EEET ECOLOGICAL ENGINEERING & ENVIRONMENTAL TECHNOLOGY

Ecological Engineering & Environmental Technology, 2025, 26(1), 8–28
https://doi.org/10.12912/27197050/195227
ISSN 2719–7050, License CC-BY 4.0

Received: 2024.10.18
Accepted: 2024.11.21
Published: 2024.12.01

Enhancing flash flood risk prediction – A case study from the Assaka watershed, Guelmim Region, Southwestern Morocco

Soukaina Talha[1*], Ahmed Akhssas[1], Abdellatif Aarab[2], Ayoub Aabi[2], Badr Berkat[1]

[1] Laboratory of Applied Geophysics, Geotechnics, Engineering Geology, and Environment, Mohammadia Engineering School, Mohammed V University in Rabat, 765, Av. Ibn Sina, Agdal, Rabat, Morocco
[2] Laboratory for Water Analysis and Modeling of Natural Resources, Mohammadia Engineering School, Mohammed V University in Rabat, BP:765 Ibn Sina Avenue Agdal, Rabat, Morocco
* Corresponding author's e-mail: soukaina_talha2@um5.ac.ma

ABSTRACT

Since the onset of the Industrial Revolution, significant climatic shifts have led to various environmental imbalances globally, notably increasing the frequency of flash floods, especially in vulnerable regions like the Assaka watershed in southwestern Morocco. This study aims to enhance flash flood risk prediction by integrating machine learning (ML) algorithms with geographic information system (GIS) technology. The random forest (RF) algorithm was employed to analyze over eight million data points, using fourteen predictors categorized into topographic (e.g., altitude, slope, topographic wetness index (TWI)), climatic (e.g., land surface temperature (LST), soil moisture index (SMI)), and geological factors (e.g., drainage density, soil type, lithology). These variables were derived from remotely sensed data and geospatial analyses. The RF model classified the Assaka watershed into five flood susceptibility levels: lowest, low, medium, high, and highest. The results indicated that the most vulnerable areas are near the watershed outlet and the main tributaries, Essayed and Oum Laachar Wadis. These regions are characterized by high land surface temperatures, low drainage density, poor soil moisture, and specific geological conditions, all of which contribute to heightened flood risk. The model's performance was evaluated using multiple metrics, achieving precision (0.968), recall (0.967), accuracy (0.967), F1 score (0.965), Kappa statistic (0.839), and an AUC of 1.0, highlighting its robustness and predictive capabilities. The originality of this study lies in its comprehensive integration of ML with GIS to develop a highly reliable flood susceptibility map for the Assaka watershed. This framework addresses existing gaps in flood risk assessment, offering a significant advancement over traditional methods through its use of advanced data-driven modeling techniques. The findings provide essential insights for prioritizing conservation and flood management strategies, contributing to better preparedness against flash floods in the Guelmim region and potentially other similar environments globally.

Keywords: geographic information system, artificial intelligence, machine learning, random forests, susceptibility mapping, Morocco, Assaka watershed, flash flood, forecasting.

INTRODUCTION

Flash floods are among the most frequent and destructive natural disasters globally, inflicting widespread damage on both urban areas and agricultural lands while also posing a significant threat to human life (He et al., 2024a; Isma et al., 2024; Saha et al., 2024). The urgency to address this challenge has intensified, particularly in the face of climate change, which exacerbates the frequency and severity of such events (Atanga and Tankpa, 2021; Ionno et al., 2024; Kundzewicz et al., 2014). Morocco, like many countries, has been increasingly vulnerable to catastrophic floods, with several regions experiencing repeated and devastating incidents over the past decades. Notably, cities like Agadir, Marrakech, and Beni Mellal have been severely impacted by flash floods, resulting in loss of life, property damage, and disruption of economic activities (Aangri et al., 2024; Cotti et al., 2022; Said and Ahmed, 2023). These regions face a complex interplay of environmental and urban factors that make them prone to flooding.

Previous approaches to managing flood risks in Morocco have largely relied on traditional methods, such as spatial data analysis and remote sensing (Sumi, Kantoush, and Saber, n.d.). Remote sensing has enabled real-time monitoring and rapid assessment of flood damage but often lacks the predictive capability required to prevent such disasters. Despite advances, conventional methods have sometimes fallen short in accurately predicting flash floods due to the complex and dynamic nature of the phenomenon (Patil et al., 2024). Thus, there is a growing need for more sophisticated and predictive models, particularly in regions where flash floods are recurrent.

In recent years, ML has emerged as a powerful tool in environmental science, offering more accurate and data-driven approaches for flood risk assessment (Mosavi et al., 2018). The use of ML models, especially when combined with GIS and spatial data, has provided more reliable predictions of flood-prone areas. However, despite the global shift toward advanced technologies, many regions in Morocco still rely on outdated or incomplete methods, making them vulnerable to increasingly severe flood events.

Among the most affected regions in Morocco is the Guelmim area in the southwest, which has endured multiple severe floods over the past five decades, with notable events in 1968, 1985, 1989, 2002, 2010, and 2014 (Bannari, e al., 2019; Khaddari et al., 2023a; Theilen-Willige et al., 2015). These recurrent disasters have led the Moroccan government to designate Guelmim as a disaster-prone area (Bannari et al., 2020; Talha et al., 2019). Several studies have been conducted to assess flood risks in this region, utilizing techniques like spatial prediction of flood hazards, land surface moisture evaluation, and the fuzzy analytical hierarchy process (FAHP) combined with GIS (Talha et al., 2019). While these methods have contributed to a better understanding of flood risks, they often lack the precision needed to mitigate future events effectively (Al-Aizari et al., 2024; Albertini et al., 2024; Tan et al., 2024a).

