Journal of Hydrology 618 (2023) 129229

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

# Soil water erosion susceptibility assessment using deep learning algorithms

Khabat Khosravi [a], Fatemeh Rezaie [b,c], James R. Cooper [d,*], Zahra Kalantari [e], Soroush Abolfathi [f], Javad Hatamiafkoueieh [g,*]

[a] Department of Earth and Environment, Florida International University, Miami, USA
[b] Geoscience Data Center, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124 Gwahak-ro, Yuseong-gu, Daejeon 34132, Republic of Korea
[c] Department of Geophysical Exploration, Korea University of Science and Technology, 217 Gajeong-ro, Yuseong-gu, Daejeon 34113, Republic of Korea
[d] Department of Geography & Planning, School of Environmental Sciences, University of Liverpool, Liverpool, UK
[e] Department of Sustainable Development, Environmental Science and Engineering (SEED), KTH Royal Institute of Technology, Stockholm, Sweden
[f] School of Engineering, University of Warwick, CV4 7AL Coventry, UK
[g] Department of Mechanics and Control Processes, Academy of Engineering, Peoples' Friendship University of Russia (RUDN University), Miklukho-Maklaya Str. 6, Moscow 117198, Russian Federation

ARTICLE INFO

This manuscript was handled by Marco Borga, Editor-in-Chief, with the assistance of Francesco Comiti, Associate Editor

*Keywords:*
Soil erosion
Deep learning
Land degradation
CNN
RNN
LSTM

ABSTRACT

Accurate assessment of soil water erosion (SWE) susceptibility is critical for reducing land degradation and soil loss, and for mitigating the negative impacts of erosion on ecosystem services, water quality, flooding and infrastructure. Deep learning algorithms have been gaining attention in geoscience due to their high performance and flexibility. However, an understanding of the potential for these algorithms to provide fast, cheap, and accurate predictions of soil erosion susceptibility is lacking. This study provides the first quantification of this potential. Spatial predictions of susceptibility are made using three deep learning algorithms – Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Long-Short Term Memory (LSTM) – for an Iranian catchment that has historically experienced severe water erosion. Through a comparison of their predictive performance and an analysis of the driving geo-environmental factors, the results reveal: (1) elevation was the most effective variable on SWE susceptibility; (2) all three developed models had good prediction performance, with RNN being marginally the most superior; (3) maps of SWE susceptibility revealed that almost 40 % of the catchment was highly or very highly susceptible to SWE and 20 % moderately susceptible, indicating the critical need for soil erosion control in this catchment. Through these algorithms, the soil erosion susceptibility of catchments can potentially be predicted accurately and with ease using readily available data. Thus, the results reveal that these models have great potential for use in data poor catchments, such as the one studied here, especially in developing nations where technical modeling skills and understanding of the erosion processes occurring in the catchment may be lacking.

## 1. Introduction

Soil water erosion (SWE) is a major cause of global land degradation and soil loss (Tang et al., 2015) through its negative impact on the organic, physical and chemical characteristics of soils (Aslam et al., 2021). Various parts of the world are under the hazard of SWE, and these areas are growing in extent as a result of climate and land use change, with areas extending beyond those of arid and semi-arid regions into humid regions (Boudjemline and Semar, 2018; Wijitkosum, 2021).

Soil is eroded by water through processes which vary in temporal and spatial scales. These processes are splash, interrill, rill and gully erosion, in which particles are detached by raindrop impact, unconcentrated flow or concentrated flow, and transported via rainsplash, interrill flow or concentrated flow (Cooper et al., 2012). These processes cause direct and indirect damage to ecosystems services through soil nutrient loss, decreasing soil productivity and reducing food production. Thus SWE is often the main driver for a number of global social and economic issues such as food insecurity, higher food prices and loss of biodiversity (Phinzi et al., 2021). Furthermore, SWE can have considerable off-site impacts, including decreasing water quality through increasing water turbidity, leading to irreparable damage to the aquatic system, reducing dam reservoir capacity, changing river morphology

* Corresponding authors.
*E-mail addresses:* khabat.khosravi@gmail.com, kkhosrav@fiu.edu (K. Khosravi), rezaie@kigam.re.kr (F. Rezaie), james.cooper@liverpool.ac.uk (J.R. Cooper), zahrak@kth.se (Z. Kalantari), Soroush.Abolfathi@warwick.ac.uk (S. Abolfathi), khatamiafkuiekh-d@rudn.ru (J. Hatamiafkoueieh).

https://doi.org/10.1016/j.jhydrol.2023.129229
Received 20 September 2022; Received in revised form 13 January 2023; Accepted 2 February 2023
Available online 6 February 2023

and increasing flood risk.

One country in which these negative impacts of SWE are being felt severely is Iran, the focus for this paper. Iran is a semi-arid country, covering a land area of 165 million hectares. The country experiences a mean annual soil erosion rate of 25 to 30 ton/ha/year, over 20 times higher than the mean global rate and four times higher than any other country (Afshar et al., 2010; Khalili Moghadam et al., 2015; Sadeghi, 2017). On average, 1 to 5 million tons of erosion takes places annually over the country's land mass (Mohammadi et al., 2021). Thus, although 50 million hectates of land is available for agriculture, only 1.3 million hectares is classed as being suitable for growing crops (Laylin, 2018). SWE is causing 400 million $m^3$ of sedimentation each year in reservoirs, resulting in an annual volume reduction of 0.5 % in reservoir capacity (Emadodin et al., 2012; Sadeghi, 2017) at an annual cost of US $0.6 billion (World Bank, 2005). This sedimentation reduces power generation and land irrigation capacity, further reducing the availability of arable lands. As a consequence, these combined impacts of SWE are costing the country US $56 to US $112 billion every year, higher than the revenue generated by oil production (Sadeghi, 2017).

Enhanced rates of SWE in Iran are due to the combined effects of over-exploitation by communities and industry, and changes in climate patterns (Sadeghi, 2017). This expliotation includes improper and unnecessary infrastructure development, land use changes and unlawful exploitation of resources. An intensification in agricultural activities over the past five decades has took place in tandeem with this unsustainable land use change. From the 1950 s until 2008, around 5 million hectares of forest was converted to farmland, pasture, and urban areas (Emadodin and Bork, 2012). There are also a number of indirect causes for these high SWE rates, including an inattention to soil value, a small number of hydrometry stations, limited short-term studies of erosion, unreliable monitoring data, a misunderstanding of the SWE processes and apathy (Sadeghi, 2017). All of these indirect causes have combined to mask the severity and intensity of the problem. The consequence is that soil is lost due to erosion approximately 19 times faster than it forms (Emadodin et al., 2012). Nearly $75 million is thought to be needed annually in soil improvement projects to reduce SWE (Akbari, 2017). Thus, in Iran, like many other countries experiencing the impacts of SWE, there is a critical need to target soil erosion control to those areas most susceptible, as well to identify areas most suitable for sustainable agricultural development.

Modeling plays a key role in identifying these areas, understanding the factors that lead to high erosion susceptibility and in testing the efficacy of soil erosion control strategies. Thus far, three main types of approaches have been used: empirical, physically-based and data-driven/machine learning models (Raza et al., 2021). Empirical models, such as the Universal Soil Loss Equation (Wischmeier and Smith, 1965), and its associated family of derivative models (e.g. Revised USLE (RUSLE; Renard et al., 1997), Modified USLE (MUSLE; Williams, 1975), and the modified Mediterranean Desertification and Land Use (MEDALUS; Abuzaid et al., 2021) estimate SWE using a prefixed set of physical parameters representing the main factors thought to affect erosion (Conoscenti et al., 2008). These models have a linear and simple structure and do not aim to simulate the physical processes of SWE. Instead, they are based on producing a mathematical algorithm that best describes the relationship between these parameters and measured erosion rates. These empirical models can perform well for the conditions upon which they are calibrated, but typically perform poorly for conditions outside those used in calibration (e.g. Cerdan et al., 2010; Kinnell, 2010; Rapp, 1963; Tan et al., 2018; Zhang et al., 1996).

Furthermore, they do not simulate soil deposition (e.g., sedimentation) and, in most cases, insufficient measured data exist to rigorously determine the single factors for all needed situations and scenarios (Auerswald et al., 2006; Wischmeier and Smith, 1978, 1965).

Physically-based models, such as Watershed Erosion Prediction Project (WEPP; Nearing et al., 1989) and KINematic runoff and EROSion (KINEROS; Woolhiser et al., 1990), attempt to mathematically represent the physics behind the processes of detachment, transportation and deposition. These models usually require many variables and detailed spatial and temporal catchment data for the model build, calibration and validation. Such information is rarely available in developing countries because the monitoring required, especially for large catchments, is costly and time consuming (Conoscenti et al., 2008). For example, in Iran, many of the hydrometric stations only measure sediment concentration for a few days during some severe storms, and the temporal pattern of sediment yield is rarely provided (Darvishan et al., 2010). Furthermore, physically-based simulations are complex, time-consuming and require highly-skilled end users that do not always reside in these countries. This issue has limited their use and success in Iran (Akhavan et al., 2010; Amiri, 2010). Moreover, concerns have been raised about their tendency to overestimate small runoff events and underestimate large runoff amounts, even when calibrated (Kinnell, 2017), and the physical basis of many commonly used modes has been questioned (Wainwright et al., 2008).

Recently, development of machine learning (ML) approaches have opened up new and exciting ways to predict environmental behavior, including soil erosion susceptibility modeling (e.g., Mosavi et al., 2020; Vu et al., 2020). These approaches have a non-linear structure and seek to find a robust relationship between input and output readily available parameters. ML techniques do not seek to explain the physical processes and mathematical reasoning for changes in environmental behavior but to recognize patterns, both expected and unexpected, within data. These patterns can highlight environmental relationships in space and time that may unveil critical details about behavior, reveal previously unsuspected relationships, or mitigate uncertainty in estimates. Furthermore, ML models are insensitive to missing data. Thus, these types of techniques are at their most beneficial in situations when physically-based models cannot be applied (e.g. lack of understanding of the underlying physics of the process) or that suffer from inadequacies due to the limitation of data. Therefore, the key advantage of ML approaches is that some parameters which are difficult or expensive to measure, such as soil erosion, could be easily predicted using other readily available factors, such as rainfall and those gained from satellite data.

Artificial Neural Network (ANN) is the oldest and most widely used ML technique in geoscience due to its computational efficiency (Abrahart et al., 2012). For example, Ebtehaj and Bonakdari (2013) applied ANN algorithms to predict sediment transport in sewers, revealing that ANN had a higher prediction power than existing empirical transport formulas. Similar results have also been found within other areas of hydrology, such as river suspended load prediction and evaporation modeling (Kisi et al., 2016; Melesse et al., 2011). However, these algorithms have slow coverage speed during the training procedure, high errors in the modeling phase, and a low convergence and generalization power (Kisi et al., 2012). Thus, ANN algorithms have poor prediction power when the range of the testing dataset is outside of the range of the training data (Kisi et al., 2016; Melesse et al., 2011), and they require a long-term dataset to achieve a reasonable result. Thus, to solve this weakness, ANN algorithms have been ensembled with fuzzy logic algorithms to create Adaptive Neural Fuzzy Inference System (ANFIS) models. ANFIS models take the advantages from both ANN and fuzzy logic methods to generally produce a better predictive capability (Ebtehaj and Bonakdari, 2014) but they are poor at finding the best weight parameters which heavily influence prediction accuracy (Tien Bui et al., 2016). Furthermore, ANFIS algorithm suffer from the need for a large number of model operators, each of which need to be set accurately, especially the weights of membership function. Also ANFIS algorithms lack a systematic approach in the design of fuzzy rules and in the choice of membership functions variables (Khosravi et al., 2018; Tien Bui et al., 2016).

