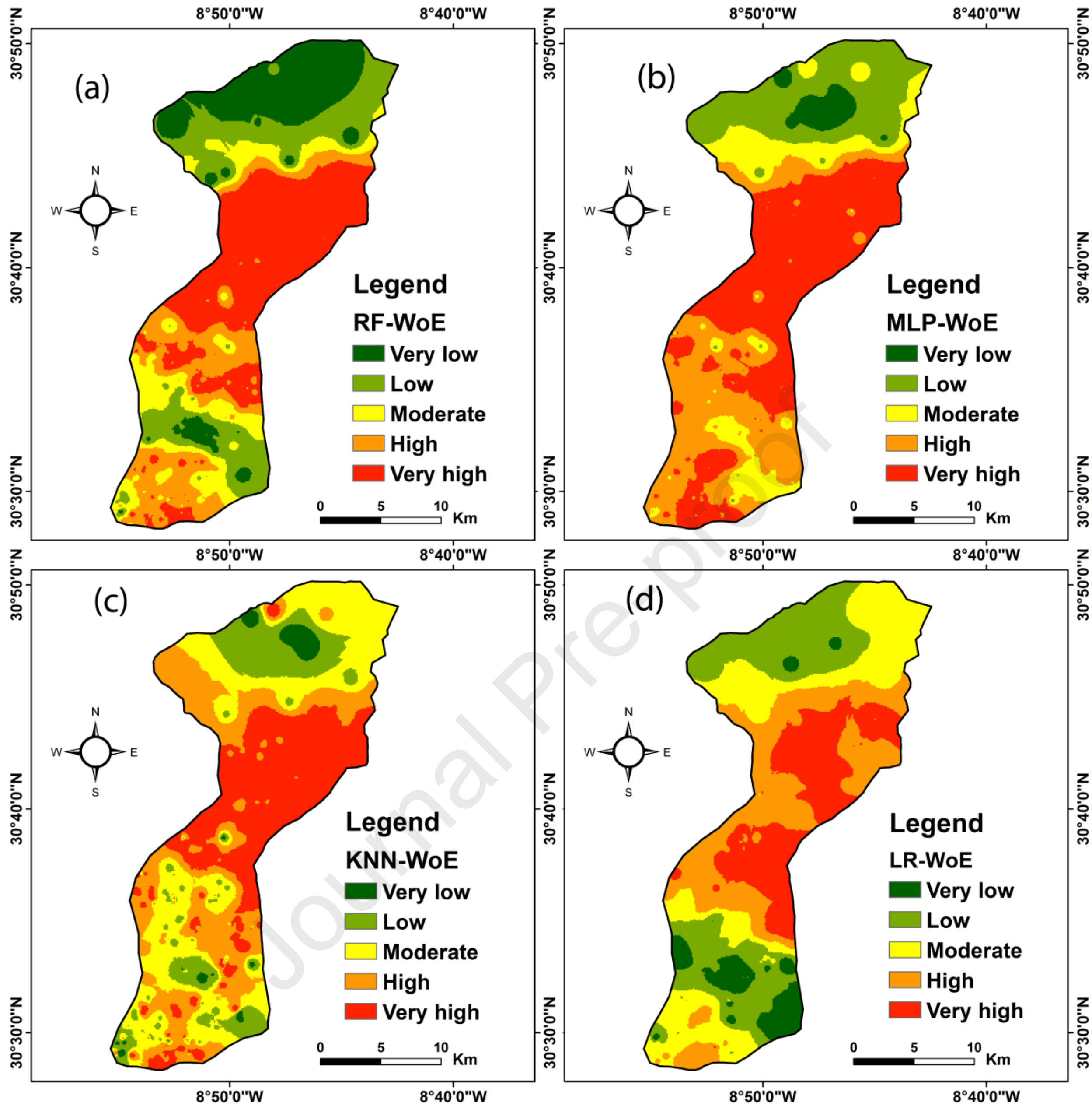


Fig. 9: Gully erosion susceptibility mapping using: (a) RF-WOE, (b) MLP-WOE, (c) KNN-WOE and LR-WOE.

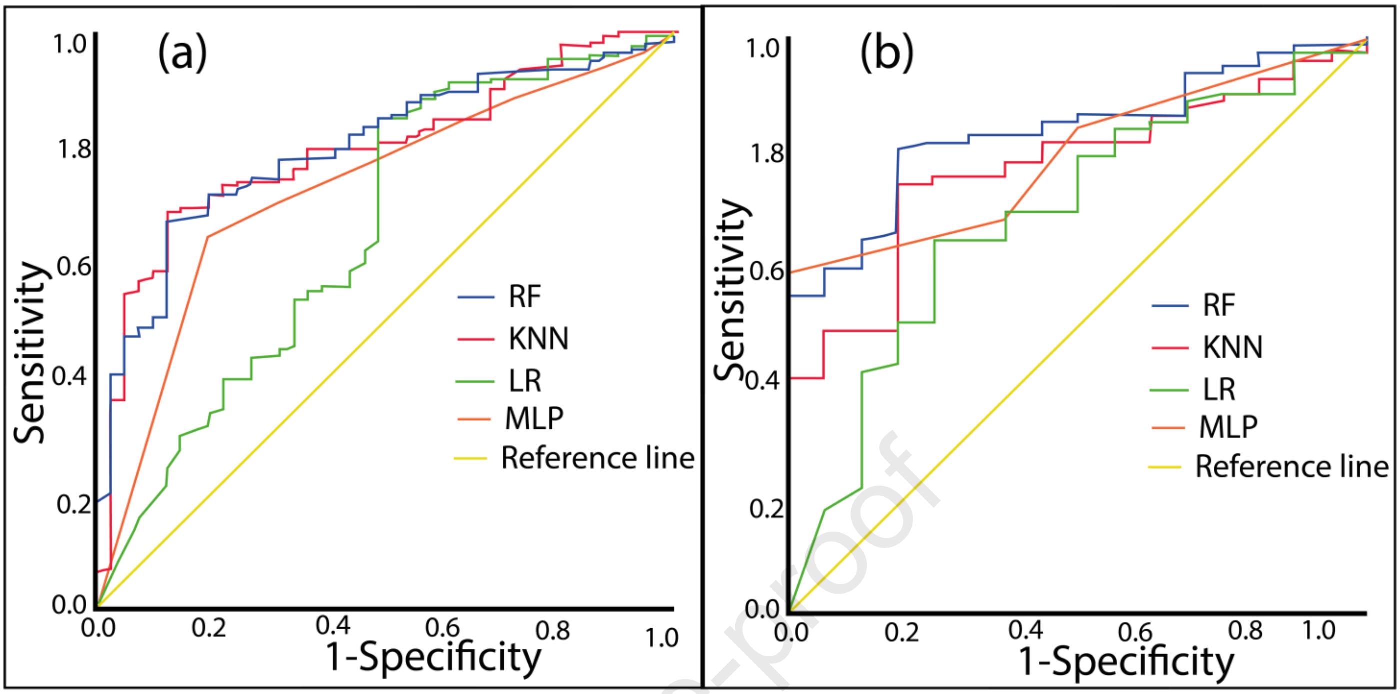


Fig. 10: (a) ROC curves of success rate and (b) ROC curves of prediction rate.