This study aims to address the limitations of conventional methods by enhancing flash flood risk prediction in the Assaka watershed within the Guelmim region. By employing the RF machine learning algorithm, renowned for its accuracy and robustness, this research seeks to develop a highly reliable flood susceptibility map for the region. The purpose of this study is to fill a critical gap in existing flood risk management by introducing a more data-driven and predictive approach, capable of accurately identifying and classifying flood-prone areas. The main objective is to assess whether integrating ML and GIS technologies can significantly improve flood prediction accuracy in comparison to traditional methods, thereby providing a valuable tool for disaster preparedness and risk management. The study hypothesizes that the random forest model, with its ability to process large and complex datasets, will provide enhanced predictive capabilities that contribute to a better understanding of flood susceptibility in the Guelmim region.

MATERIALS AND METHODS

Study area

The subject of this study (Fig. 1 and Fig. 2), the Assaka watershed, is located in southern Morocco within the Guelmim-Oued Noun region, under the supervision of the Hydraulic Basin Agency of Souss Massa ABHSM of Morocco. This hydraulic sub-watershed of Guelmim covers an area of 6.862 square kilometers, with a perimeter of 597 kilometers (Khaddari et al., 2022; Mahmouhi et al., 2016; Mathieu et al., 2004). It is characterized by an arid to semi-arid desert climate, with an average annual temperature of 19 °C, and annual rainfall typically reaching up to 145 mm (Bannari et al., 2019; Said, Ahmed, and Kharrim, 2023; Said, Ahmed, Lahcen, Rhita, n.d.). According to the 2014 census, the region supports a population of over 18.000 people (HCP of Morocco, 2014). The watershed has experienced several catastrophic flood and flash flood events in recent years, with notable occurrences in 1968, 1985, 1989, 2002, 2010, and 2014 (Bannari et al., 2020; Bannari et al., 2019; Talha et al., 2019). These events have caused widespread damage, including the destruction of

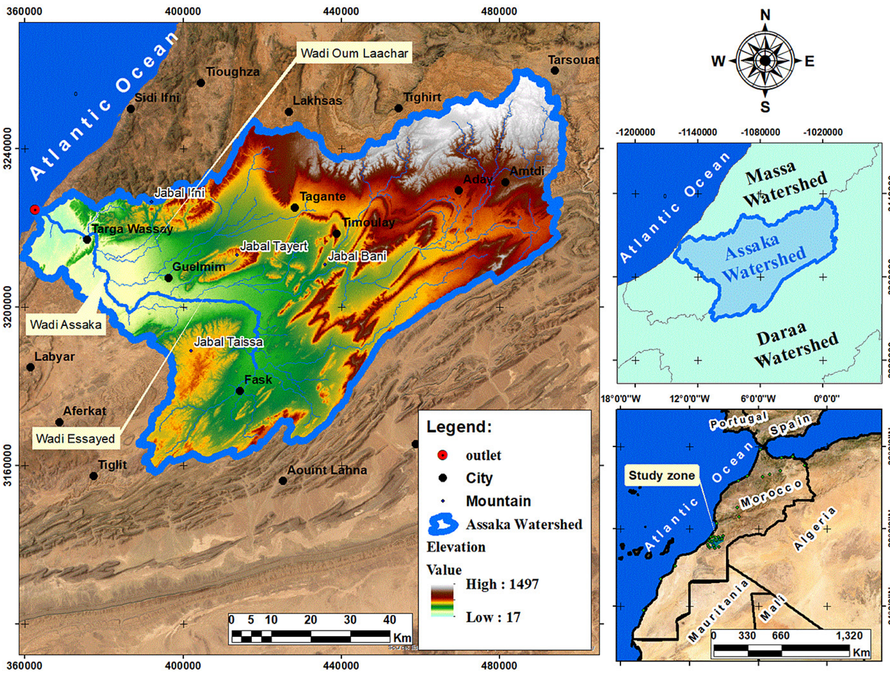


Figure 1. Spatial representation of the study region

infrastructure, loss of lives, and displacement of residents.

The 2014 floods were particularly severe, resulting in the Moroccan government declaring Guelmim a disaster zone (Talha et al., 2019). The Assaka watershed features rugged terrain in its northeastern part, with elevations ranging from sea level along the western Atlantic coast to over 1.100 meters in the mountainous areas (Theilen-Willige et al., 2015). The watershed is fed by the confluence of two rivers, Oum Laachar and Essayed, which flow into the Atlantic Ocean. The landscape is dominated by depressions and broad valleys with flat bottoms primarily located at elevations between 300 and 600 meters (Khaddari et al., 2023a; Mahmouhi et al., 2016;

Figure 2. Flood impact on oases, farmlands, and a classroom in the study area (November 24, 2014) (Bannari et al., 2016)

Mathieu et al., 2004). Recent flash flood events in the region highlight the vulnerability of this area. For instance, during the floods of 2014, the region experienced significant flooding, particularly around the Essayed and Oum Laachar tributaries. These floods damaged roads, homes, and agricultural lands, leaving long-lasting impacts on local livelihoods. The watershed's steep terrain and limited drainage capacity exacerbate the impact of heavy rainfall, making flash floods more frequent and severe. Photos of the aftermath of these floods show the devastation caused, with submerged roads and washed-away bridges becoming a common sight. The watershed's slope map categorizes the terrain into five classes, with values ranging from 0 to over 64 degrees. The slope generally decreases from the northeast toward the southwest, with large areas featuring gentle to moderate slopes (less than 11 degrees) (Talha et al., 2019). Morphometric analysis highlights that wide valleys and depressions at elevations between 350 and 400 meters, along with tableland regions characterized by gentle slopes, are significant features that influence flood dynamics (Talha et al., 2019).