Another powerful ML technique commonly used is support vector machine (SVM). This algorithm has been used in a range of applications in hydrology, such as rainfall-runoff forecasting (e.g. Dibike et al., 2001) and suspended sediment load prediction (Çimen, 2008). The developed

models have a high prediction power (Ganguli and Reddy, 2014; Goyal et al., 2014), but suffer from having a lot of hyper-parameters making model implementation difficult (Ahmad et al., 2018).

Group Method of Data Handling (GMDH) is another powerful and flexible neuron-based algorithm. The GMDH algorithm relates to the deterministic self-organizing method group, where the principle of a black box, connectionism and induction is used (Anastasakis and Mort, 2001). Applications include land subsidence susceptibility mapping (Panahi et al., 2022) and flood modeling (Dodangeh et al., 2020). However the weakness of the GMDH algorithm lies in its fixed configuration, using a deterministic approach to find the optimal partition of datasets and parameters (Robinson, 1998).

Although these ML models have been applied to model behaviors in a wide range of environmental settings, they all suffer from the need to determine accurate weights in the membership function and the optimal values for hyper-parameters. Since trial and error is not possible to determine the exact weights, bio-inspired algorithms have been applied to optimize these ML techniques. For example, Angileri et al. (2016) applied the stochastic gradient tree boost (SGT) for SWE modeling in Italy, revealing excellent reliability accuracy. Sajedi-Hosseini et al. (2018) applied the Fuzzy Decision-Making Trial and Evaluation Laboratory (Fuzzy DEMATEL) approach for SWE modeling and mapping at Noor-Rood catchment in Iran and found this approach had a reasonable prediction power. In the same catchment, Mosavi et al. (2020) utilized several machine learning models – weighted subspace random forest (WSRF), Gaussian process with a radial basis function kernel (Gausspr-radial) and Naive Bayes (NB) – for SWE susceptibility mapping. Their results showed the WSRF model outperformed the other models, followed by Gausspprradial, and NB. Yousefi et al. (2021) applied three ML models of Random Forest (RF), classification and regression tree (CART), and SVM to model land degradation in Alborz Mountains, Iran; RF outperformed the CART and SVM models. Although tree-based models have higher performance than ANN, ANFIS and SVM models, trees are sensitive to noisy data (Tien Bui et al., 2012) making them less suitable to catchments that lack continuous erosion monitoring or where this monitoring is sparse in spatial extent.

Since these studies were conducted, a new type of ML model has been developed, namely deep learning algorithms. These algorithms have a greater flexibility than traditional ML models and thus a higher predictive performance (Ghorbanzadeh et al., 2019). Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Long-Short Term Memory (LSTM) are amongst the popular deep learning algorithms. Examples of application of the CNN algorithm include flood modeling (Khosravi et al., 2020) and landslide susceptibility assessment (Thi Ngo et al., 2021). The RNN algorithm has been applied for landslide susceptibility assessment (Li et al., 2021), and LSTM for flood modeling (Fang et al., 2021). A significant gap exists in understanding the potential of these deep learning algorithms, and in the identification of the most flexible and accurate algorithm for SWE susceptibility prediction.

The present paper, therefore, aims to fill this gap in understanding by achieving the following objectives, using the heavily eroded Noor-Rood catchment in Iran as a case study: (1) Delineate SWE susceptibility areas using three deep learning algorithms techniques, namely CNN, RNN and LSTM; (2) compare the predictive power of these data-driven models; and (3) perform an analysis of effectiveness of the geo-environmental driving variables through Information Gain Attribute Evaluator (IGAE) feature selection. The performance of these particular algorithms is tested for the following reasons: (1) a CNN model automatically detects the important features without any human supervision. (2) RNN can process inputs of any length; and (3) LSTM that enables a model to "remember" past information. The research offers new insight into which deep learning algorithms offer the potential to provide relatively cheap and fast predictions of SWE susceptibility in situations when understanding of the physical processes at play may not be well understood, such as the catchment studied here.

## 2. Study area

Noor-Rood catchment is one of the main sub-catchments of the Haraz watershed, located southwest of Amol city, Mazandaran province, Northern Iran (Fig. 1; Eastern longitudes 51° 26″ 13″ to 52° 18″ 21″ and Northern latitudes 36° 00″ 58″ to 36° 16″ 18″). The catchment is fully mountainous (elevation between 721 and 4333 m), has an area of 1298 $km^2$ and experiences a mean annual precipitation of 504 mm (Solaimani and Hadian Amr, 2008). The study area is mostly covered by the central Elborz and Shemshak (shale, marl and sandstone) and Karaj (tuff and shale) formations. Historically the catchment has suffered intense SWE area due to flooding, and human-induced processes (e.g., land use change, over-grazing, intensive agriculture and etc.) causing high turbidity in Haraz River, high sedimentation and a decrease in water capacity in the Haraz dam, located downstream of the catchment. Since the building of large dams in the Haraz watershed are costly and of critical importance to irrigation and energy generation, delineating SWE prone-areas in the study area is necessary for aiding the location of new dams and mitigating the impacts of SWE on existing dam capacity.

## 3. Methodology

The conceptual framework for modeling SWE susceptibility is shown in Fig. 2. Four main steps are used:

1. Data collection of SWE historical data;
2. Extraction of SWE geo-environmental variables;
3. Generating SWE susceptibility maps using deep learning CNN, LSTM and RNN algorithms;
4. Model evaluation and comparing the results across algorithms using the area under the receiver operating characteristic curve (AUC).

### 3.1. SWE inventory map

Spatial modeling through the machine learning approach has a binary classification system (i.e., SWE occurrence/non-occurrence). The first and vital step in this binary modeling was determining the spatial relationship between SWE historical data and geo-environmental factors in order to determine the effectiveness of each on SWE susceptibility. The historical data was derived from a survey of soil erosion in 2017 by Sajedi-Hosseini et al. (2018), in which recordings of sheet, rill, interrill, gully erosion, and mass movements were made at 116 locations, (Fig. 1); and equally occurrences of no erosion were recorded at 116 locations.

### 3.2. Construction of the training and testing datasets

In the next step, values of 0 and 1 were allocated to locations that have experienced no SWE and SWE, respectively. Then, all data was split into two parts (Chung and Fabbri, 2003) in a ratio of 70:30 for training and testing, respectively, based on the most commonly used ratio. The first 70 % was used for model development, and the remaining 30 % for model evaluation. In the final step, the training dataset was overlaid with all the SWE geo-environment conditioning factors to extract their attribute values for modeling.

### 3.3. SWE geo-environmental factor

Based on a body of previous work on SWE susceptibility modeling (Sajedi-Hosseini et al. 2018; Aslam et al., 2021; Conoscenti et al., 2008; Mosavi et al., 2020; Rosskopf et al., 2020), the characteristics of the study area and data availability, 14 SWE geo-environmental conditioning factors were considered (Fig. 3a–n). After initial selection of these input factors, their effectiveness was investigated. Parameters with a low or null value can cause lower model performance. To identify if this was the case within the training dataset, an Information Gain

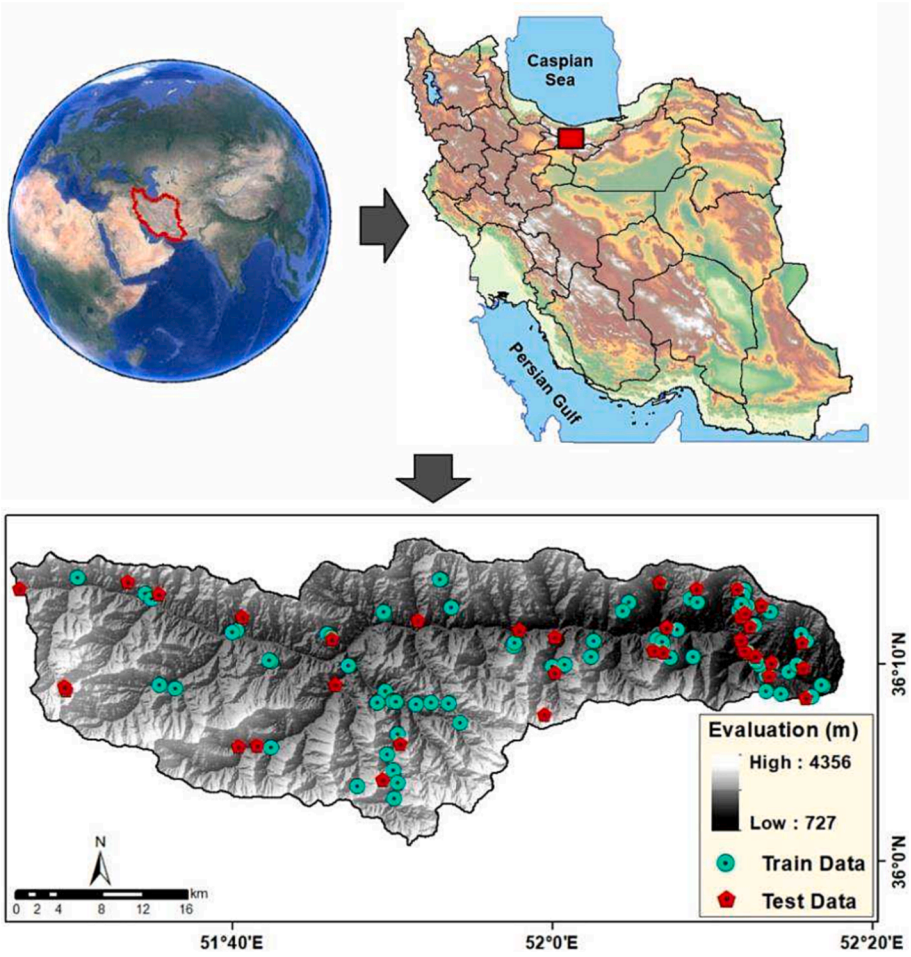


**Fig. 1.** The location of Noor-Rood catchment and the SWE sites used to train and test the deep learning models.

Attribute Evaluator (IGAE) feature selection technique was applied to determine input variable importance. The IGAE technique evaluates the worth of an attribute by measuring the information gain with respect to the class, as follows:

$$IGAE(Class, Attribute) = H(Class) - H(Class|Attribute) \quad (1)$$

where H is the information entropy. More information about this technique can be found in Novakovic (2009) and Trabelsi et al. (2017).

Through the use of this technique, all 14 geo-environmental factors were considered effective and all considered in the next step of model development. These factors were elevation, ground slope, slope aspect, plan curvature, topographic wetness index (TWI), stream power index (SPI), distance from river, slope length and steepness factor (LS factor), rainfall erosivity (RE), hydrologic soil group (HSG), Normalized Difference Vegetation Index (NDVI), land use, soil texture and geology. Ground slope, elevation, slope aspect, LS, plan curvature, TWI, and SPI were gained directly from ASTER Global Digital Elevation Model (DEM) of the study area with a 30 m × 30 m resolution (https://earthexplorer.usgs.gov/). The distance from river factor was determined in ArcGIS 10.2 using a digital river network. The RE factor was calculated from recorded rainfall in the study area using a rainfall gradient approach, due to the lack of rainfall intensity data, according to the following equation (Yu and Rosewell, 1996):

$$RE = 0.0483(P)^{1.61} \quad (2)$$

where $P$ is the mean annual precipitation (mm). The mean annual precipitation was estimated based on data recorded from 1976 to 2016 by six rain gauges in the catchment. Further information about the rainfall records can be found in Mosavi et al. (2020). HSG (Table A in Supplementary material) and soil texture data was acquired in a shape-file format from Mazandaran Regional Water Authority. NDVI and land use factors were measured using Landsat 8 OLI/TIRS (Operational Land Imager/Thermal Infrared Sensor) imagery from June 2016 (Sajedi-Hosseini et al., 2018). The geology dataset was obtained from the national geology map of Iran at a scale of 1:100,000 (Table B in Supplementary material).