The performance of the developed models was evaluated using the receiver operating characteristic (ROC) curve analysis.

For RF model, the area under the curve (AUC) values are equals to 0.800 and 0.838 in the training and testing sets, respectively, as shown in Fig. 10.a and 10.b and Table 5. The model's accuracy of over 80% in the study area indicates that it is appropriate for mapping gully erosion susceptibility. The accuracy of the classification models decreases in the following order: MLP, KNN, and LR with AUC of 0.796, 0.777, and 0.692, respectively, on the testing set. The study demonstrated that the RF model exhibited good performance in classification problems compared to other models, which is consistent with the findings of several previous studies (Avand et al., 2019; Rahmati et al., 2017; Saha et al., 2020).

Table 5 Model statistical measures assigned to the training and testing datasets

	RF- WoE		MLP- WoE		KNN- WoE		LR- WoE	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Accuracy	96.744	86.869	92.093	79.208	81.395	84.466	92.093	80.808
Precision	99.408	98.611	91.979	88.095	82.212	87.368	91.979	88.095
Sensitivity	96.552	85.542	98.851	87.059	98.276	95.402	98.851	89.157
Specificity	97.561	93.750	63.415	37.500	9.756	25.000	63.415	37.500
AUC	0.800	0.838	0.729	0.796	0.796	0.777	0.655	0.692

The RF model is a powerful and well-functioning model that has been demonstrated to be robust and consistent with previous research. It is a sophisticated technique in spatial sciences that has the ability to utilize multiple input variables and produce high accuracy predictions for various classes. It has the ability to use explanatory variables and identify nonlinear relationships between independent and dependent variables, making it a strong model for environmental hazard assessment. Compared to other models, the RF model has the advantage of being able to handle large datasets and manage numerous input variables efficiently. Its accurate machine learning algorithms make it a highly accurate classifier for many datasets. In this study, the importance of variables for gully erosion mapping for the El Ouaar watershed was performed based on the RF model. The variable importance values were LULC (0.06), NDVI (0.03), TWI (0.09), distance to Road (0.13), distance to Stream (0.11), Rainfall (0.19), Distance to fault (0.01), Plan curvature (0.04), Lithology (0.21), Aspect (0.07) and Slope (0.04) and Elevation (0.03). The most important factors for gully susceptibility mapping in the Ouaar watershed were lithology (0.21) and rainfall (0.19), while distance to fault (0.01), was the least important (Fig. 11).

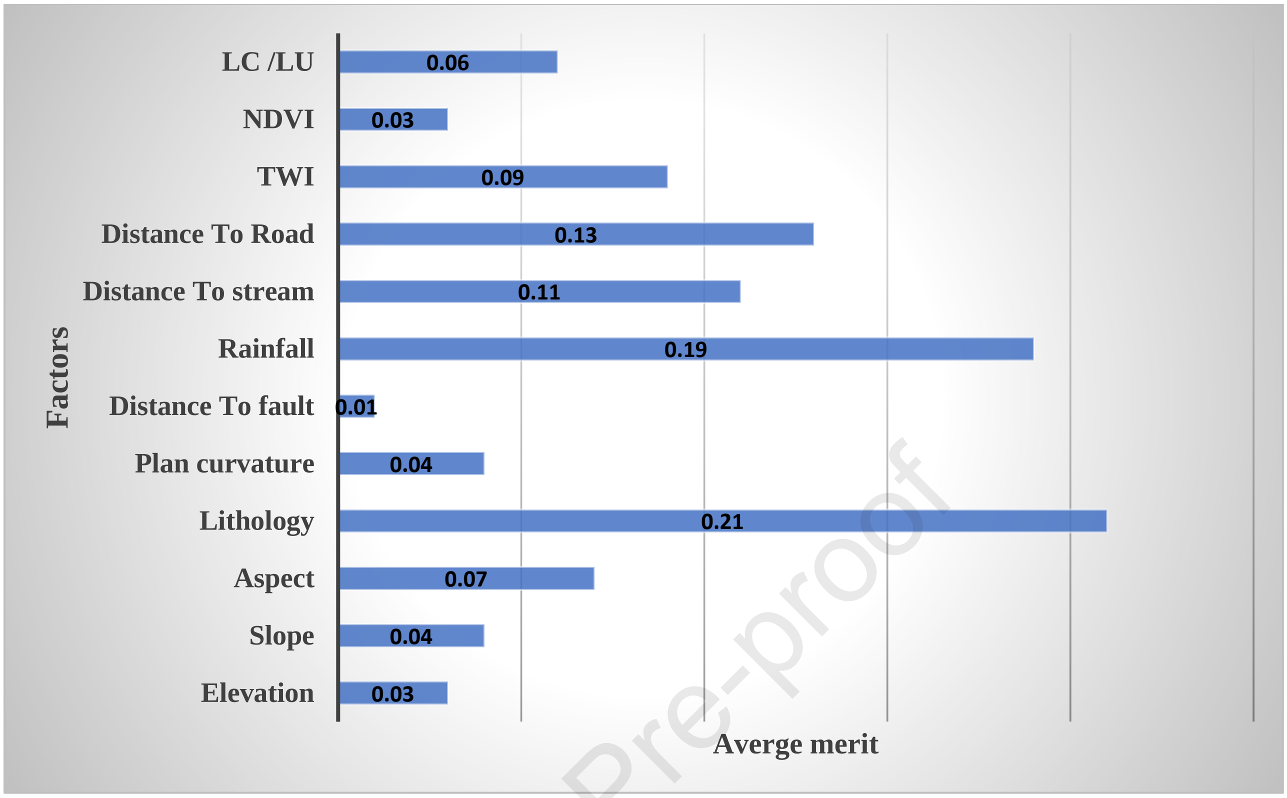


Fig. 11: The importance of conditioning factors.

6. Discussion

Gully erosion susceptibility models based on machine learning algorithms have been recognized as an effective tool for soil ecosystem management worldwide (Arabameri et al., 2021; Roy and Saha 2022; Wang et al., 2022) . In general, among other machine learning techniques, RF with its capacity to efficiently handle large datasets, non-linear parameters, categorical and continuous data, over-fitting, outliers, and multiple features was reported to produce the best performance in terms of high accuracy (Chen et al., 2021; Hembram et al., 2021; Lana et al., 2022; Pourghasemi et al., 2020; Saha et al., 2021)

The RF model has more precision than the other models, according to the results of this analysis, in creating a map of susceptibility to gully erosion (Fig. 10 and Table 5). An approach to modelling and analysing numerical data that includes both independent and dependent variables is called the analysis of regression. In order to forecast the future behaviour of the dependent variable, regression analysis aims to represent the dependent variable as an independent function of variables, coefficients, and error values. When the link between the dependent and independent variables is positive in certain areas of the research area and negative in other areas, it is obvious that logistic regression cannot accurately and precisely detect the relationship.