Method

The methodology followed in this study is summarized in the flowchart presented in (Fig. 3) It involves the preparation of a flood information map, identification of the factors contributing to flash flood conditions, creation of a flood susceptibility model using RF algorithms, evaluation of model accuracy, and the generation of a comprehensive flash flood susceptibility map (Al-Aizari et al., 2024; Ganjirad and Delavar, 2023; Tan et al., 2024a).

Data used

The primary datasets employed in this research include data from the Landsat-8 operational land imager (OLI) sensor, the digital elevation model (DEM), and the flood information map.

Digital elevation model (DEM)

Altitude, represented as altimetry, is a fundamental descriptor of terrestrial topography. It is a crucial factor influencing the occurrence of flash

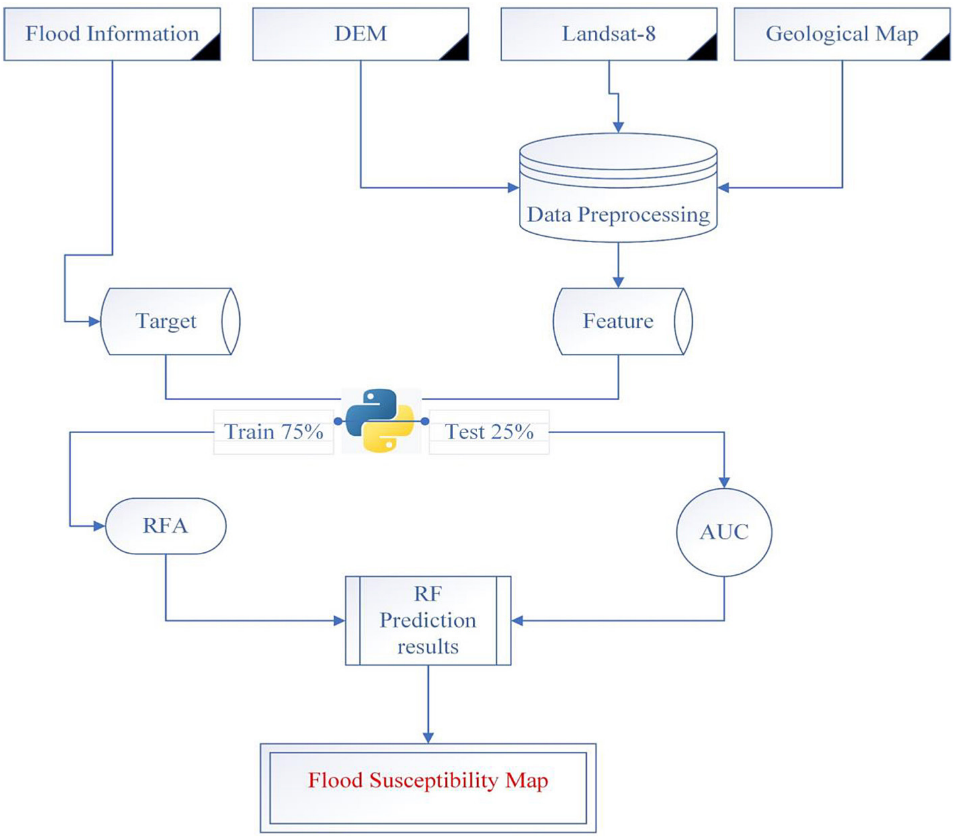


Figure 3. Process for flash flood susceptibility mapping in the ASSAKA watershed using GIS and machine learning

floods (Avand 2022; Muthusamy et al., 2021). Generally, the frequency of flash floods tends to increase as elevation decreases, rendering lower areas more prone to flooding conditions (Hawker et al., 2018; Tarekegn et al., 2010). For this case study, elevation data with a spatial resolution of 30 by 30 meters was collected from the Global Data Explorer. The elevation in the study area varies significantly, ranging from as low as 17 meters to heights exceeding 1000 meters. This variation in elevation plays an essential role in analyzing flood susceptibility and terrain characteristics. (Fig. 5b).

Landsat-8 oli sensor data

Landsat-8 is the eighth satellite in the LANDSAT series and is equipped with two main sensors: the OLI and the thermal infrared sensor (TIRS). The OLI sensor captures images in nine spectral bands, including visible, near-infrared, and shortwave infrared, with a spatial resolution of 30 by 30 meters. This allows for comprehensive coverage of large terrestrial regions while maintaining sufficient resolution to identify and categorize various surface features. Additionally, the TIRS was integrated into the mission to measure Earth's thermal energy. It operates in two bands, ten and eleven, providing resolutions ranging from 30 to 100 meters. (Mao et al., 2020; Sivrikaya et al., 2024; Sun et al., 2021).

Flood information map

To collect flood data for predicting flash flood susceptibility using the random forest machine learning algorithm (Fig. 5o), the methodology is organized into several crucial steps. Initially, flood conditioning factors are prepared using GIS tools and geographic datasets. Next, the analytical hierarchy process (AHP) is employed to assign weights to various criteria. In the third step, these factors are transformed into a normalized range between 0 and 1 through the Fuzzy Membership algorithm with linear functions. The weighted criteria obtained from the AHP are then combined with the normalized factors. Ultimately, this complete dataset is used in the Random Forest algorithm to generate a susceptibility map for flash floods (Talha et al., 2019).