### 3.4. Model description

#### 3.4.1. Convolutional Neural Network (CNN)

The CNN method has the structure of ordinary feed-forward neural networks, with each hidden layer made of neurons fully connected to all neurons in the previous and next layer. Neurons in a single layer operate completely independently and do not share any connections. Weight sharing and local connection are two special characteristics of the CNN model, enhancing the network training efficiency and improving its capability to effectively decrease the network's free parameters (Du et al., 2022).

The generalized architecture of the CNN model consists of three main types of layers: convolutional, pooling, and fully-connected layers. In the convolution layer, a linear operator, such as convolution, or a nonlinear activation function, such as rectified linear unit (ReLU; defined as $f(x) = \max(0, x)$; Yamashita et al., 2018) separates and identifies the various features of the dataset for analysis. These functions are able to capture the spatial and temporal dependencies between data. The primary aim of the pooling layer is to decrease the size of the

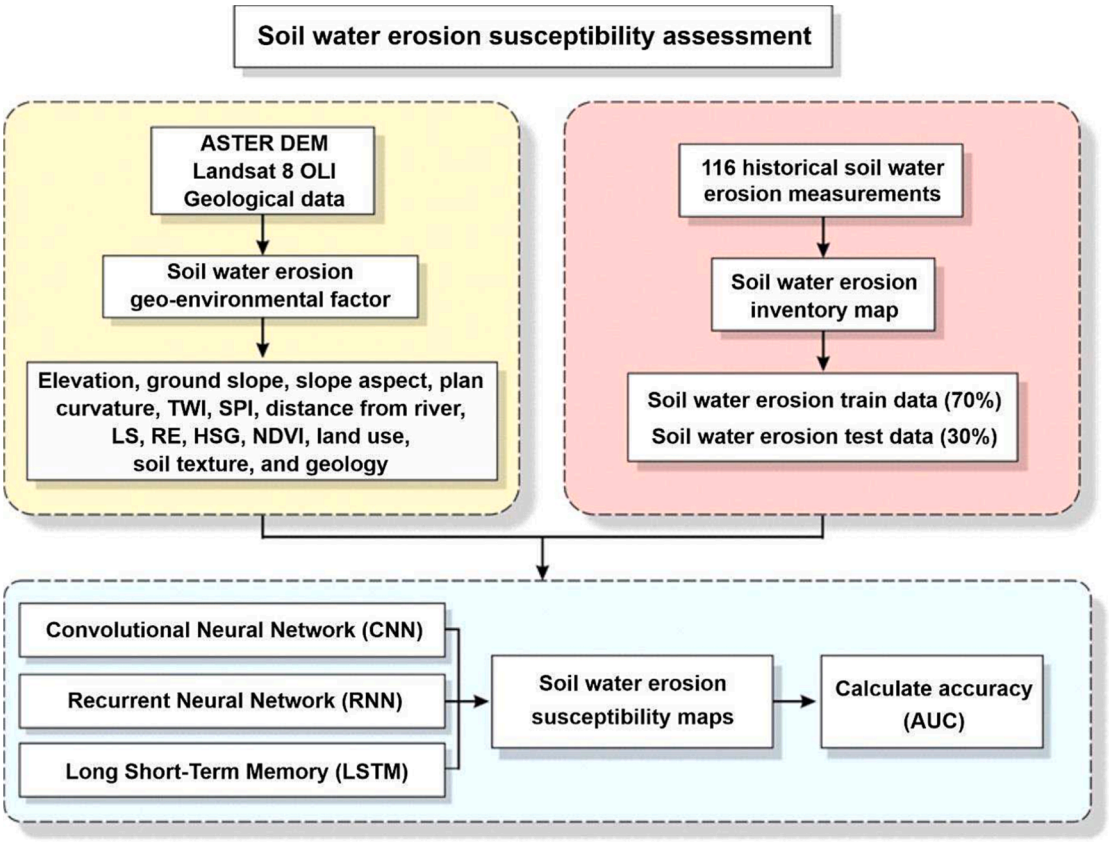


**Fig. 2.** Conceptual model development framework.

convolved feature map to reduce the computational costs, and to handle overfitting issues. This step is performed by decreasing the connections between layers, independently operating on each feature map. Max, Min and average pooling are the most common form of pooling layer used to reduce data size, whilst still preserving the most important features within them (Pally and Samadi, 2022). For example, max pooling, a nonlinear downsampling operation, computes the largest value in each patch of a feature map (Venkatappareddy et al., 2021). Thus, the pooling layer generalizes the features extracted by the convolution layer, and helps the networks to recognize the features independently. The fully-connected layer is the final part of CNN model which maps the extracted features into final output using a softmax activation function to convert the vector of numbers into the vector of probabilities. The output of the function reveals the probability of each class (Li et al., 2021).

### 3.4.2. Recurrent Neural Network (RNN)

RNN is a class of ANN model, derived from feed-forward networks (FFN), where connections between nodes form a directed or undirected graph along a temporal sequence. RNN algorithms are powerful and robust because they use their internal state (memory) to process variable length sequences of inputs. RNNs are best applied to sequential or time-series data, but by extracting the contextual information from the data, RNN models can be efficaciously applied to data classification. The RNN structure is composed of several successive recurrent layers. A FFN allocates a weight to the input parameters, similar to other deep learning models, but the key difference is that RNN algorithms, using the internal memory, assign this weight to both current and previous inputs.

RNN models benefit from having loops in hidden layers, and these loops have a significant effect on the training capability of the model. Feedback loops from the output layer ($X_1(t)$, …, $X_{nj}$) are evaluated using the networks input ($Y(t)$). For each new input, the output values are calculated and then sent back to the modeling process as an adjusted/modified input. This approach continues until a fixed and steady output is obtained. The mathematical approaches are as follows:

$$a_i(t) = \sum \varphi_{ji} Y_i(t) + \sum \sigma_{ji} h_i(t-1), j = 1, \ldots, n_H \quad (3)$$

$$h_i(t) = F(a_i(t)), j = 1, \ldots, n_H \quad (4)$$

$$b_i(t) = \sum a_{ji} h_i(t), j = 1, \ldots, n_j \quad (5)$$

$$X_i(t) = G(b_j(t)), j = 1, \ldots, n_j \quad (6)$$

where $\varphi_{ji}$, $\sigma_{ji}$ and $a_{ji}$ are weights, $F$ is defined as the non-linear transformation parameter, and $n_H$ is the hidden number of neurons. Generally, hidden layers through networks input [i.e., $Y(t)$] calculate the adjusted network weights.

### 3.4.3. Long-Short term memory (LSTM)

LSTM is another type of RNN, capable of learning order dependence in sequence prediction problems. This algorithm extends the internal memory and helps the model to recall events from a long period of time. Standard RNN models, due to the issue of gradient vanishing, are unable to prepare long memories. In contrast, LSTM models are able to train and learn from important events with lags of unknown duration between important events in a time series. LSTM uses three ways to allocate the weights: forget the information, let in new information and hold the information that affects the output. The model also has three main layers or gates: input (control the input information in a memory cell), output (maintain control over the outgoing information throughout the remainder of the networks) and forget (control the input from the previous memory and determine if it should be deleted based on the preceding cell condition). This information is calculated through the following equations:

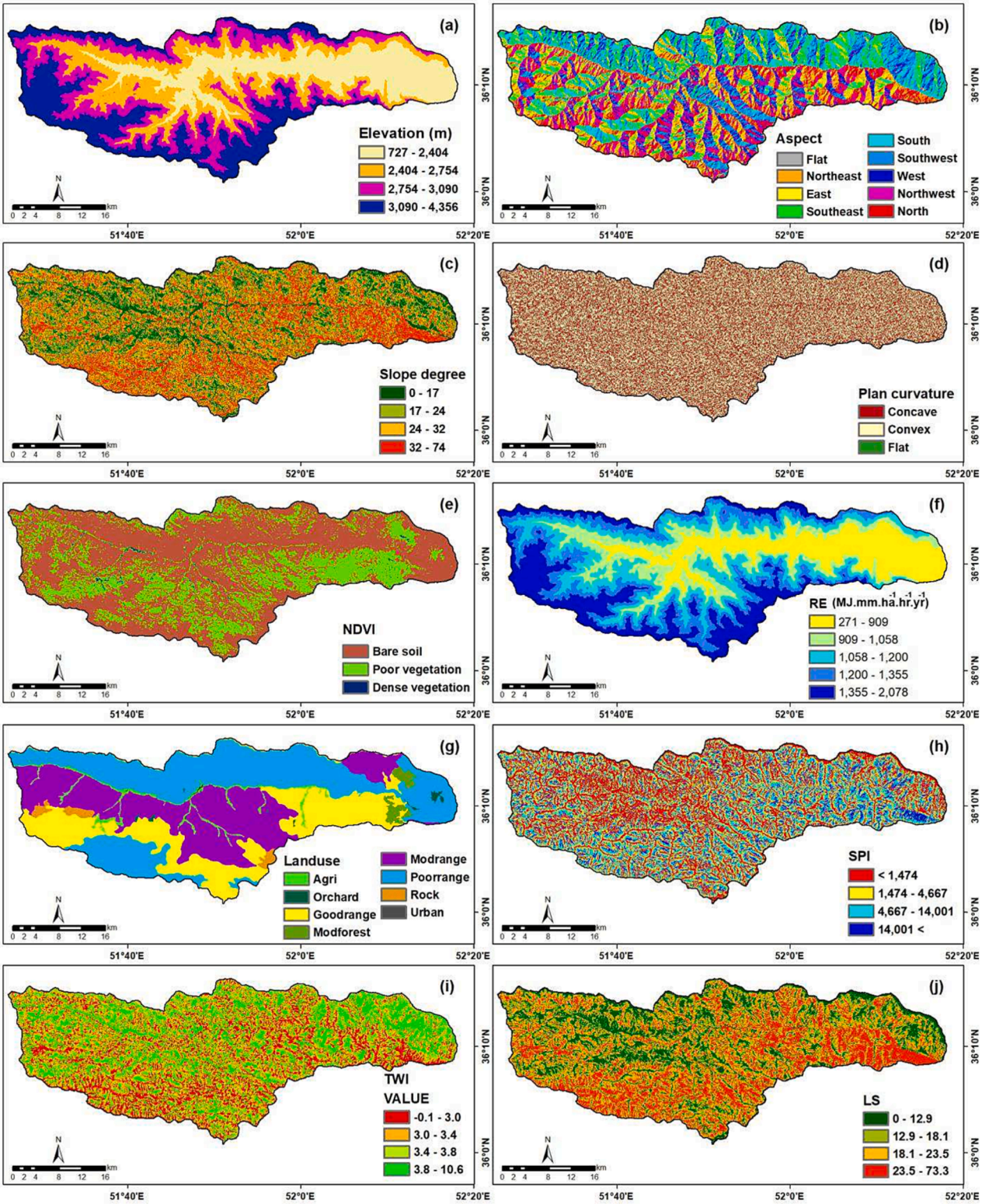


**Fig. 3.** Spatial variation in SWE geo-environmental factors used for model development, a) elevation, b) aspect, c) slope degree, d) plan curvature, e) NDVI, f) RE, g) landuse, h) SPI, i) TWI, j) LS, k) distance from river, l) hydrologic soil group, m) soil texture, and n) geology.