This study, in agreement with recently published studies, also found that the RF algorithm is the most suitable model for mapping gully erosion susceptibility in the El Ouaar watershed based on different performance criteria. Although several studies reported that other machine learning models, for example, boosted regression tree (BRT) (Amiri et al., 2019) or extreme gradient boosting (XGBoost) (Yang et al., 2021), generated better performance compared to the RF model, the additional advantage of RF lies in its ability to evaluate the importance of each conditioning factor in modelling process. This beneficial feature makes RF-based models widely used for modelling processes in general and suitable for gully erosion susceptibility mapping in particular. In addition, studies that better explain the model performance through the analysis of insight mechanisms such as the distributions of variables and their interaction, rather than a pure comparison based on the statistical criteria (e.g., RMSE and MAE), are strongly encouraged.

The present study indicated that lithology, rainfall, and distance to stream and road were the most important variables influencing gully erosion susceptibility modelling. These findings are largely in agreement with recent research (Rahmati et al., 2016; Amiri et al., 2019; Tien Bui et al., 2019; Chen et al., 2021) where rainfall, lithology, and the distance from streams/rivers are generally more important variables contributing to gully erosion than other conditioning factors. The results from the four machine learning models used in this study also confirmed that the regions with moderate rainfall, elevation, and slope but close to streams and roads are located in very high gully erosion susceptibility areas. In contrast, the influence of LULC on gully erosion was found to be not significantly strong (ranked 7th out of 12 variables) in this study. However, it is worth noting that bare lands cover most of the study area. The results show a greater concentration of gully erosion in areas with bare land in this study, which is in line with the report by Lei et al. (2020) and Chen et al. (2020). The impact of land use/land cover on gully activity was also reported previously (Vandekerckhove et al., 2003) ; nevertheless, there remain many doubts about the features, such as mismanagement (Hosseinalizadeh et al., 2019) , that induce subsurface gully development. Therefore, future research should explore further the uncertainties of LULC variables in the development of gully erosion.

Recently, hybrid models based on the combination of two or more techniques have been highly recommended for gully susceptibility prediction and mapping (Arabameri et al., 2020; Hitouri et al., 2022; Roy and Saha 2022). This study also developed a hybrid machine learning model for gully susceptibility prediction by integrating the Weight of Evidence (WoE) with Multilayer Perceptron (MLP-WoE), Logistic Regression (LR-WoE), K-Nearest Neighbors (KNN-WoE) and Random Forest (RF-WoE). While a comparison with stand-alone models was not performed by the present study, other authors have reported that hybrid models can deliver better and perfect results (Arabameri et al., 2020; Hembram et al., 2021). Tien Bui et al. (2019) proposed a hybrid RF-ADTree model based on RF and alternating decision tree (ADTree) algorithms that were able to significantly improve the prediction accuracy of the stand-alone ADTree model. Roy and Saha (2022) also indicated that the integrated RSS-RBFnn and RTF-RBFnn models, i.e., radial basis function neural network (RBFnn) combined with random sub-space (RSS) and rotation forest (RTF), showed better results than the single RBFnn model for gully erosion susceptibility maps. With the significant increase in the number of machine learning algorithms, future comparative evaluations of single and hybrid models are important for the assessment of performance and accuracy, since different modelling techniques may produce very different results and performances. In addition, hybrid models developed from a combination with deep learning (Band et al. 2020; Chen et al., 2021) can be promising studies of gully erosion susceptibility.

The geological and geographical situation of the Elouar watershed contribute to the risk of gully erosion. Indeed, this watershed is part of the Souss Basin. The plain of this basin is filled with recent Quaternary deposits, consisting of loose to semi-compact layers that are sensitive to erosion. These deposits are composed of lithostratigraphic units U1 to U4 mentioned above (Aït Hssaïne, 1994; Ambroggi, 1963; Hssaisoune et al., 2012). Climate factors also play an important role in this area. The study area is characterized by an arid and semi-arid climate, influenced by the geographical location, especially the High Atlas Mountains to the north, the Anti-Atlas Mountains to the south, and the Atlantic Ocean to the east. The variation in rainfall between the north and south promotes strong water currents in the rivers due to the altitude difference between the upstream area of the study area, which can reach over 3000 meters, and its downstream area, which can be as low as 194 meters, facilitating water transport and contributing to the gully erosion risks. Overgrazing and anthropogenic factors also contribute to erosion risks in this area, and laws should be enacted to regulate irresponsible land use practices. The effect of this phenomenon is manifested in the destruction of infrastructure (roads, bridges, houses...), which results in significant material and economic losses for the country. For instance, the Faculty of Sharia and Law in Taroudant, constructed in a highly area-prone to erosion by ravines, has led to the destruction of walls and the emergence of erosive areas both inside and outside of the faculty (Fig. 12).

Fig. 12: Photos showing the influence of gully erosion on the infrastructure within the study area.

To minimize this influence, several measures have been taken to address the issue of gully erosion in the Taroudant region. These include the construction of walls along the banks of the Elouar river and the gullies that are particularly prone to erosion. Additionally, the construction of sidewalks next to roads and the implementation of greenhouse agriculture have also been considered as part of the actions to mitigate and overcome the spread of this phenomenon (Fig. 13).

Fig. 13: Photos showing some solutions implemented to prevent the distribution of gully erosion in the study area.

In this context, our study underscores the critical importance of the developed models in the field of gully erosion mapping. However, we recognize that there is a room for improvement in future research. For instance, in our study, we based on freely available GIS data due to the lack of high-resolution datasets. We acknowledge that this choice of data has certain limitations, specifically regarding data resolution for controlling factors, which resulted in the low accuracy and low precision of the developed models. It should be noted that accurate and detailed gully erosion requires high spatial resolutions (Garosi et al., 2018; Rahmati et al., 2016). In addition, it is essential to highlight the significance of data quality and recognizing the uncertainties associated with using controlling factors with different pixel sizes is of utmost importance. Although many studies have investigated the susceptibility of gully erosion based on different pixel size of some controlling factors (Garosi et al., 2018; Rahmati et al., 2016). There is still ongoing debate regarding the most appropriate pixel size to consider when examining controlling factors for gully mapping susceptibility.

7. Conclusion

In conclusion, this study successfully achieved its goal of using multi-collinearity analysis to identify significant factors in gully erosion, creating hybrid machine learning models to map erosion-prone areas, employing k-fold cross-validation to mitigate randomness, and assessing the capability and robustness of the models using ROC. This study shows that the weight of evidence is very important in identifying the most suitable conditioning factors to generate an effective map of gully erosion susceptibility in the El Ouaar watershed. The results showed that RF-WoE obtained the best performance (AUC = 0.8), followed by KNN-WoE (AUC = 0.796), then MLP-WoE (AUC = 0.729) and LP-WoE (AUC = 0.655), respectively. These results showed that the good precision obtained is due to the fact that each type of erosion has its own set of conditioning factors, which must be evaluated separately. The results obtained from this work provide planners and researchers with an appropriate perspective on the effect of conditioning factors in future analysis. Further research could explore the use of other machine learning techniques and consider additional factors to improve the accuracy of gully erosion prediction models.