Flood factors

Based on insights from previous studies (Bakhtiari et al., 2023; He et al., 2024b; Tan et al., 2024b), fourteen critical flash flood conditioning predictors have been identified and utilized in this case study. The factors analyzed include altitude, slope, aspect, drainage density, soil type, lithology, and land use and land cover (LU/LC), along with hydrological indices such as flow accumulation (FA), stream power index (SPI), topographic position index (TPI), topographic wetness index (TWI), and curvature. Additionally, climatic variables such as land surface temperature (LST) and soil moisture index (SMI) are considered. This diverse range of factors is derived from established research and plays a critical role in evaluating flash flood susceptibility.

Drainage density

In watershed environments, drainage density (Fig. 5d) significantly influences water circulation when rainfall occurs (Bhattacharjee, n.d.; Ngai et al., 2024; Yao et al., 2017). This factor plays a crucial role in predicting the likelihood of flash floods. Lower drainage systems can lead to various issues, including watershed overflow and persistent flooding in specific areas. Given the substantial impact of drainage density, it is essential to accurately measure it. In this case, drainage density was derived using the line density tool from QGIS software. For the Research Area, the resulting drainage density map, displayed in (Fig. 5d), ranged from 0 to over 2.30 km/km^2 at a 30 by 30-meter resolution.

Land use land cover

Land use and land cover (LULC) (Fig. 5k) at various scales is a valuable asset for resource managers, offering insights drawn from a broad range of satellite imagery to satisfy user requirements and support decision-making (Bano, 2024; Ghouldan et al., 2023). In this study, land use is a critical factor in determining runoff velocity, which motivated its inclusion in the analysis. The vulnerability to floods in LULC contexts differs for each land cover type due to the distinct effects of these covers on water runoff and absorption. Presented below is a detailed analysis of the flood vulnerability of the aforementioned LULC classes (Mawasha and Britz, 2022a). Urban Regions Surfaces like concrete and asphalt cannot absorb water, making these regions more vulnerable to flooding (Allafta and Opp, 2021). As a result, there is more runoff and the possibility of flooding(Abd El-Hamidet al., 2021). Bodies of water because of the high water content in these places, flooding is a real possibility in the areas immediately around them in the event of heavy rains or an influx of water from further upstream (Mawasha and Britz, 2022b). The flora Vegetation helps absorb water and stabilizes the soil, minimizing runoff, therefore areas with lush vegetation tend to be less vulnerable to floods. Agricultural Lands: The vulnerability of these regions to flooding might differ. Reduce your risk of flooding by using well-managed agricultural fields that use effective soil techniques to absorb large volumes of water. On the other hand, floods might occur more often on unmanaged property that has exposed or compacted soil (Ait El Haj et al., 2023; Parvin et al., 2024). No Coverage Bare ground, particularly compacted or poorly permeable soil, makes an area prone to floods. More water will flow off and more people will be in danger of flooding since there is very little vegetation to soak it up. Therefore, according to LULC classifications, places with lush vegetation are often less vulnerable to flooding than those with built-up areas or bare land (Abd El-Hamid et al., 2021; Allafta and Opp, 2021; Mawasha and Britz, 2022a). A land use map for the study area was created using Landsat 8 OLI satellite imagery, which has a spatial resolution of 30 meters.This map categorizes the region into five key classes: Built-up areas, water bodies, vegetation, agricultural lands, and bare soil.

Slope

Flash floods are closely linked to the degree of slope (Fig. 5c) (Hermawan et al., 2024; Iresh et al., 2024; Shayannejad and Ostad, n.d.), which is a critical morphological characteristic. The slope has a direct impact on the velocity of surface runoff, which in turn affects flash flood susceptibility. In this study, a slope degree map was produced from the digital elevation model (DEM) raster using QGIS. The slope values across the study area range from 0 to more than 64 degrees, with a spatial resolution of 30 meters.

Soil moisture index (SMI)

Soil moisture (SM) (Fig. 5a) is an important physiographic feature used in numerous hydrological applications and acts as an early warning sign for potential flash floods. Regions with a low soil moisture index (SMI) tend to retain less water, indicating limited moisture absorption capacity and an increased risk of flash flooding (Saha et al., 2018; Zhang and Zhou, 2016). On the other hand, regions with a high Soil Moisture Index (SMI) demonstrate better moisture absorption abilities, making them less prone to flash floods. In this study, SMI maps were generated using Landsat-8 OLI data with a spatial resolution of 30 meters. The analysis showed that SMI values are notably high in mountainous areas, highlighting their critical role in evaluating flood risk.

Soil type

Soil (Fig. 5m) is a porous medium composed of three phases: water, air, and minerals. It is also home to numerous microorganisms and macroorganisms from both the plant and animal kingdoms. Insufficient water absorption by soil plays a significant role in the development of floods (Basri et al., 2022; Loeb et al., 2007). Therefore, recognizing soil types within a study area is of paramount importance. Creating a soil map involves utilizing data from the FAO. In this research, the resulting map classified the soils into six categories: Calcaric Fluvisols (Jc), Chromic Luvisols (Lc), Lithosols (I), Yermosols (Y), and Haplic Yermosols (Yh). The map was subsequently re-projected from the WGS84 coordinate system to UTM coordinates.

Aspect

The slope aspect (Fig. 5e) is defined as the direction in which the terrain's maximum slope occurs and is considered a crucial parameter for flood susceptibility analysis in many studies (Buttle et al., 2016; Kang et al., 2021; Lawford et al., 1995). The aspect map was generated in QGIS using data from the DEM.