$$F_t = \sigma(W_{xf}X_t + W_{hf}H_{t-1} + B_f) \quad (7)$$

$$I_t = \sigma(W_{xi}X_t + W_{hi}H_{t-1} + B_i) \quad (8)$$

$$\overline{C}_t = \sigma(W_{xc}X_t + W_{hc}H_{t-1} + B_c) \quad (9)$$

$$C_t = F_t{*}C_{t-1} + I_t{*}\overline{C}_t \quad (10)$$

$$O_t = \sigma(W_{xo}X_t + W_{ho}H_{t-1} + B_o) \quad (11)$$

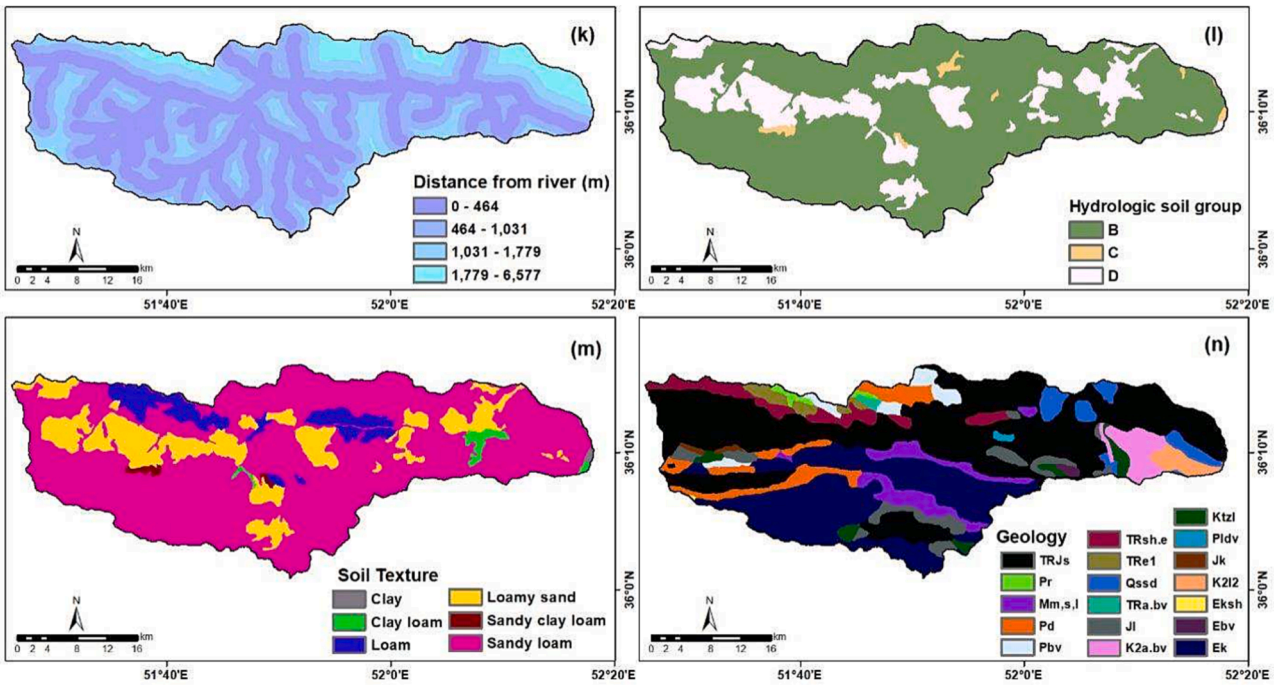


**Fig. 3.** (*continued*).

$$H_t = o_t \tanh(C_t) \tag{12}$$

$$Y_t = \sigma(W_{hy} H_t + B_y) \tag{13}$$

$$\sigma(x) = \frac{1}{1 + \exp^{-x}} \tag{14}$$

where $X_t$, $Y_t$, $I_t$, $F_t$, $O_t$, $C_t$, $\overline{C}_t$ and $\sigma$ are input vector, output vector, input gate, forget gate, output gate, finishing state in memory block, temporary and sigmoid function respectively. $W_{xf}, W_{xi}, W_{xc}$, and $W_{xo}$ are input weight matrices, $W_{hf}, W_{hi}, W_{hc}$, and $W_{ho}$ are recurrent weights matrices, $W_{hy}$ is the output weight and $B_f$, $B_i$, $B_c$, $B_o$ and $B_y$ are related bias vectors.

### 3.5. SWE susceptibility map generation

After developing the three models using the training data, they were used to calculate SWE indices (SWEI) for all pixels in the study area. These SWEI values were classified based on the quantile classification scheme to very low, low, moderate, high, and very high to produce a SWE susceptibility map.

### 3.6. Model performance evaluation

The prediction power of each of the developed models was evaluated quantitatively using the powerful and reliable receiver operating characteristic (ROC) curve method, using Statistical Package for the Social Sciences (SPSS) software. Model evaluation was assessed during both model development (success-rate ROC curve) using the training dataset, and during model evaluation using the testing dataset (prediction-rate ROC curve). Success-rate ROC curves only reveal the efficiency of the built model, whereas prediction-rate ROC curves, using data not used in model development, reveals how good the model is at prediction and thus shows the model generalization power. Area Under the ROC curve (AUC) measures the entire two-dimensional area underneath the ROC curve, providing an aggregate measure of performance across all possible classification thresholds. AUC varies between 0 (the lowest model prediction power) and 1 (ideal model with the highest performance). Table 1, based on Yesilnacar (2005), shows how ranges of AUC values can be interpreted for classifying model quality.

**Table 1**
Interpreting AUC values.

| AUC quantitative values | Quality interpretation |
|---|---|
| 0.5–0.6 | Unsatisfactory |
| 0.6–0.7 | Satisfactory |
| 0.7–0.8 | Good |
| 0.8–0.9 | Very good |
| 0.9–1 | Excellent |

## 4. Result and analysis

### 4.1. Effectiveness of input variables on SWE susceptibility

The effectiveness of the 14 input variables on SWE susceptibility, assessed using the IGAE feature selection technique, is presented in a Radar-chart (Fig. 4). This chart shows that elevation was the parameter with the highest effectiveness on SWE (0.30), followed by RE (0.16), NDVI (0.143), TWI (0.142), plan curvature (0.141), ground slope (0.12), geology (0.099), aspect (0.091), SPI (0.073), distance from river (0.070), land use (0.067), soil texture (0.030), LS (0.020), and HSG (0.016).

### 4.2. SWE susceptibility maps

The delineated SWE susceptibility maps showed the middle and eastern part of the catchment had a high and very high susceptibility to SWE, while the rest of the study area, which surrounded the susceptible areas, had a lower susceptibility (Fig. 5a-c).

According to the ROC model evaluation technique, all models, in terms of success rate and prediction rate, had a very good prediction power ($0.8 < AUC < 0.9$) (Fig. 6). All models had the same performance in the training phase ($AUC = 0.85$), but RNN was marginally superior in the testing phase ($AUC = 0.83$). Thus, the RNN model had the highest generalization power, albeit only slightly.

Differences in the three maps are subtle. The maps produced by the

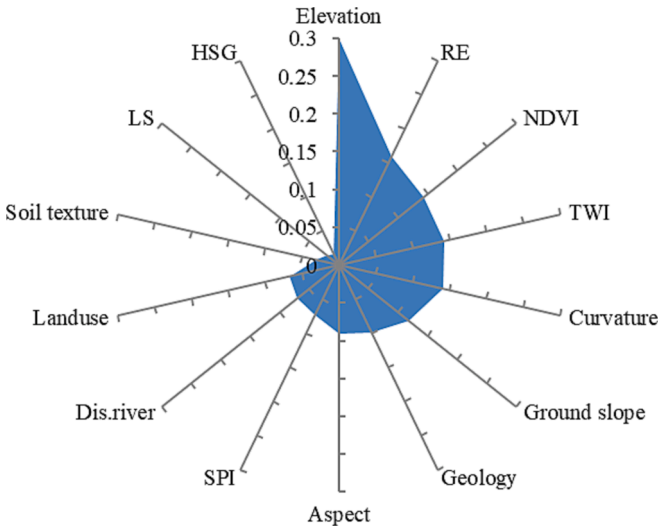


**Fig. 4.** Radar-chart showing variable effectiveness on SWE.

LSTM and RNN models were most alike, producing similar predictions for the percentage of catchment occupied by very low, high, very high susceptibility. The CNN model predicted a higher percentage of very high susceptibility areas than the other two models, but a slightly lower proportion of very low susceptibility (Fig. 7). The CNN model overestimated the very high susceptibility classes, while underestimated the moderate classes. The LSTM model overestimated the very low susceptibility classes while underestimated low and moderate susceptibility classes (Figs. 5 and 7). According to the predictions made by the RNN model, 19.8 % of the study area had a very high SWE susceptibility, 19.7 %, 20.8 %, 20.8 %, and 18.9 % had a high, moderate, low and very low susceptibility (Fig. 7). Overall, 39.5 % of the catchment had a high and very high susceptibility to SWE.

## 5. Discussion

### 5.1. Effectiveness of geo-environmental factors

To examine the effectiveness of each of the geo-environmental factors, a Frequency Ratio (FR) approach was used. FR is a bivariate statistical approach that calculates the probabilistic relationship between SWE susceptibility and each of the factors. The FR is the ratio of the area where SWE occurs to the total study area for a given attribute. For each range or type of factor, FR was calculated as follows:

$$FR = \frac{D_i / \sum_{i=1}^{N} D_i}{A_i / \sum_{i=1}^{N} A_i}$$

where $D_i$ is the area of SWE of the $i$-th category (very low, low, medium, high, very high), $A_i$ is the area of the $i$-th category for a certain factor and $N$ is the category number of the factor. A derived FR value of more than 1 indicates a strong and positive relationship between the concerned class of the selected factor data layer and high SWE susceptibility; on the other hand, a frequency ratio value of$<1$ suggests a poor and negative correlation between SWE susceptibility and the concerned class of the factor data layer, and low susceptibility.

The FR values and the elevation map (Fig. 3a) show that areas with lower and mid-elevations (727–2404 m) had a strong correlation with the highest SWE susceptibility, particularly due to mass movement. Areas with higher elevations may usually be thought to more susceptible, due to higher rainfall and steeper slope, thinner soils and bedrock outcrops. However precipitation falls as snow in these higher elevations. Since snow has less erosivity than rainfall and can acts as an armor to protect soil from erosion, these higher elevation areas therefore experience less erosion than the lower and mid-elevations. Thus RE had a poor correlation with SWE (Fig. 3f). Furthermore livestock farming is denser in low elevations, resulting in overgrazing, a reduction in vegetation cover (especially during early spring when immature is damaged), soil compaction, lower infiltration rates, lower soil retention and higher runoff, combining to make soil more susceptible to rain-splash, sheet, rill and gully erosion (Sternberg et al., 2000; Orgill et al., 2018; Narantsetseg et al., 2018; Donovan and Monaghan; 2021).

Areas with low NDVI, representing soil with little or no vegetation, were strongly positively correlated with SWE corresponding with higher susceptibility (Fig. 8). Whereas areas with high NVDI were negatively correlated because vegetation cover reduced the volume and intensity of rainfall reaching the soil surface (e.g. Vaezi et al., 2016) and the detachment of sediment in overland flow. In the long term, vegetation also increases soil organic matter, and improves soil physical properties (e.g. Puigdefabregas et al., 1999), further reducing soil erosivity.