Ethical Approval: Not applicable

Consent to Participate: Not applicable

Consent to Publish: Not applicable

Funding: No external funding

Competing Interests: This manuscript has not been published or presented elsewhere in part or entirety and is not under consideration by another journal. There are no conflicts of interest to declare.

Acknowledgement: The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work, under the Research Groups Funding program grant code (NU/RG/SERC/12/21).

Availability of data and materials: The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Abraham, M.T., Satyam, N., Lokesh, R., Pradhan, B., Alamri, A., 2021. Factors affecting landslide susceptibility mapping: Assessing the influence of different machine learning approaches, sampling strategies and data splitting. Land 10, 989.

Abu El-Magd, S.A., Ali, S.A., Pham, Q.B., 2021. Spatial modeling and susceptibility zonation of landslides using random forest, naïve bayes and K-nearest neighbor in a complicated terrain. Earth Science Informatics 14, 1227–1243.

Aït Hssaïne, A., 1994. Géomorphologie et Quaternaire du piémont de Taroudant-Ouled Teïma, Maroc (PhD Thesis). Thèse inédite, université de Montréal.

Ajit, P., 2016. Prediction of employee turnover in organizations using machine learning algorithms. algorithms 4, C5.

Ali, S.A., Parvin, F., Pham, Q.B., Khedher, K.M., Dehbozorgi, M., Rabby, Y.W., Anh, D.T., Nguyen, D.H., 2022. An ensemble random forest tree with SVM, ANN, NBT, and LMT for landslide susceptibility mapping in the Rangit River watershed, India. Natural Hazards 1–33.

Ali, S.A., Parvin, F., Pham, Q.B., Vojtek, M., Vojteková, J., Costache, R., Linh, N.T.T., Nguyen, H.Q., Ahmad, A., Ghorbani, M.A., 2020. GIS-based comparative assessment of flood susceptibility mapping using hybrid multi-criteria decision-making approach, naïve Bayes tree, bivariate statistics and logistic regression: a case of Topl’a basin, Slovakia. Ecological Indicators 117, 106620.

Ali, S.A., Parvin, F., Vojteková, J., Costache, R., Linh, N.T.T., Pham, Q.B., Vojtek, M., Gigović, L., Ahmad, A., Ghorbani, M.A., 2021. GIS-based landslide susceptibility modeling: A comparison between fuzzy multi-criteria and machine learning algorithms. Geoscience Frontiers 12, 857–876.

Amare, S., Langendoen, E., Keesstra, S., Ploeg, M. van der, Gelagay, H., Lemma, H., Zee, S.E. van der, 2021. Susceptibility to gully erosion: applying random forest (RF) and frequency ratio (FR) approaches to a small catchment in Ethiopia. Water 13, 216.
Ambroggi, R., 1963. Etude géologique du versant méridional du Haut Atlas occidental et de la plaine du Souss. Editions marocaines et internationales.
Amiri, M., Pourghasemi, H.R., Ghanbarian, G.A., Afzali, S.F., 2019. Assessment of the importance of gully erosion effective factors using Boruta algorithm and its spatial modeling and mapping using three machine learning algorithms. Geoderma 340, 55–69.
Arabameri, A., Cerda, A., Tiefenbacher, J.P., 2019. Spatial pattern analysis and prediction of gully erosion using novel hybrid model of entropy-weight of evidence. Water 11, 1129.
Arabameri, A., Chandra Pal, S., Costache, R., Saha, A., Rezaie, F., Seyed Danesh, A., Pradhan, B., Lee, S., Hoang, N.-D., 2021. Prediction of gully erosion susceptibility mapping using novel ensemble machine learning algorithms. Geomatics, Natural Hazards and Risk 12, 469–498.
Arabameri, A., Pradhan, B., Rezaei, K., Yamani, M., Pourghasemi, H.R., Lombardo, L., 2018a. Spatial modelling of gully erosion using evidential belief function, logistic regression, and a new ensemble of evidential belief function–logistic regression algorithm. Land Degradation & Development 29, 4035–4049.
Arabameri, A., Rezaei, K., Pourghasemi, H.R., Lee, S., Yamani, M., 2018b. GIS-based gully erosion susceptibility mapping: a comparison among three data-driven models and AHP knowledge-based technique. Environmental earth sciences 77, 1–22.
Arabameri, A., Saha, S., Roy, J., Tiefenbacher, J.P., Cerda, A., Biggs, T., Pradhan, B., Thi Ngo, P.T., Collins, A.L., 2020. A novel ensemble computational intelligence approach for the spatial prediction of land subsidence susceptibility. Science of The Total Environment 726, 138595. https://doi.org/10.1016/j.scitotenv.2020.138595
Avand, M., Janizadeh, S., Naghibi, S.A., Pourghasemi, H.R., Khosrobeigi Bozchaloei, S., Blaschke, T., 2019. A comparative assessment of random forest and k-nearest neighbor classifiers for gully erosion susceptibility mapping. Water 11, 2076.
Azareh, A., Rahmati, O., Rafiei-Sardooi, E., Sankey, J.B., Lee, S., Shahabi, H., Ahmad, B.B., 2019. Modelling gully-erosion susceptibility in a semi-arid region, Iran: Investigation of applicability of certainty factor and maximum entropy models. Science of the Total Environment 655, 684–696.
Azedou, A., Lahssini, S., Khattabi, A., Meliho, M., Rifai, N., 2021. A methodological comparison of three models for gully erosion susceptibility mapping in the rural municipality of El Faid (Morocco). Sustainability 13, 682.
Band, S.S., Janizadeh, S., Chandra Pal, S., Saha, A., Chakrabortty, R., Shokri, M., Mosavi, A., 2020. Novel ensemble approach of deep learning neural network (DLNN) model and particle swarm optimization (PSO) algorithm for prediction of gully erosion susceptibility. Sensors 20, 5609.

Belasri, A., Lakhouili, A., 2016. Estimation of soil erosion risk using the universal soil loss equation (USLE) and geo-information technology in Oued El Makhazine Watershed, Morocco. Journal of Geographic Information System 8, 98.

Belayneh, M., Yirgu, T., Tsegaye, D., 2020. Current extent, temporal trends, and rates of gully erosion in the Gumara watershed, Northwestern Ethiopia. Global Ecology and Conservation 24, e01255.

Bilotta, G.S., Brazier, R.E., Haygarth, P.M., 2007. The impacts of grazing animals on the quality of soils, vegetation, and surface waters in intensively managed grasslands. Advances in agronomy 94, 237–280.

Bonham-Carter, G.F., Agterberg, F.P., Wright, D.F., 1988. Integration of geological datasets for gold exploration in Nova Scotia. Photogrammetric Engineering and Remote Sensing 54, 1585–1592.