Curvature

Curvature (Fig. 5f) is identified as another influential conditioning factor for flood susceptibility and is derived from the digital elevation model (DEM) using QGIS (Heerdegen and Beran, 1982; Jeong et al., 2003). It is categorized into three types: concave, convex, and flat surfaces. As a critical element affecting runoff flow, curvature can play a significant role in assessing areas prone to flooding (Malik et al., 2020; Tehrany et al., 2013).

Topographic wetness index (TWI)

The TWI (Fig. 5h), introduced by Beven and Kirkby in 1979, quantifies the spatial distribution of moisture across a watershed (De Risi et al., 2018;

Riadi et al., 2018). It essentially measures the potential for water accumulation at any given pixel within the watershed, calculated using the formula:

$$TWI = \ln(\frac{A}{\tan \beta}) \tag{1}$$

where: A represents the specific catchment area (m^2 per meter), while β indicates the slope gradient (in degrees). The topographic wetness index (TWI) was calculated using GIS software, offering insights into regions with greater moisture retention, which may be more prone to flooding (De Risi et al., 2014; Lee and Rezaie, 2022).

Topographic position index (TPI)

This morphometric factor, derived from the DEM (Fig. 5i), serves as a valuable indicator by quantifying the elevation difference between a specific cell and the mean elevation of its neighboring area within a defined radius (Calderón et al., 2020). The values are divided into five distinct classes utilizing the natural breaks method (Mokarram and Hojati, n.d.).

Stream power index

The stream power index (SPI) (Fig. 5g) measures the potential energy of flowing water to cause erosion and is calculated based on the slope and catchment area (flow accumulation) (Elmahdy et al., 2020). The results are visible in a raster with a 30-meter resolution.

Flow accumulation (FA)

Regions near flow accumulation paths (Fig. 5l), particularly those with substantial upstream water accumulation, are at a higher risk of flooding. The flow accumulation factor, obtained from the Digital Elevation Model (DEM), indicates the total weight of all upstream cells that contribute water flow to each downslope cell in the raster output ('FA2', n.d.; Ikirri et al., 2022; Zingaro et al., 2020). Cells with high flow accumulation signify areas of concentrated water flow, making them useful for detecting flow channels. In contrast, cells with zero flow accumulation represent local topographic elevations, which can be used to identify ridges or peaks.

Lithology

Lithology (Fig. 5n) significantly influences flood formation by affecting the permeability and water infiltration capacity of the terrain (Heitmuller et al., 2015; Langston and Temme, 2019; Mavromatis et al., 2016). The lithological distribution map of the Assaka watershed categorizes the region based on the permeability of the rock types. In this classification, lower values represent areas with higher permeability, whereas higher values indicate less permeable zones. The analysis of the lithological map shows that the study area is primarily composed of shales, limestones, sandstones, and quartzites, spanning several geological periods from the Precambrian to the Quaternary. This variety of lithological units has a direct impact on the hydrological behaviour of the area.

Land surface temperature (LST)

The presence of extensive impermeable surfaces, such as concrete and asphalt, often leads to increased land surface temperatures (Fig. 5j) in urban and developed regions (Diallo et al., 2019; Rahaman and Shermin, 2022). Because water cannot penetrate these surfaces, runoff is greater and flooding is more likely, particularly during periods of heavy rain. Furthermore, elevated LST can worsen evaporation and influence regional weather patterns, which in turn may cause heavier downpours (Diallo et al., 2024; Silvestro et al., 2013). regions with more vegetation or bodies of water tend to have lower land surface temperatures because they are better able to absorb and retain moisture compared to urbanized regions (Atta, 2023; Getirana et al., 2021). Soil and vegetation act as a natural flood control system, collecting water and slowly releasing it, therefore these circumstances often lessen the likelihood of flooding. Greater runoff from impermeable surfaces and less efficient land absorption makes regions with warmer land surfaces more likely to experience floods (Atta, 2023; Diallo et al., 2019, 2024; Getirana et al., 2021).

Data preprocessing

This stage aims to prepare the data for use in a machine learning algorithm and is divided into two sub-steps: data preparation followed by data cleaning.

Data preparation

fourteen raster (Fig. 5) images were processed by converting them into numerical values using Python libraries such as Pandas and NumPy within a GIS environment. This transformation yielded a dataset in which each pixel is linked with its respective X and Y coordinates, along with values for various factors (Masmoudi et al., 2021). The quality of this data is crucial for achieving accurate results. To ensure data integrity, it's essential to clean the data before use. A fundamental aspect of this cleaning process is the removal of outliers that significantly differ from the majority (Dietterich, 1990; F.Y et al., 2017). The challenge in identifying and eliminating these outliers stems from the lack of universally accepted statistical rules for their detection. Successful outlier detection relies on thorough knowledge of the subject matter and a comprehensive understanding of the data collection process (Mahesh, 2020).

Machine learning algorithm

Humans learn from past experiences, and now machines can be trained to do likewise. This concept forms the core of machine learning—a branch of artificial intelligence that allows machines to learn from historical data (Malakouti, 2023).