High values of TWI were correlated strongly with SWE, since areas with high TWI can indicate areas where there is an accumulation of water, possibly due to the routing of overland flow, resulting in a higher likelihood of particle detachment through rilling and gullying. Flat slopes were also associated with areas with high SWE susceptibility, since, with all things constant (such as soil thickness and vegetation cover), lower angle slopes are most often associated with lower infiltration capacities, and thus higher rates of runoff (e.g. Chen and Young, 2006). Two lithological units, Pldv (Rhyolitic to rhyodacitic volcanic rocks) and K2l2 (Thick-bedded to massive limestone (Maastrichtian)) were highly correlated with areas of high SWE, with all other units displaying a weak association (Fig. 8). Slopes with different aspect receive differing amounts of solar radiation; southerly aspects correspond strongly with areas of high SWE, since these areas receive more solar radiation, reducing soil wetness, aggregate stability and vegetation cover, and more crusting of the soil surface, reducing infiltration capacities and increasing runoff (e.g. Marques and Mora, 1992; Fang and Guo, 2015). All other factors had a poor correlation with SWE susceptibility.

### 5.2. Comparison in model performance

Since inputs were the same for each of the models, the difference in performance between algorithms is attributable to their different computational structures. RNN and LSTM models outperformed the CNN model because they both have an internal memory that allows them to 'remember' important information about the input they receive, enhancing their precision in predicting what's coming next. Although the generalization ability of CNN is enhanced due to the use of a pooling layer, their fitting ability is reduced when the inputs are low-dimensional vectors, leading to information loss (Liu et al., 2019).

CNN and RNN algorithms operate similarly by introducing sparsity and reusing the same neurons and weights. However in the CNN model latent patterns in data are detected using convolution operations, whereas in the RNN model the specific sequences in input data are found by considering the relationship between the current and the previous state, allowing RNNs to feed results back into the network. Also, in CNN models, the input size and resulting output are fixed but in RNN models they vary. Further, the RNN algorithm benefits from a parameter sharing technique, allowing RNN models to share parameters across different time steps. In addition, the incorporation of convolutional layers can stretch its efficiency to nearby pixels.

The CNN and LSTM algorithms are feed forward neural networks, but a backpropagation training algorithm is used in the RNN model providing the ability to more effectively interpret temporal information and capture long-term dependencies in data, as well as obtain a minimum error function value (Ajitha et al., 2022). When these differences are taken together, this greater complexity in the RNN computational structure accounts for the enhanced predictive power of RNN models over CNN models (Mutlu et al., 2019).

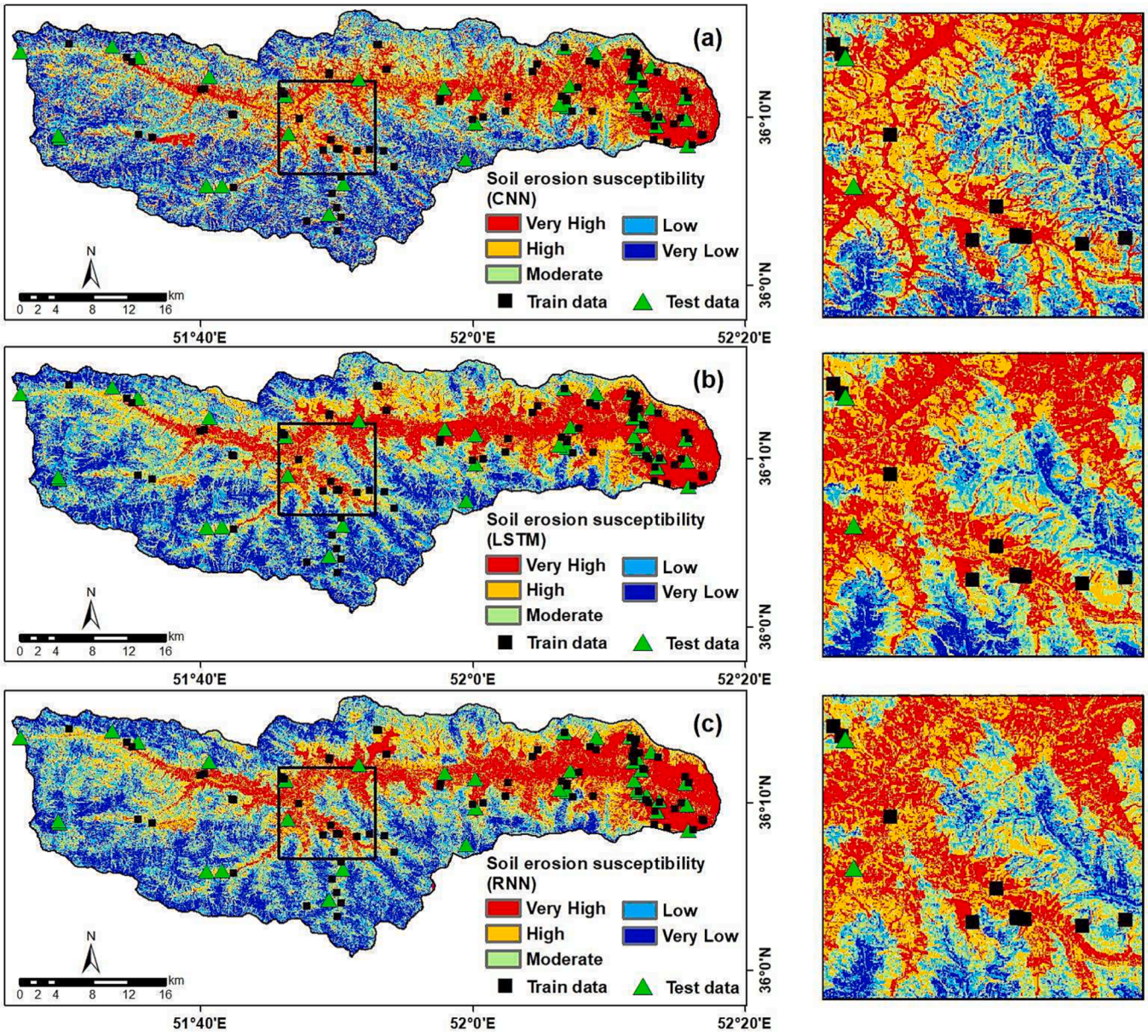


**Fig. 5.** SWE susceptibility maps based on the different deep learning models: a) CNN, b) LSTM and c) RNN.

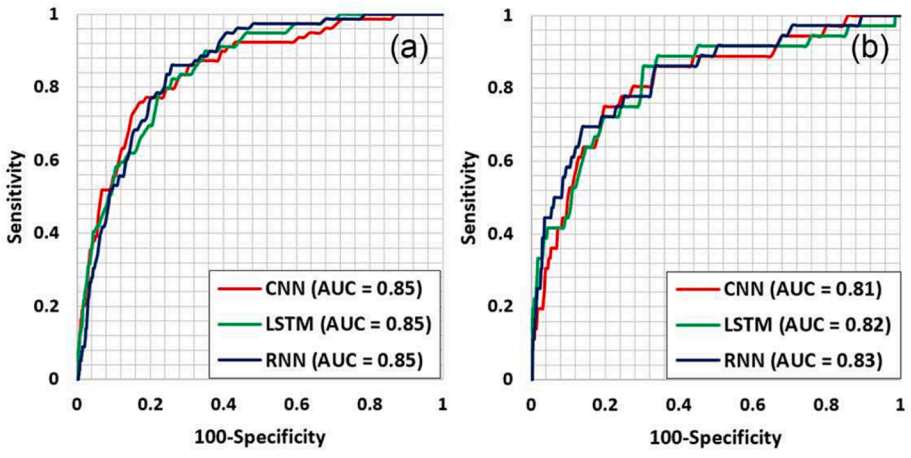


**Fig. 6.** ROC curves for (a) success rate and (b) prediction rate.

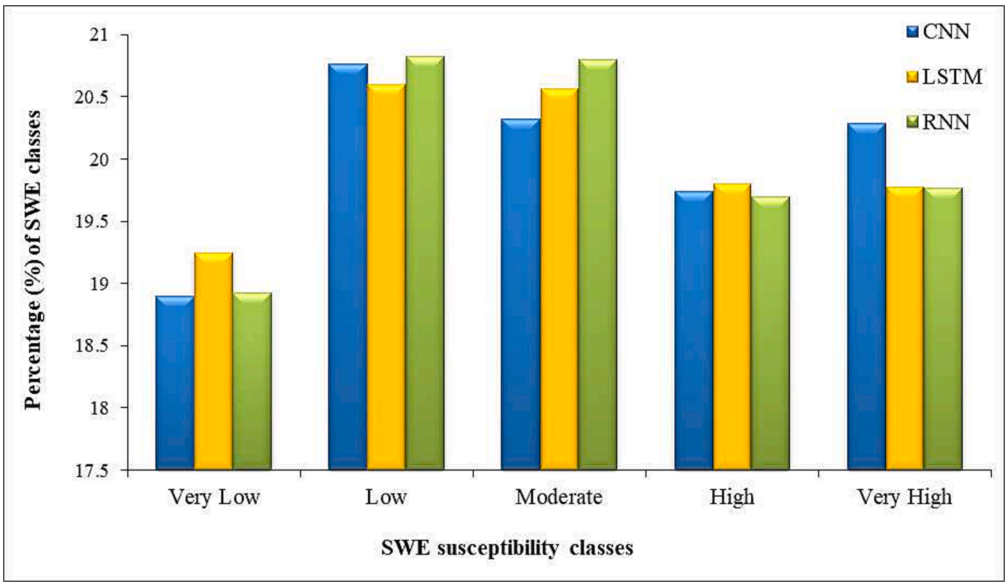


**Fig. 7.** A histogram showing the percentage of SWE classes that fall into the susceptibility class for each deep learning model.

A direct comparison between our findings with those of previous studies is not possible as deep learning algorithms have not yet to be applied for SWE modeling and mapping. Several studies have applied different ML models for SWE mapping however. Yousefi et al. (2021) applied three models of RF, CART and SVM model for SWE susceptibility mapping in a different catchment in Iran and reported these models have accuracies of between 0.89 and 0.96 based on the ROC-AUC method. These values are higher than reported in this study, despite the use of more powerful algorithms, resulting from the greater complexity in the relationships between the geo-environmental factors and erosion processes in the catchment studied in the present paper. In particular, snow in mountainous areas, high livestock density and deforestation in lowland areas, and little variation in soil grain size, resulted in some factors, such as RE, land use and soil texture, that would normally be well correlated with SWE to be not well correlated, and there was an unexpected relation with slope elevation. In the same catchment, Sajedi-Hosseini et al. (2018) applied a Fuzzy DEMATEL approach to SWE mapping with an accuracy of 0.83. Although their method has higher performance, the generated map had higher uncertainty, because the weights were calculated based on the expert opinion. In contrast, Mosavi et al. (2020) applied WSRF, Gausssprradial, and NB methods for SWE in the same study area and found predictive accuracy was lower; WSRF with Kappa index of 0.81 had a higher performance than Gausssprradial (0.76) and NB with Kappa (0.71). Also Abolhasani et al. (2022) implemented RF, boosted regression tree (BRT), SVM, and classification and regression tree (CART) algorithms, with AUC values of 0.81, 0.76, 0.71 and 0.63 for SWE mapping at Qazvin Plain, Iran. The three models developed in the current study performed more strongly.

### 5.3. *Applying deep learning models to map SWE susceptibility*

The choice of the 'best' predictive model is most often a compromise between model prediction accuracy and model complexity, with the later, in deep learning models, most closely related to the data input requirements. The major advantages of the deep learning models developed in this paper are their simplicity, and their ease and in expense to build and run, unlike physically-based models, whilst providing little compromise on model performance. In other words, the deep learning models provided good prediction performance based on inputs that are readily available from satellite imagery and national rainfall monitoring without the need for catchment or hillslope-scale monitoring. Thus, the results reveal that these models have great potential for use in SWE susceptibility assessment in data poor catchments, especially in developing nations where technical modeling skills and understanding of the hydrologic and erosion processes occurring in the catchment may be lacking.