Bouslihim, Y., 2020. Hydrological and soil erosion modeling using SWAT model and Pedotransfert Functions: A case study of Settat-Ben Ahmed watersheds, Morocco (PhD Thesis). Université Hassan Ier Settat (Maroc).

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Chakrabortty, R., Pal, S.C., 2023. Systematic review on gully erosion measurement, modelling and management: Mitigation alternatives and policy recommendations. Geological Journal.

Chen, W., Lei, X., Chakrabortty, R., Pal, S.C., Sahana, M., Janizadeh, S., 2021. Evaluation of different boosting ensemble machine learning models and novel deep learning and boosting framework for head-cut gully erosion susceptibility. Journal of Environmental Management 284, 112015.

Chen, W., Li, H., Hou, E., Wang, S., Wang, G., Panahi, M., Li, T., Peng, T., Guo, C., Niu, C., Xiao, L., Wang, J., Xie, X., Ahmad, B. Bin, 2018. GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. Science of The Total Environment 634, 853–867. https://doi.org/10.1016/j.scitotenv.2018.04.055

Chen, Y., Qin, S., Qiao, S., Dou, Q., Che, W., Su, G., Yao, J., Nnanwuba, U.E., 2020. Spatial predictions of debris flow susceptibility mapping using convolutional neural networks in Jilin Province, China. Water 12, 2079.

Choubin, B., Moradi, E., Golshan, M., Adamowski, J., Sajedi-Hosseini, F., Mosavi, A., 2019. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Science of the Total Environment 651, 2087–2096.

Chu, L., Wang, L.-J., Jiang, J., Liu, X., Sawada, K., Zhang, J., 2019. Comparison of landslide susceptibility maps using random forest and multivariate adaptive regression spline models in combination with catchment map units. Geosciences Journal 23, 341–355.

Conforti, M., Aucelli, P.P., Robustelli, G., Scarciglia, F., 2011. Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (Northern Calabria, Italy). Natural hazards 56, 881–898.
Conoscenti, C., Angileri, S., Cappadonia, C., Rotigliano, E., Agnesi, V., Märker, M., 2014. Gully erosion susceptibility assessment by means of GIS-based logistic regression: A case of Sicily (Italy). Geomorphology 204, 399–411. https://doi.org/10.1016/j.geomorph.2013.08.021
Davis, C.R., Trevatt, A.E., McGoldrick, R.B., Parrott, F.E., Mohanna, P.-N., 2016. How to train plastic surgeons of the future. Journal of Plastic, Reconstructive & Aesthetic Surgery 69, 1134–1140.
Dijon, R., 1966. Reconnaissance hydrogéologique et ressources en eau du bassin des oueds Seyad-Ouarg-Noun, Maroc Sud-occidental (PhD Thesis). Montpellier.
d'Oleire-Oltmanns, S., Marzolff, I., Tiede, D., Blaschke, T., 2014. Detection of gully-affected areas by applying object-based image analysis (OBIA) in the region of Taroudannt, Morocco. Remote Sensing 6, 8287–8309.
Elmoulat, M., Ait Brahim, L., 2018. Landslides susceptibility mapping using GIS and weights of evidence model in Tetouan-Ras-Mazari area (Northern Morocco). Geomatics, Natural Hazards and Risk 9, 1306–1325.
Fressard, M., Thiery, Y., Maquaire, O., 2014. Which data for quantitative landslide susceptibility mapping at operational scale? Case study of the Pays d'Auge plateau hillslopes (Normandy, France). Natural Hazards and Earth System Sciences 14, 569–588. https://doi.org/10.5194/nhess-14-569-2014
Gafurov, A.M., Yermolayev, O.P., 2020. Automatic gully detection: Neural networks and computer vision. Remote Sensing 12, 1743.
Gardner, M.W., Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric environment 32, 2627–2636.
Garosi, Y., Sheklabadi, M., Pourghasemi, H.R., Besalatpour, A.A., Conoscenti, C., Van Oost, K., 2018. Comparison of differences in resolution and sources of controlling factors for gully erosion susceptibility mapping. Geoderma 330, 65–78.
Ghorbani Nejad, S., Falah, F., Daneshfar, M., Haghizadeh, A., Rahmati, O., 2017. Delineation of groundwater potential zones using remote sensing and GIS-based data-driven models. Geocarto international 32, 167–187.
Ghorbanzadeh, O., Blaschke, T., Aryal, J., Gholaminia, K., 2020a. A new GIS-based technique using an adaptive neuro-fuzzy inference system for land subsidence susceptibility mapping. Journal of Spatial Science 65, 401–418. https://doi.org/10.1080/14498596.2018.1505564
Ghorbanzadeh, O., Shahabi, H., Mirchooli, F., Valizadeh Kamran, K., Lim, S., Aryal, J., Jarihani, B., Blaschke, T., 2020b. Gully erosion susceptibility mapping (GESM) using

machine learning methods optimized by the multi-collinearity analysis and K-fold cross-validation. Geomatics, Natural Hazards and Risk 11, 1653–1678.
Ghorbanzadeh, O., Valizadeh Kamran, K., Blaschke, T., Aryal, J., Naboureh, A., Einali, J., Bian, J., 2019. Spatial prediction of wildfire susceptibility using field survey gps data and machine learning approaches. Fire 2, 43.
Ghosh, A., Maiti, R., 2021. Soil erosion susceptibility assessment using logistic regression, decision tree and random forest: study on the Mayurakshi river basin of Eastern India. Environmental Earth Sciences 80, 1–16.
Golestani, G., Issazadeh, L., Serajamani, R., 2014. Lithology effects on gully erosion in Ghoori chay Watershed using RS & GIS. Int. J. Biosci 4, 71–76.
Gomiero, T., 2016. Soil degradation, land scarcity and food security: Reviewing a complex challenge. Sustainability 8, 281.
Hancock, G.R., Evans, K.G., 2010. Gully, channel and hillslope erosion–an assessment for a traditionally managed catchment. Earth surface processes and landforms 35, 1468–1479.
Hembram, T.K., Saha, S., Pradhan, B., Abdul Maulud, K.N., Alamri, A.M., 2021. Robustness analysis of machine learning classifiers in predicting spatial gully erosion susceptibility with altered training samples. Geomatics, Natural Hazards and Risk 12, 794–828. https://doi.org/10.1080/19475705.2021.1890644
Hitouri, S., Varasano, A., Mohajane, M., Ijlil, S., Essahlaoui, N., Ali, S.A., Essahlaoui, A., Pham, Q.B., Waleed, M., Palateerdham, S.K., 2022. Hybrid Machine Learning Approach for Gully Erosion Mapping Susceptibility at a Watershed Scale. ISPRS International Journal of Geo-Information 11, 401.
Hong, H., Naghibi, S.A., Moradi Dashtpagerdi, M., Pourghasemi, H.R., Chen, W., 2017. A comparative assessment between linear and quadratic discriminant analyses (LDA-QDA) with frequency ratio and weights-of-evidence models for forest fire susceptibility mapping in China. Arabian Journal of Geosciences 10, 1–14.
Hosmer, D.W., Lemeshow, S., Sturdivant, R.X., 2000. Introduction to the logistic regression model. Applied logistic regression 2, 1–30.
Hosseinalizadeh, M., Kariminejad, N., Chen, W., Pourghasemi, H.R., Alinejad, M., Behbahani, A.M., Tiefenbacher, J.P., 2019. Gully headcut susceptibility modeling using functional trees, naïve Bayes tree, and random forest models. Geoderma 342, 1–11.
Hssaisoune, M., Boutaleb, S., Benssaou, M., Tagma, T., Fasskaoui, M.E., Bouchaou, L., 2012. Analyse géophysique et structurale de l'aquifère de la plaine du Souss-Massa: synthèse et conséquences hydrogéologiques 20.
Huang, D., Su, L., Zhou, L., Tian, Y., Fan, H., 2023. Assessment of gully erosion susceptibility using different DEM-derived topographic factors in the black soil region of Northeast China. International Soil and Water Conservation Research 11, 97–111. https://doi.org/10.1016/j.iswcr.2022.04.001
Hughes, A., Prosser, I., Stevenson, J., Scott, A., Lu, H., Gallant, J., Moran, C., 2001. Gully Erosion Mapping for the National Land and Water Resources Audit.