Random forest

The model employed in this case study is RF. A random forest consists of a set of decision trees, described as $\{h(x, \theta), \theta = 1,...\}$. Each $h(x, \theta)$ represents an individual decision tree, where θ denotes a specific tree within the collection. These trees independently issue predictions ($h(x, \theta)$) for a given input x (Kim et al., 2018; Savargiv et al., 2021; Zhu et al., 2018). The final prediction from the random forest is derived by aggregating the outputs of all these trees, typically by choosing the most common class among the different tree predictions (Avci et al., 2023; Han et al., 2022). Random Forests are effective across a wide array of problems. The core concept is to utilize multiple decision trees to mitigate the tendency of a single tree to overfit specific data segments (Boulesteix et al., 2012; Probstet al., 2018). By amalgamating various individual decision trees into a collective ensemble, a random forest can average out individual errors, thus minimizing the risk of overfitting (Naghibi, et al., 2017; Papineni et al., 2021). An advantage of random forests is that they do not necessitate data pre-processing. However, to achieve optimal performance, It is crucial to adjust key model parameters, such as the maximum depth of the trees and the maximum number of features considered (Patil and Singh, 2014).

Evaluation of the model's performance

Assessment is an essential part of the probabilistic modeling process; without it, the model's reliability cannot be ensured (Krause et al., 2017). For effective model training and evaluation, it is important to split the dataset into two parts: a training set and a testing set. In this case study, a split of 75% for training and 25% for testing is recommended (Kamal and Bablu, n.d.). Additionally, using a seed ensures that the results can be reproduced during this split. It is crucial to note that the percentage of data allocated for testing is very important (Azari et al., 2022). Ideally, it should be between 20% and 40% of the total data. Exceeding 40% may compromise the reliability of the tests, while less than 20% might be considered for large datasets. However, staying within these values is advised for more reliable and representative results (Coopera et al., 1997; Garosi et al., 2019). The method used to separate the data primarily involves random division, often referred to as the "holdout method". This random operation Partitions the data into two sets: one designated for training and the other for testing. In this example, a parameter of 0.25 means that 25% of the data is reserved for testing and 75% for training (Ibrahim and Bennett, 2014; Moein et al., 2023). There is no specific mathematical formula for this function, but it relies on statistical concepts to randomly divide the data, ensuring that both the training and test sets accurately represent the initial dataset. After splitting the data, it is necessary to standardize the features by centering and scaling them (Ismail et al., 2021; Oytun et al., 2020). Data standardization is a relatively straightforward mathematical operation that involves centering the data around zero and scaling it to have a unit variance. Here are the mathematical formulas for standardizing the data. To standardize a dataset X: Calculate the mean (k) of each feature, where xi represents each value of the feature and m Represents the overall count of samples (Equation 2). Calculate the standard deviation (ω) of each feature (Equation 3). The formula to standardize a specific value x is (Equation 4), This means that for each value x in your dataset, you subtract the mean of the feature (k) and divide it by the standard deviation (ω) of that feature. This centers the data around zero and scales it to have a variance of 2 (Coopera et al., 1997; Garosi et al., 2019; Ibrahim and Bennett, 2014).

$$k = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{2}$$

$$(3)\ \omega = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(x_i - k)^2} \tag{3}$$

$$x_{standardisé} = \frac{x - k}{\omega} \tag{4}$$

The evaluation process involves using multiple metrics such as Recall (Equation 6), F1 Score (Equation 8), Receiver Operating Characteristic Curve (Equation 10), Precision (Equation 5), Kappa Index (Equation 9), and Accuracy (Equation 7). These performance metrics are widely used in the literature.

$$Pr = \frac{TrPs}{TrPs + FaPs} \tag{5}$$

$$Re = \frac{TrPs}{TrPs + FaNe} \tag{6}$$

$$Ac = \frac{TrPs + TrNe}{TrPs + FaNe + TrNe + FaPs} \tag{7}$$

$$Fs = \frac{(2 * Pr * Re)}{(Pr + Re)} \tag{8}$$

$$Ka = \frac{P_d - P_{ep}}{1 - P_{ep}} \tag{9}$$

$$AUC = \frac{(\sum TrPs + \sum TrNe)}{(Pr + Nu)} \tag{10}$$

TrPs refers to true positives, *FaPs* to false positives, *TrNe* to true negatives, and *FaNe* to false negatives. *Pd* indicates the relative agreement among evaluators, *Pep* represents the expected probability of agreement by chance, *d* is the total number of flood pixels, and *Nu* is the total number of non-flood pixels. For the research Site case, the model's performance during the test period was rigorously evaluated using several key metrics. The results are as follows: precision: 0.968, recall: 0.967, accuracy: 0.967, F1 score: 0.965, Kappa statistic: 0.839, and AUC: 1. These metrics collectively indicate that the model performs exceptionally well. The high precision and recall rates reflect the model's capability to accurately determine positive instances, while the accuracy demonstrates its overall effectiveness. The F1 Score highlights a well-balanced performance between precision and recall, and the Kappa statistic