The major disadvantages of these types of models however are twofold. First, like all statistical methods, the developed models only relate directly to the catchments being considered, and thus their application to other catchments may prove inappropriate. Future studies should apply the developed models to catchments with differing rainfall, soils, land use, land cover, geology and morphology to discover whether this is the case. Second, due to their 'black-box' structure, they provide poor explanatory power, and thus are unable to extract understanding of the processes that cause changes in SWE susceptibility. With these considerations in mind, the use of deep learning models may not simply lie in predicting erosion, but integrating these techniques into process-based models to help identify and optimize model parameters and mitigate uncertainty in model estimates (e.g., Vojinovic et al., 2013), help recognize patterns within satellite data to unveil critical details about behavior, and possibly reveal new environmental relationships. Future studies should seek to explore this potential.

This study has considered geo-environmental factors that affect erosion. Where data is available, future studies should consider human-related factors in deep learning models, such as livestock density (e.g., Evans, 1997), arable yields (e.g., Gliessman, 2004) the location of road-networks (e.g., Deng et al., 2011; Keshkamat et al., 2013) and distance from water resources (e.g., Mirzabaev et al., 2016). For example, Yousefi et al. (2016) reported that the rate of soil degradation around water resources was higher than in other parts more distant from water supplies. Deng et al. (2011) demonstrated that when areas are composed of relatively high quality grassland, roads lead to soil erosion and degradation, whereas when grassland resources are sparse, access to a road results in soil restoration. In addition, due to a lack of rainfall intensity data, the current study had to take a simplified approach to account for rainfall as a factor, using mean annual rainfall data to generate the RE factor. This approach does not account for the role of individual storm events, and their magnitude and duration, on soil erosion. Nor does it consider how preceding events affect susceptibility in proceeding events. Thus a focus for future studies should be on applying deep-learning models to more data-rich catchments, using observations of rainfall intensity to generate the RE factor with higher accuracy (Capolongo et al. 2008; Panagos et al. 2015; Petroselli et al., 2021). Since deep learning models are data-driven, one can postulate that with

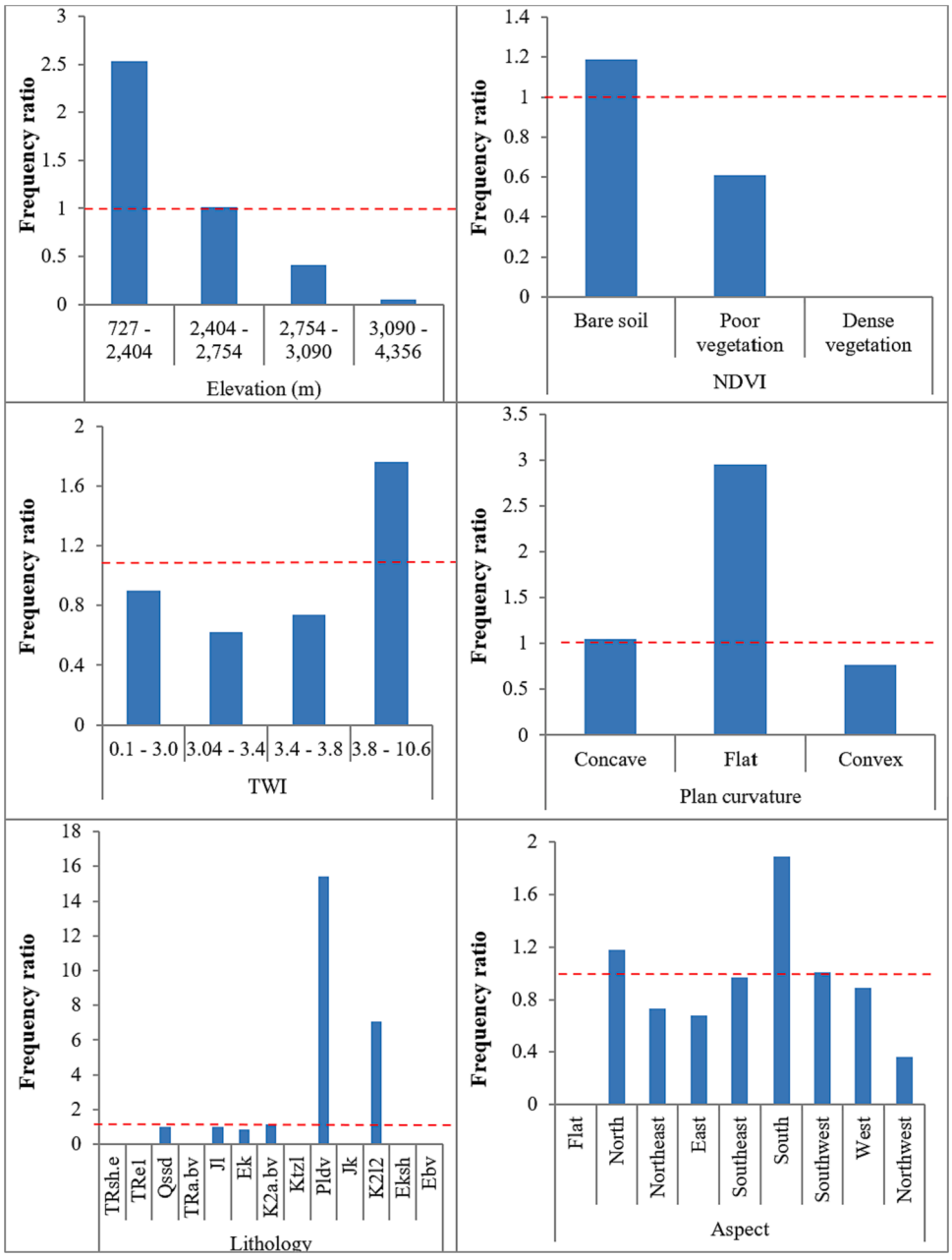


**Fig. 8.** Effectiveness of each of the classes of the most important factors geo-environmental factors on SWE susceptibility, based on FR method.

hourly rainfall data incorporated with the kind of satellite data used in this current study, it should be possible to build more powerful, more accurate models of soil erosion susceptibility over large spatial extents, providing an important tool for targeting soil erosion control to those areas most susceptible, and for identifying areas most suitable for sustainable agricultural development.

## 6. Conclusion

Soil erosion by water is a major cause of global land degradation and soil loss. Accurate predictions of erosion susceptibility are critical for protecting soils and targeting efforts to mitigate the impacts of erosion on ecosystem services, water quality, flooding and infrastructure. Using satellite data, rainfall, soil and other several readily available data, this paper has quantified for the first time, the potential of deep learning models to provide accurate predictions of soil water erosion (SWE) susceptibility. Three state-of-the-art deep learning algorithms – Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Long-Short Term Memory (LSTM) – were applied to assess susceptibility in an Iranian catchment that has historically experienced severe erosion. The main findings were as follows:

1. Elevation was the most effective geo-environmental variable on SWE susceptibility, followed by rainfall erosivity, normalized difference vegetation index, topographic wetness index, plan curvature, ground slope, geology, aspect, stream power index, distance from river, land use, soil texture, slope length and steepness factor, and hydrologic soil groups
2. Model evaluation revealed that all three developed models had good prediction performance, with RNN being marginally the most superior.
3. Maps of SWE susceptibility revealed that almost 40 % of the catchment was considered to be highly or very highly susceptible to SWE

and 20 % moderately susceptible, indicating the critical need for soil erosion control in this catchment to reduce susceptibility.
4. SWE mainly occurs in areas with elevations lower than 2700 m, bare soils, high TWI, flat curvature, a southerly aspect, and rhyolitic to rhyodacitic volcanic rocks

The strength of these algorithms lies in their ease to implement, use of readily available satellite and rainfall data, and being inexpensive to build and run in comparison to physically-based models, whilst providing little compromise on model performance. Together, these findings reveal that deep learning models have great potential for use in SWE susceptibility assessment, especially in situations when understanding of the physical processes at play may not be well understood or field monitoring data is lacking. Thus, understanding more about this potential for different catchments and input variables represents a vital research avenue for hydrologists.

## Funding

This publication has been supported by the RUDN University Scientific Projects Grants System, project No 202235-2-000. In addition, James Cooper was partially supported by two UK Natural Environment Research Council grants (NE/S01697X/1 and NE/V008404/1).

## CRediT authorship contribution statement

**Khabat Khosravi:** Conceptualization, Methods, Software, Write the manuscript, Review and editing. **Fatemeh Rezaie:** Methods, Software. **James R. Cooper:** Write the manuscript, Review and editing. **Zahra Kalantari:** Write the manuscript. **Soroush Abolfathi:** Write the manuscript. **Javad Hatamiafkoueieh:** Write the manuscript.

## 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

The authors wish to thank Dr. Bahram Choubin for sharing the SWE dataset.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2023.129229.

## References

Abolhasani, A., Zehtabian, G., Khosravi, H., Rahmati, O., Alamdarloo, E.H., D'Odorico, P., 2022. A new conceptual framework for spatial predictive modelling of land degradation in a semiarid area. L. Degrad. Dev. 33 (17), 3358–3374.

Abrahart, R.J., Anctil, F., Coulibaly, P., Dawson, C.W., Mount, N.J., See, L.M., Shamseldin, A.Y., Solomatine, D.P., Toth, E., Wilby, R.L., 2012. Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. Prog. Phys. Geogr. Earth Environ. 36, 480–513. https://doi.org/10.1177/0309133312444943.

Abuzaid, A.S., AbdelRahman, M.A.E., Fadl, M.E., Scopa, A., 2021. Land Degradation Vulnerability Mapping in a Newly-Reclaimed Desert Oasis in a Hyper-Arid Agro-Ecosystem Using AHP and Geospatial Techniques. Agronomy 11, 1426. https://doi.org/10.3390/agronomy11071426.

Afshar, F.A., Ayoubi, S., Jalalian, A., 2010. Soil redistribution rate and its relationship with soil organic carbon and total nitrogen using 137Cs technique in a cultivated complex hillslope in western Iran. J. Environ. Radioact. 101, 606–614. https://doi.org/10.1016/j.jenvrad.2010.03.008.

Ahmad, M.W., Reynolds, J., Rezgui, Y., 2018. Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. J. Clean. Prod. 203, 810–821. https://doi.org/10.1016/j.jclepro.2018.08.207.

Ajitha, A., Goel, M., Assudani, M., Radhika, S., Goel, S., 2022. Design and development of Residential Sector Load Prediction model during COVID-19 Pandemic using LSTM based RNN. Electr. Power Syst. Res. 212, 108635 https://doi.org/10.1016/j.epsr.2022.108635.

Akbari, A.-M., 2017. Soil erosion in Iran 2.5 times the world average [WWW Document]. Tehran Times. URL https://www.tehrantimes.com/news/418381/Soil-erosion-in-Iran-2-5-times-the-world-average#:~:text=TEHRAN — Iran has a mean, deputy agriculture minister has said.

Akhavan, S., Abedi-Koupai, J., Mousavi, S.-F., Afyuni, M., Eslamian, S.-S., Abbaspour, K. C., 2010. Application of SWAT model to investigate nitrate leaching in Hamadan-Bahar Watershed. Iran. Agric. Ecosyst. Environ. 139, 675–688. https://doi.org/10.1016/j.agee.2010.10.015.

Amiri, F., 2010. Estimate of Erosion and Sedimentation in Semi-arid Basin using Empirical Models of Erosion Potential within a Geographic Information System. Air. Soil Water Res. 3, ASWR.S3427. https://doi.org/10.4137/ASWR.S3427.