Hussain, M.A., Chen, Z., Kalsoom, I., Asghar, A., Shoaib, M., 2022. Landslide Susceptibility Mapping Using Machine Learning Algorithm: A Case Study Along Karakoram Highway (KKH), Pakistan. J Indian Soc Remote Sens 50, 849–866. https://doi.org/10.1007/s12524-021-01451-1

Issaka, S., Ashraf, M.A., 2017. Impact of soil erosion and degradation on water quality: a review. Geology, Ecology, and Landscapes 1, 1–11.

Jaafari, A., Janizadeh, S., Abdo, H.G., Mafi-Gholami, D., Adeli, B., 2022. Understanding land degradation induced by gully erosion from the perspective of different geoenvironmental factors. Journal of Environmental Management 315, 115181.

Jiang, C., Fan, W., Yu, N., Nan, Y., 2021. A new method to predict gully head erosion in the Loess Plateau of China based on SBAS-InSAR. Remote Sensing 13, 421.

Karami, A., Khoorani, A., Noohegar, A., Shamsi, S.R.F., Moosavi, V., 2015. Gully Erosion Mapping Using Object-Based and Pixel-Based Image Classification MethodsGully Erosion Mapping. Environmental & Engineering Geoscience 21, 101–110.

Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization.

Lana, J.C., Castro, P. de 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. https://doi.org/10.1016/j.geomorph.2022.108159

Lei, X., Chen, W., Avand, M., Janizadeh, S., Kariminejad, N., Shahabi, Hejar, Costache, R., Shahabi, Himan, Shirzadi, A., Mosavi, A., 2020. GIS-based machine learning algorithms for gully erosion susceptibility mapping in a semi-arid region of Iran. Remote Sensing 12, 2478.

Lin, G.-W., Chen, H., 2012. The relationship of rainfall energy with landslides and sediment delivery. Engineering geology 125, 108–118.

Lucà, F., Conforti, M., Robustelli, G., 2011. Comparison of GIS-based gullying susceptibility mapping using bivariate and multivariate statistics: Northern Calabria, South Italy. Geomorphology 134, 297–308.

Meliho, M., Khattabi, A., Mhammdi, N., 2018. A GIS-based approach for gully erosion susceptibility modelling using bivariate statistics methods in the Ourika watershed, Morocco. Environmental Earth Sciences 77, 1–14.

Merghadi, A., Yunus, A.P., Dou, J., Whiteley, J., ThaiPham, B., Bui, D.T., Avtar, R., Abderrahmane, B., 2020. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews 207, 103225.

Mohsin, M., Ali, S.A., Shamim, S.K., Ahmad, A., 2022. A GIS-based novel approach for suitable sanitary landfill site selection using integrated fuzzy analytic hierarchy process and machine learning algorithms. Environmental Science and Pollution Research 29, 31511–31540.

Momm, H.G., Bingner, R.L., Wells, R.R., Wilcox, D., 2012. AGNPS GIS-based tool for watershed-scale identification and mapping of cropland potential ephemeral gullies. Applied engineering in agriculture 28, 17–29.

Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrological processes 5, 3–30.

Moore, I.D., Wilson, J.P., 1992. Length-slope factors for the Revised Universal Soil Loss Equation: Simplified method of estimation. Journal of soil and water conservation 47, 423–428.

Mosaid, H., Barakat, A., Bustillo, V., Rais, J., 2022. Modeling and Mapping of Soil Water Erosion Risks in the Srou Basin (Middle Atlas, Morocco) Using the EPM Model, GIS and Magnetic Susceptibility. Journal of Landscape Ecology 15, 126–147.

Nhu, V.-H., Shirzadi, A., Shahabi, H., Singh, S.K., Al-Ansari, N., Clague, J.J., Jaafari, A., Chen, W., Miraki, S., Dou, J., 2020. Shallow landslide susceptibility mapping: A comparison between logistic model tree, logistic regression, naïve bayes tree, artificial neural network, and support vector machine algorithms. International journal of environmental research and public health 17, 2749.

Nir, N., Knitter, D., Hardt, J., Schütt, B., 2021. Human movement and gully erosion: Investigating feedback mechanisms using Frequency Ratio and Least Cost Path analysis in Tigray, Ethiopia. PloS one 16, e0245248.

O'brien, R.M., 2007. Quality & Quantity. A caution regarding rules of thumb for variance inflation factors 41, 673–690.

Oliveira, G.G., Pedrollo, O.C., Castro, N.M., 2015. Simplifying artificial neural network models of river basin behaviour by an automated procedure for input variable selection. Engineering Applications of Artificial Intelligence 40, 47–61.

Paola, J.D., Schowengerdt, R.A., 1995. A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. International Journal of remote sensing 16, 3033–3058.

Park, S., Im, J., Jang, E., Rhee, J., 2016. Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. Agricultural and forest meteorology 216, 157–169.

Parvin, F., Ali, S.A., Calka, B., Bielecka, E., Linh, N.T.T., Pham, Q.B., 2022. Urban flood vulnerability assessment in a densely urbanized city using multi-factor analysis and machine learning algorithms. Theoretical and Applied Climatology 1–21.

Paul, G.C., Saha, S., 2019. Spatial prediction of susceptibility to gully erosion in Jainti River basin, Eastern India: a comparison of information value and logistic regression models. Modeling Earth Systems and Environment 5, 689–708.

Pham, B.T., Nguyen-Thoi, T., Qi, C., Van Phong, T., Dou, J., Ho, L.S., Van Le, H., Prakash, I., 2020. Coupling RBF neural network with ensemble learning techniques for landslide susceptibility mapping. Catena 195, 104805.