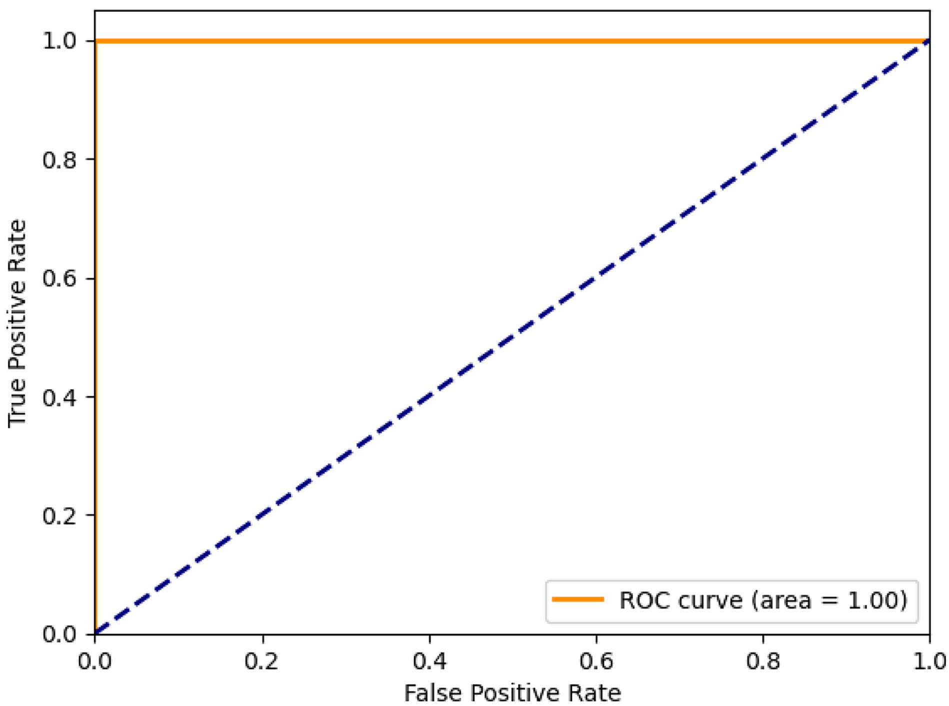


Figure 4. Receiver operating characteristic (ROC) curve

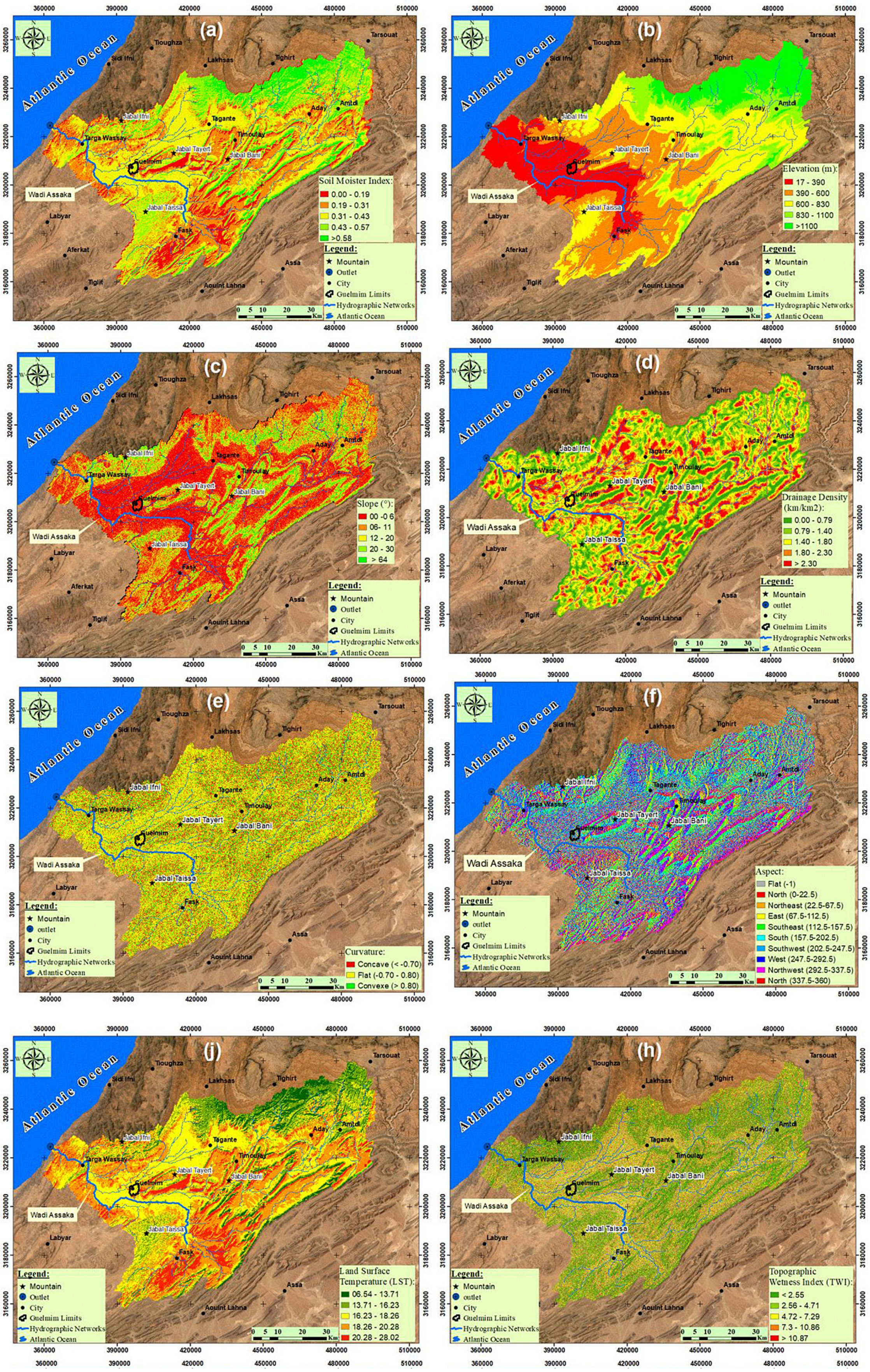


Figure 5. (a) soil moisture index (SMI); (b) altitude; (c) slope; (d) drainage density; (e) aspect; (f) curvature; (g) stream power index (SPI); (h) topographic wetness index (TWI)