Anastasakis, L., Mort, N., 2001. The Development of Self-Organization Techniques in Modelling: A Review of the Group Method of Data Handling (GMDH). United Kingdom.

Angileri, S.E., Conoscenti, C., Hochschild, V., Märker, M., Rotigliano, E., Agnesi, V., 2016. Water erosion susceptibility mapping by applying Stochastic Gradient Treeboost to the Imera Meridionale River Basin (Sicily, Italy). Geomorphology 262, 61–76. https://doi.org/10.1016/j.geomorph.2016.03.018.

Aslam, B., Maqsoom, A., Salah Alaloul, W., Ali Musarat, M., Jabbar, T., Zafar, A., 2021. Soil erosion susceptibility mapping using a GIS-based multi-criteria decision approach: Case of district Chitral, Pakistan. Ain Shams Eng. J. 12, 1637–1649. https://doi.org/10.1016/j.asej.2020.09.015.

Auerswald, K., Kainz, M., Fiener, P., 2006. Soil erosion potential of organic versus conventional farming evaluated by USLE modelling of cropping statistics for agricultural districts in Bavaria. Soil Use Manag. 19, 305–311. https://doi.org/10.1111/j.1475-2743.2003.tb00320.x.

Boudjemline, F., Semar, A., 2018. Assessment and mapping of desertification sensitivity with MEDALUS model and GIS – Case study: basin of Hodna, Algeria. J. Water L. Dev. 36, 17–26. https://doi.org/10.2478/jwld-2018-0002.

Capolongo, D., Diodato, N., Mannaerts, C.M., Piccarreta, M., Strobl, R.O., 2008. Analyzing temporal changes in climate erosivity using a simplified rainfall erosivity model in Basilicata (southern Italy). J. Hydrol. 356 (1–2), 119–130.

Cerdan, O., Govers, G., Le Bissonnais, Y., Van Oost, K., Poesen, J., Saby, N., Gobin, A., Vacca, A., Quinton, J., Auerswald, K., Klik, A., Kwaad, F.J.P.M., Raclot, D., Ionita, I., Rejman, J., Rousseva, S., Muxart, T., Roxo, M.J., Dostal, T., 2010. Rates and spatial variations of soil erosion in Europe: A study based on erosion plot data. Geomorphology 122, 167–177. https://doi.org/10.1016/j.geomorph.2010.06.011.

Chen, L., Young, M., 2006. Green-Ampt infiltration model for sloping surfaces. Water Resour. Res. 42 (7) https://doi.org/10.1029/2005WR004468.

Chung, C.-J.-F., Fabbri, A.G., 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. Nat. Hazards 30, 451–472. https://doi.org/10.1023/B:NHAZ.0000007172.62651.2b.

Çimen, M., 2008. Estimation of daily suspended sediments using support vector machines. Hydrol. Sci. J. 53, 656–666. https://doi.org/10.1623/hysj.53.3.656.

Conoscenti, C., Di Maggio, C., Rotigliano, E., 2008. Soil erosion susceptibility assessment and validation using a geostatistical multivariate approach: a test in Southern Sicily. Nat. Hazards 46, 287–305. https://doi.org/10.1007/s11069-007-9188-0.

Cooper, J.R., Wainwright, J., Parsons, A.J., Onda, Y., Fukuwara, T., Obana, E., Kitchener, B., Long, E.J., Hargrave, G.H., 2012. A new approach for simulating the redistribution of soil particles by water erosion: A marker-in-cell model. J. Geophys. Res. 117 (F4), n/a–n/a.

Darvishan, A., Sadeghi, S., Gholami, L., 2010. Efficacy of Time-Area Method in simulating temporal variation of sediment yield in Chehelgazi watershed, Iran. Ann. Warsaw Univ. Life Sci. - SGGW. L. Reclam. 42, 51–60. https://doi.org/10.2478/v10060-008-0064-8.

Deng, X., Huang, J., Huang, Q., Rozelle, S., Gibson, J., 2011. Do roads lead to grassland degradation or restoration? A case study in Inner Mongolia, China. Environ. Dev. Econ. 16, 751–773. https://doi.org/10.1017/S1355770X11000180.

Dibike, Y.B., Velickov, S., Solomatine, D., Abbott, M.B., 2001. Model Induction with Support Vector Machines: Introduction and Applications. J. Comput. Civ. Eng. 15, 208–216. https://doi.org/10.1061/(ASCE)0887-3801(2001)15:3(208).

Dodangeh, E., Panahi, M., Rezaie, F., Lee, S., Tien Bui, D., Lee, C.-W., Pradhan, B., 2020. Novel hybrid intelligence models for flood-susceptibility prediction: Meta optimization of the GMDH and SVR models with the genetic algorithm and harmony search. J. Hydrol. 590, 125423 https://doi.org/10.1016/j.jhydrol.2020.125423.

Donovan, M., Monaghan, R., 2021. Impacts of grazing on ground cover, soil physical properties and soil loss via surface erosion: A novel geospatial modelling approach. J. Environ. Manage. 287, 112206 https://doi.org/10.1016/j.jenvman.2021.112206.

Du, R., Liu, W., Fu, X., Meng, L., Liu, Z., 2022. Random noise attenuation via convolutional neural network in seismic datasets. Alexandria Eng. J. 61, 9901–9909. https://doi.org/10.1016/j.aej.2022.03.008.

Ebtehaj, I., Bonakdari, H., 2013. Evaluation of Sediment Transport in Sewer using Artificial Neural Network. Eng. Appl. Comput. Fluid Mech. 7, 382–392. https://doi.org/10.1080/19942060.2013.11015479.