Pham, Q., Pal, S., Chakrabortty, R., Norouzi, A., Golshan, M., Ogunrinde, T., Janizadeh, S., Khedher, K., Tran Anh, D., 2021. Evaluation of various boosting ensemble algorithms for predicting flood hazard susceptibility areas Evaluation of various boosting ensemble algorithms for predicting flood hazard susceptibility areas. Geomatics, Natural Hazards and Risk 12, 2607–2628. https://doi.org/10.1080/19475705.2021.1968510

Pham, Q.B., Ali, S.A., Bielecka, E., Calka, B., Orych, A., Parvin, F., \Lupikasza, E., 2022. Flood vulnerability and buildings' flood exposure assessment in a densely urbanised city: comparative analysis of three scenarios using a neural network approach. Natural Hazards 1–39.

Poesen, J., Nachtergaele, J., Verstraeten, G., Valentin, C., 2003. Gully erosion and environmental change: importance and research needs. Catena 50, 91–133.

Pourghasemi, H.R., Moradi, H.R., Fatemi Aghda, S.M., Gokceoglu, C., Pradhan, B., 2014. GIS-based landslide susceptibility mapping with probabilistic likelihood ratio and spatial multi-criteria evaluation models (North of Tehran, Iran). Arab J Geosci 7, 1857–1878. https://doi.org/10.1007/s12517-012-0825-x

Pourghasemi, H.R., Sadhasivam, N., Kariminejad, N., Collins, A.L., 2020. Gully erosion spatial modelling: Role of machine learning algorithms in selection of the best controlling factors and modelling process. Geoscience Frontiers 11, 2207–2219.

Pradhan, B., Oh, H.-J., Buchroithner, M., 2010. Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. Geomatics, Natural Hazards and Risk 1, 199–223. https://doi.org/10.1080/19475705.2010.498151

Quarteroni, A., Veneziani, A., 2003. Analysis of a geometrical multiscale model based on the coupling of ODE and PDE for blood flow simulations. Multiscale Modeling & Simulation 1, 173–195.

Quevedo, R.P., Maciel, D.A., Uehara, T.D.T., Vojtek, M., Renno, C.D., Pradhan, B., Vojtekova, J., Pham, Q.B., 2021. Consideration of spatial heterogeneity in landslide susceptibility mapping using geographical random forest model. Geocarto International 1–24.

Rahmati, O., Haghizadeh, A., Pourghasemi, H.R., Noormohamadi, F., 2016. Gully erosion susceptibility mapping: the role of GIS-based bivariate statistical models and their comparison. Natural Hazards 82, 1231–1258.

Rahmati, O., Kalantari, Z., Ferreira, C.S., Chen, W., Soleimanpour, S.M., Kapović-Solomun, M., Seifollahi-Aghmiuni, S., Ghajarnia, N., Kazemabady, N.K., 2022. Contribution of physical and anthropogenic factors to gully erosion initiation. Catena 210, 105925.

Rahmati, O., Tahmasebipour, N., Haghizadeh, A., Pourghasemi, H.R., Feizizadeh, B., 2017. Evaluating the influence of geo-environmental factors on gully erosion in a semi-arid region of Iran: An integrated framework. Science of The Total Environment 579, 913–927. https://doi.org/10.1016/j.scitotenv.2016.10.176

Roy, J., Saha, S., 2021. Integration of artificial intelligence with meta classifiers for the gully erosion susceptibility assessment in Hinglo river basin, Eastern India. Advances in Space Research 67, 316–333.

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

Saha, S., Roy, J., Arabameri, A., Blaschke, T., Bui, D.T., 2020. Machine learning-based gully erosion susceptibility mapping: A case study of eastern India. Sensors (Switzerland) 20. https://doi.org/10.3390/s20051313

Saha, S., Roy, J., Pradhan, B., Hembram, T.K., 2021. Hybrid ensemble machine learning approaches for landslide susceptibility mapping using different sampling ratios at East Sikkim Himalayan, India. Advances in Space Research 68, 2819–2840.

Scherr, S.J., 2000. A downward spiral? Research evidence on the relationship between poverty and natural resource degradation. Food policy 25, 479–498.

Shafizadeh-Moghadam, H., Tayyebi, A., Helbich, M., 2017. Transition index maps for urban growth simulation: application of artificial neural networks, weight of evidence and fuzzy multi-criteria evaluation. Environmental monitoring and assessment 189, 1–14.

Shit, P.K., Bhunia, G.S., Pourghasemi, H.R., 2020. Gully erosion susceptibility mapping based on bayesian weight of evidence, in: Gully Erosion Studies from India and Surrounding Regions. Springer, pp. 133–146.

Sissakian, V., Al-Ansari, N., Knutsson, S., 2013. Sand and dust storm events in Iraq. Journal of Natural Science 5, 1084–1094.

Tairi, A., Elmouden, A., Bouchaou, L., Aboulouafa, M., 2021. Mapping soil erosion–prone sites through GIS and remote sensing for the Tifnout Askaoun watershed, southern Morocco. Arabian Journal of Geosciences 14, 1–22.

Tien Bui, D., Shirzadi, A., Shahabi, H., Chapi, K., Omidavr, E., Pham, B.T., Talebpour Asl, D., Khaledian, H., Pradhan, B., Panahi, M., 2019. A novel ensemble artificial intelligence approach for gully erosion mapping in a semi-arid watershed (Iran). Sensors 19, 2444.

Turner, B.L., Menendez III, H.M., Gates, R., Tedeschi, L.O., Atzori, A.S., 2016. System dynamics modeling for agricultural and natural resource management issues: Review of some past cases and forecasting future roles. Resources 5, 40.

Vandekerckhove, L., Poesen, J., Govers, G., 2003. Medium-term gully headcut retreat rates in Southeast Spain determined from aerial photographs and ground measurements. CATENA, Gully Erosion and Global Change 50, 329–352. https://doi.org/10.1016/S0341-8162(02)00132-7

Wang, Z., Zhang, G., Wang, C., Xing, S., 2022. Assessment of the gully erosion susceptibility using three hybrid models in one small watershed on the Loess Plateau. Soil and Tillage Research 223, 105481.

Wassie, S.B., 2020. Natural resource degradation tendencies in Ethiopia: a review. Environmental systems research 9, 1–29.