Ebtehaj, I., Bonakdari, H., 2014. Performance Evaluation of Adaptive Neural Fuzzy Inference System for Sediment Transport in Sewers. Water Resour. Manag. 28, 4765–4779. https://doi.org/10.1007/s11269-014-0774-0.
Emadodin, I., Bork, H.R., 2012. Degradation of soils as a result of long-term human-induced transformation of the environment in Iran: an overview. J. Land Use Sci. 7, 203–219. https://doi.org/10.1080/1747423X.2011.560292.
Emadodin, I., Narita, D., Bork, H.R., 2012. Soil degradation and agricultural sustainability: an overview from Iran. Environ. Dev. Sustain. 14, 611–625. https://doi.org/10.1007/s10668-012-9351-y.
Evans, R., 1997. Soil erosion in the UK initiated by grazing animals. Appl. Geogr. 17, 127–141. https://doi.org/10.1016/S0143-6228(97)00002-7.
Fang, H., Guo, M., 2015. Aspect-induced differences in soil erosion intensity in a gullied hilly region on the Chinese Loess Plateau. Environ. Earth Sci. 74, 5677–5685.
Fang, Z., Wang, Y., Peng, L., Hong, H., 2021. Predicting flood susceptibility using LSTM neural networks. J. Hydrol. 594, 125734 https://doi.org/10.1016/j.jhydrol.2020.125734.
Ganguli, P., Reddy, M.J., 2014. Ensemble prediction of regional droughts using climate inputs and the SVM-copula approach. Hydrol. Process. 28, 4989–5009. https://doi.org/10.1002/hyp.9966.
Ghorbanzadeh, O., Meena, S.R., Blaschke, T., Aryal, J., 2019. UAV-based slope failure detection using deep-learning convolutional neural networks. Remote Sens. 11, 2046. https://doi.org/10.3390/rs11172046.
Gliessman, S.R., 2004. Integrating Agroecological Processes into Cropping Systems Research. J. Crop Improv. 11, 61–80. https://doi.org/10.1300/J411v11n01_04.
Goyal, M.K., Bharti, B., Quilty, J., Adamowski, J., Pandey, A., 2014. Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. Expert Syst. Appl. 41, 5267–5276. https://doi.org/10.1016/j.eswa.2014.02.047.
Keshkamat, S.S., Tsendbazar, N.E., Zuidgeest, M.H.P., Shiirev-Adiya, S., van der Veen, A., van Maarseveen, M.F.A.M., 2013. Understanding transportation-caused rangeland damage in Mongolia. J. Environ. Manage. 114, 433–444. https://doi.org/10.1016/j.jenvman.2012.10.043.
Khalili Moghadam, B., Jabarifar, M., Bagheri, M., Shahbazi, E., 2015. Effects of land use change on soil splash erosion in the semi-arid region of Iran. Geoderma 241–242, 210–220. https://doi.org/10.1016/j.geoderma.2014.11.025.
Khosravi, K., Panahi, M., Tien Bui, D., 2018. Spatial prediction of groundwater spring potential mapping based on an adaptive neuro-fuzzy inference system and metaheuristic optimization. Hydrol. Earth Syst. Sci. 22, 4771–4792. https://doi.org/10.5194/hess-22-4771-2018.
Khosravi, K., Panahi, M., Golkarian, A., Keesstra, S.D., Saco, P.M., Bui, D.T., Lee, S., 2020. Convolutional neural network approach for spatial prediction of flood hazard at national scale of Iran. J. Hydrol. 591, 125552 https://doi.org/10.1016/j.jhydrol.2020.125552.
Kinnell, P.I.A., 2010. Event soil loss, runoff and the Universal Soil Loss Equation family of models: A review. J. Hydrol. 385, 384–397. https://doi.org/10.1016/j.jhydrol.2010.01.024.
Kinnell, P.I.A., 2017. A comparison of the abilities of the USLE-M, RUSLE2 and WEPP to model event erosion from bare fallow areas. Sci. Total Environ. 596–597, 32–42. https://doi.org/10.1016/j.scitotenv.2017.04.046.
Kisi, O., Dailr, A.H., Cimen, M., Shiri, J., 2012. Suspended sediment modeling using genetic programming and soft computing techniques. J. Hydrol. 450–451, 48–58. https://doi.org/10.1016/j.jhydrol.2012.05.031.
Kisi, O., Genc, O., Dinc, S., Zounemat-Kermani, M., 2016. Daily pan evaporation modeling using chi-squared automatic interaction detector, neural networks, classification and regression tree. Comput. Electron. Agric. 122, 112–117. https://doi.org/10.1016/j.compag.2016.01.026.
Laylin, D., 2018. Environmental and wildlife degradation in Iran [WWW Document]. Atl. Counc. URL https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/environmental-and-wildlife-degradation-in-iran-2/.
Li, H., Xu, Q., He, Y., Fan, X., Yang, H., Li, S., 2021. Temporal detection of sharp landslide deformation with ensemble-based LSTM-RNNs and Hurst exponent. Geomatics. Nat. Hazards Risk 12, 3089–3113. https://doi.org/10.1080/19475705.2021.1994474.
Liu, H., Lang, B., Liu, M., Yan, H., 2019. CNN and RNN based payload classification methods for attack detection. Knowl.-Based Syst. 163, 332–341. https://doi.org/10.1016/j.knosys.2018.08.036.
Marques, M.A., Mora, E., 1992. The influence of aspect on runoff and soil loss in a Mediterranean burnt forest (Spain). CATENA 19 (3–4), 333–344.
Melesse, A.M., Ahmad, S., McClain, M.E., Wang, X., Lim, Y.H., 2011. Suspended sediment load prediction of river systems: An artificial neural network approach. Agric. Water Manag. 98, 855–866. https://doi.org/10.1016/j.agwat.2010.12.012.
Mirzabaev, A., Ahmed, M., Werner, J., Pender, J., Louhaichi, M., 2016. Rangelands of Central Asia: challenges and opportunities. J. Arid Land 8, 93–108. https://doi.org/10.1007/s40333-015-0057-5.
Mohammadi, S., Balouei, F., Haji, K., Khaledi Darvishan, A., Karydas, C.G., 2021. Country-scale spatio-temporal monitoring of soil erosion in Iran using the G2 model. Int. J. Digit. Earth 14, 1019–1039. https://doi.org/10.1080/17538947.2021.1919230.
Mosavi, A., Sajedi-Hosseini, F., Choubin, B., Taromideh, F., Rahi, G., Dineva, A., 2020. Susceptibility Mapping of Soil Water Erosion Using Machine Learning Models. Water 12, 1995. https://doi.org/10.3390/w12071995.
Mutlu, B., Nefeslioglu, H.A., Sezer, E.A., Akcayol, M.A., Gokceoglu, C., 2019. An experimental research on the use of recurrent neural networks in landslide susceptibility mapping ISPRS. Int. J. Geo-Inf. 8 (12), 578.
Narantsetseg, A., Kang, S., Ko, D., 2018. Livestock grazing and trampling effects on plant functional composition at three wells in the desert steppe of Mongolia. J. Ecol. Environ. 42, 13. https://doi.org/10.1186/s41610-018-0075-2.
Nearing, M.A., Foster, G.R., Lane, L.J., Finkner, S.C., 1989. A process-based soil erosion model for USDA-water erosion prediction project technology. Trans. ASAE 32, 1587–1593.
Novakovic, J., 2009. Using Information Gain Attribute Evaluation to Classify Sonar Targets. In: 17th Telecommunications Forum TELFOR 2009. Serbia, Belgrade, pp. 1351–1354.
Orgill, S.E., Condon, J.R., Conyers, M.K., Morris, S.G., Alcock, D.J., Murphy, B.W., Greene, R.S.B., 2018. Removing Grazing Pressure from a Native Pasture Decreases Soil Organic Carbon in Southern New South Wales, Australia. L. Degrad. Dev. 29, 274–283. https://doi.org/10.1002/ldr.2560.
Pally, R.J., Samadi, S., 2022. Application of image processing and convolutional neural networks for flood image classification and semantic segmentation. Environ. Model. Softw. 148, 105285 https://doi.org/10.1016/j.envsoft.2021.105285.
Panagos, P., Ballabio, C., Borrelli, P., Meusburger, K., Klik, A., Rousseva, S., Tadić, M.P., Michaelides, S., Hrabalíková, M., Olsen, P., Aalto, J., Lakatos, M., Rymszewicz, A., Dumitrescu, A., Beguería, S., Alewell, C., 2015. Rainfall erosivity in Europe. Sci. Total Environ. 511, 801–814.
Panahi, M., Khosravi, K., Golkarian, A., Roostaei, M., Barzegar, R., Omidvar, E., Rezaie, F., Saco, P.M., Sharifi, A., Jun, C., Bateni, S.M., Lee, C.-W., Lee, S., 2022. A country-wide assessment of Iran's land subsidence susceptibility using satellite-based InSAR and machine learning. Geocarto Int. 1–23 https://doi.org/10.1080/10106049.2022.2086631.
Petroselli, A., Apollonio, C., Luca, D.L., Salvaneschi, P., Pecci, M., Marras, T., Schirone, B., 2021. Comparative Evaluation of the Rainfall Erosivity in the Rieti Province, Central Italy, Using Empirical Formulas and a Stochastic Rainfall Generator. Hydrology 8 (4), 171. https://doi.org/10.3390/hydrology8040171.
Phinzi, K., Ngetar, N.S., Ebhuoma, O., 2021. Soil erosion risk assessment in the Umzintlava catchment (T32E), Eastern Cape, South Africa, using RUSLE and random forest algorithm. South Afr. Geogr. J. 103, 139–162. https://doi.org/10.1080/03736245.2020.1716838.
Puigdefabregas, J., Sole, A., Gutierrez, L., del Barrio, G., Boer, M., 1999. Scales and processes of water and sediment redistribution in drylands: results from the Rambla Honda field site in Southeast Spain. Earth-Sci. Rev. 48, 39–70. https://doi.org/10.1016/S0012-8252(99)00046-X.
Rapp, J., 1963. Error assessment of the revised universal soil loss equation using natural runoff plot data. University of Arizona.
Raza, A., Ahrends, H., Habib-Ur-Rahman, M., Gaiser, T., 2021. Modeling Approaches to Assess Soil Erosion by Water at the Field Scale with Special Emphasis on Heterogeneity of Soils and Crops. Land 10, 422. https://doi.org/10.3390/land10040422.
Renard, K.G., Foster, G.R., Weesies, G.A., Mccool, D.K., Yoder, D.C., 1997. Predicting soil erosion by water: a guide to conservation planning with the revised universal soil loss equation (RUSLE).
Robinson, C., 1998. Multi-objective optimisation of polynomial models for time series prediction using genetic algorithms and neural networks. University of Sheffield.
Rosskopf, C.M., Di Iorio, E., Circelli, L., Colombo, C., Aucelli, P.P.C., 2020. Assessing spatial variability and erosion susceptibility of soils in hilly agricultural areas in Southern Italy. Int. Soil Water Conserv. Res. 8, 354–362. https://doi.org/10.1016/j.iswcr.2020.09.005.
Sadeghi, S.H.R., 2017. Soil erosion in Iran: State of the art, tendency and solutions. J. Agric. For. 63 https://doi.org/10.17707/AgricultForest.63.3.04.
Sajedi-Hosseini, F., Choubin, B., Solaimani, K., Cerdà, A., Kavian, A., 2018. Spatial prediction of soil erosion susceptibility using a fuzzy analytical network process: Application of the fuzzy decision making trial and evaluation laboratory approach. L. Degrad. Dev. 29, 3092–3103. https://doi.org/10.1002/ldr.3058.
Solaimani, K., Hadian Amr, M.A., 2008. Application of IRS-1D Data in Water Erosion Features Detection (Case Study: Nour Roud Catchment, Iran). Pakistan J. Biol. Sci. 11, 1893–1900. https://doi.org/10.3923/pjbs.2008.1893.1900.
Sternberg, M., Gutman, M., Perevolotsky, A., Ungar, E.D., Kigel, J., 2000. Vegetation response to grazing management in a Mediterranean herbaceous community: a functional group approach. J. Appl. Ecol. 37, 224–237. https://doi.org/10.1046/j.1365-2664.2000.00491.x.
Tan, Z., Leung, L.R., Li, H., Tesfa, T., 2018. Modeling Sediment Yield in Land Surface and Earth System Models: Model Comparison, Development, and Evaluation. J. Adv. Model. Earth Syst. 10, 2192–2213. https://doi.org/10.1029/2017MS001270.
Tang, Q., Xu, Y., Bennett, S.J., Li, Y., 2015. Assessment of soil erosion using RUSLE and GIS: a case study of the Yangou watershed in the Loess Plateau, China. Environ. Earth Sci. 73, 1715–1724. https://doi.org/10.1007/s12665-014-3523-z.
Thi Ngo, P.T., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Kariminejad, N., Cerda, A., Lee, S., 2021. Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. Geosci. Front. 12, 505–519. https://doi.org/10.1016/j.gsf.2020.06.013.
Tien Bui, D., Pradhan, B., Lofman, O., Revhaug, I., 2012. Landslide Susceptibility Assessment in Vietnam Using Support Vector Machines, Decision Tree, and Naïve Bayes Models. Math. Probl. Eng. 2012, 1–26. https://doi.org/10.1155/2012/974638.
Tien Bui, D., Pradhan, B., Nampak, H., Bui, Q.T., Tran, Q.A., Nguyen, Q.P., 2016. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibilitgy modeling in a high-frequency tropical cyclone area using GIS. J. Hydrol. 540, 317–330. https://doi.org/10.1016/j.jhydrol.2016.06.027.

Trabelsi, M., Meddouri, N., Maddouri, M., 2017. A New Feature Selection Method for Nominal Classifier based on Formal Concept Analysis. Proc. Comput. Sci. 112, 186–194. https://doi.org/10.1016/j.procs.2017.08.227.

Vaezi, A.R., Hasanzadeh, H., Cerdà, A., 2016. Developing an erodibility triangle for soil textures in semi-arid regions, NW Iran. CATENA 142, 221–232. https://doi.org/10.1016/j.catena.2016.03.015.

Venkatappareddy, P., Culli, J., Srivastava, S., Lall, B., 2021. A Legendre polynomial based activation function: An aid for modeling of max pooling. Digit. Signal Process. 115, 103093 https://doi.org/10.1016/j.dsp.2021.103093.

Vojinovic, Z., Abebe, Y., Ranasinghe, R., Vacher, A., Martens, P., Mandl, D., et al., 2013. A machine learning approach for estimation of shallow water depths from optical satellite images and sonar measurements. Journal of Hydroinformatics 15 (4), 1408–1424.

Vu, D.T., Tran, X.-L., Cao, M.-T., Tran, T.C., Hoang, N.-D., 2020. Machine learning based soil erosion susceptibility prediction using social spider algorithm optimized multivariate adaptive regression spline. Measurement 164, 108066. https://doi.org/10.1016/j.measurement.2020.108066.

Wainwright, J., Parsons, A.J., Müller, E.N., Brazier, R.E., Powell, D.M., Fenti, B., 2008. A transport-distance approach to scaling erosion rates: 1. Background and model development. Earth Surf. Process. Landforms 33, 813–826. https://doi.org/10.1002/esp.1624.

Wijitkosum, S., 2021. Factor influencing land degradation sensitivity and desertification in a drought prone watershed in Thailand. Int. Soil Water Conserv. Res. 9, 217–228. https://doi.org/10.1016/j.iswcr.2020.10.005.

Williams, J.R., 1975. Sediment-yield prediction with Universal Equation using runoff energy factor.

Wischmeier, W.H., Smith, D.D., 1965. Predicting rainfall-erosion losses from cropland east of the Rocky Mountains:guide for selection of practices for soil and water conservation. Agricultural Research Service, U. S. Dept of Agriculture in cooperation with Purdue Agricultural Experiment Station.

Wischmeier, W.H., Smith, D.D., 1978. Predicting rainfall erosion losses: a guide to conservation planning. Science and Education Administration, U.S. Dept. of Agriculture.

Woolhiser, D.A., Smith, R.E., Goodrich, D.C., 1990. KINEROS:a kinematic runoff and erosion model: documentation and user manual.

World Bank, 2005. Islamic Republic of Iran, Cost Assessment of Environmental Degradation.

Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K., 2018. Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629. https://doi.org/10.1007/s13244-018-0639-9.

Yesilnacar, E.K., 2005. The Application of Computational Intelligence to Landslide Susceptibility Mapping in Turkey. University of Melbourne.

Yousefi, S., Moradi, H., Boll, J., Schönbrodt-Stitt, S., 2016. Effects of road construction on soil degradation and nutrient transport in Caspian Hyrcanian mixed forests. Geoderma 284, 103–112. https://doi.org/10.1016/j.geoderma.2016.09.002.

Yousefi, S., Pourghasemi, H.R., Avand, M., Janizadeh, S., Tavangar, S., Santosh, M., 2021. Assessment of land degradation using machine-learning techniques: A case of declining rangelands. L. Degrad. Dev. 32, 1452–1466. https://doi.org/10.1002/ldr.3794.

Yu, B., Rosewell, C.J., 1996. Technical Notes: A Robust Estimator of the R-factor for the Universal Soil Loss Equation. Trans. ASAE 39, 559–561. https://doi.org/10.13031/2013.27535.

Zhang, X.C., Nearing, M.A., Risse, L.M., McGregor, K.C., 1996. Evaluation of WEPP Runoff And Soil Loss Predictions Using Natural Runoff Plot Data. Trans. ASAE 39, 855–863. https://doi.org/10.13031/2013.27570.