Yang, A., Wang, C., Pang, G., Long, Y., Wang, L., Cruse, R.M., Yang, Q., 2021. Gully erosion susceptibility mapping in highly complex terrain using machine learning models. ISPRS International Journal of Geo-Information 10. https://doi.org/10.3390/ijgi10100680

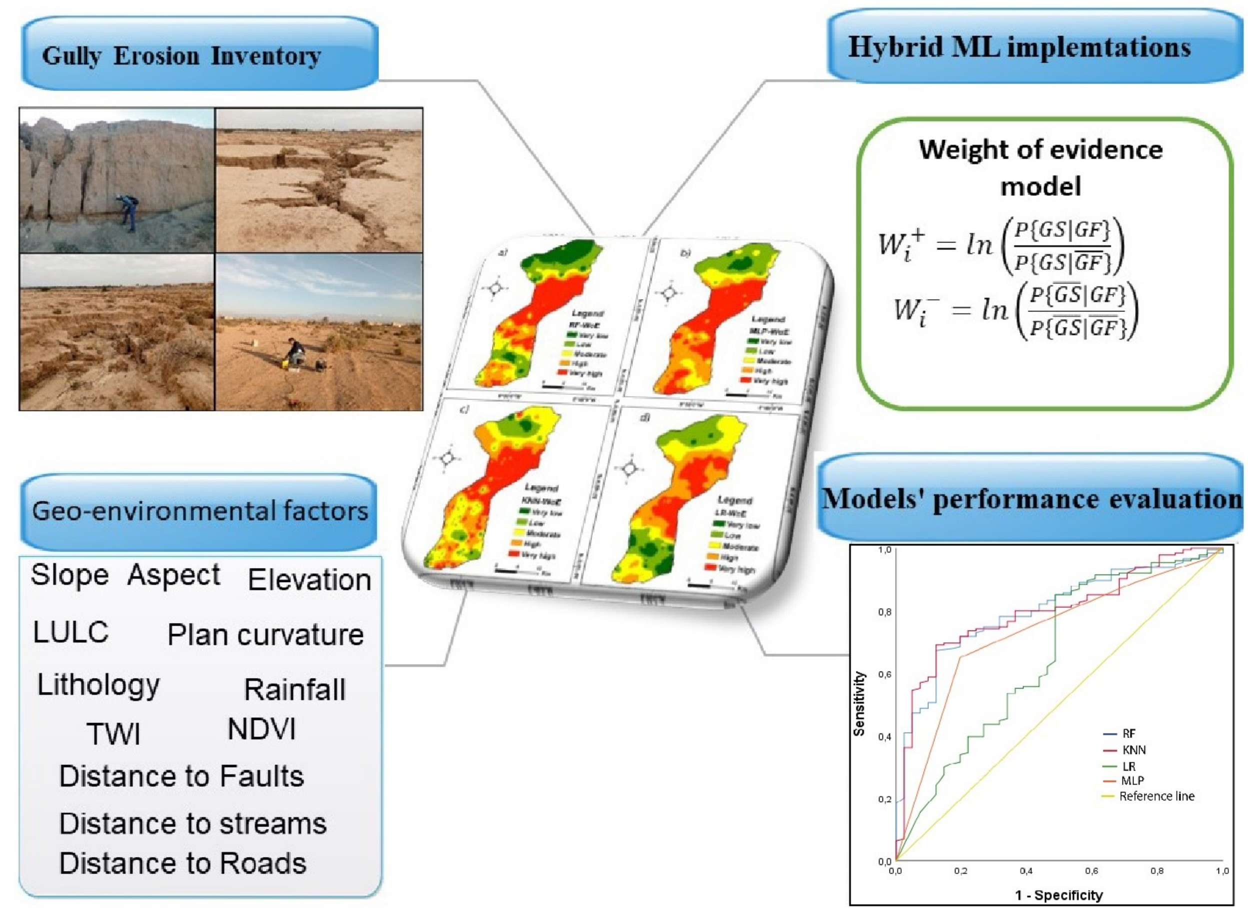
Yin, J., Su, S., Xun, J., Tang, T., Liu, R., 2020. Data-driven approaches for modeling train control models: Comparison and case studies. ISA Transactions 98, 349–363. https://doi.org/10.1016/j.isatra.2019.08.024

Yuan, L., Sinshaw, T., Forshay, K.J., 2020. Review of Watershed-Scale Water Quality and Nonpoint Source Pollution Models. Geosciences 10, 25. https://doi.org/10.3390/geosciences10010025

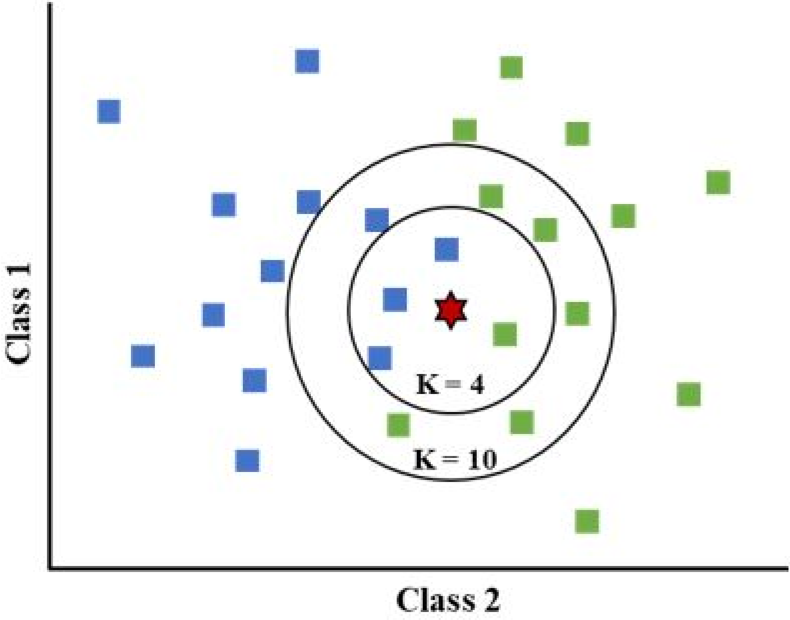
Zabihi, M., Mirchooli, F., Motevalli, A., Darvishan, A.K., Pourghasemi, H.R., Zakeri, M.A., Sadighi, F., 2018. Spatial modelling of gully erosion in Mazandaran Province, northern Iran. Catena 161, 1–13.

Zhang, S., Cheng, D., Deng, Z., Zong, M., Deng, X., 2018. A novel kNN algorithm with data-driven k parameter computation. Pattern Recognition Letters, Special Issue on Pattern Discovery from Multi-Source Data (PDMSD) 109, 44–54. https://doi.org/10.1016/j.patrec.2017.09.036

Zhou, Q., Liu, X., 2004. Analysis of errors of derived slope and aspect related to DEM data properties. Computers & Geosciences 30, 369–378.



Gully Erosion Inventory
Hybrid ML implemtations
Weight of evidence model
$W_i^+ = ln\left(\frac{P\{GS|GF\}}{P\{GS|\overline{GF}\}}\right)$
$W_i^- = ln\left(\frac{P\{\overline{GS}|GF\}}{P\{\overline{GS}|\overline{GF}\}}\right)$
Geo-environmental factors
Slope
Aspect
Elevation
LULC
Plan curvature
Lithology
Rainfall
TWI
NDVI
Distance to Faults
Distance to streams
Distance to Roads
Models' performance evaluation
Sensitivity
1 - Specificity
0,0
0,2
0,4
0,6
0,8
1,0
RF
KNN
LR
MLP
Reference line


K = 4
K = 10
Class 1
Class 2