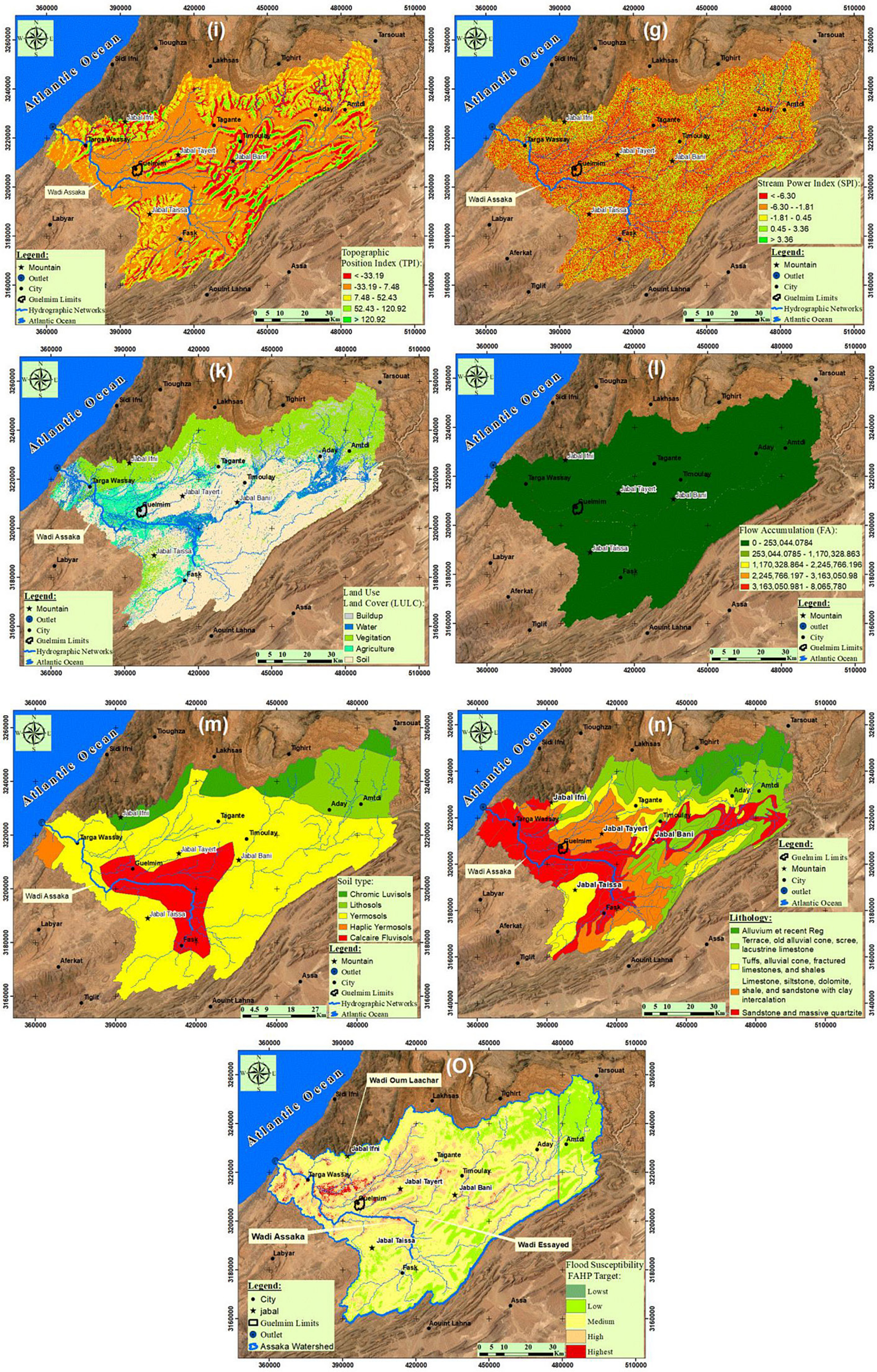


Figure 5. (i) topographic position index (TPI); (j) land surface temperature (LST); (k) land use and land cover (LULC); (l) flow accumulation (FA); (M) soil type; (N) lithology; (O) target

signifies a strong level of agreement beyond what would be expected by chance. The perfect AUC (Fig. 4) further underscores the model's superior capability in distinguishing between the classes, confirming its robust and reliable performance.

RESULTS AND DISCUSSION

Flash flood risk map

The final model for the Assaka basin, which incorporates various factors along with the random forest approach, generated a map identifying five zones prone to flooding (Fig. 6). The map indicates that residential areas located near the watershed outlet, particularly near rivers and the main tributaries Essayed and Oum Laachar Wadis (highlighted in red in the northwest), are highly vulnerable to flooding. These areas fall into the highest flash flood susceptibility category and are primarily situated in the southeastern section of the basin. The highest-risk zone, covering 14.34% of the basin's area, is primarily around the Essayed and Assaka wadis and downstream of the wadis that form these tributaries, such as the Essayed wadi. The high-risk zone occupies a vast random extent along the wadis that are tributaries of Essayed and Assaka, representing thus 35.06% of the total area. Additionally, the medium-risk zone encompasses a large part of urban areas and wadis, covering 28.80%. The other two zones, with low and very low risk of flooding, are primarily located in the northeastern part of the Assaka basin, occupying 21.52% of the area, where there are mainly high to very high slopes (Fig. 5c, Fig. 8a, Fig. 8b). These regions with high flood potential are characterized by very low elevations ranging between 17 and 390 meters, low drainage density of 0 to 0.79 km/km², and predominance of built-up areas and water bodies on the land use and land cover map. Additionally, these areas exhibit a low soil moisture index between 0 and 0.31, gentle slopes ranging from 0 to 0.6 degrees, and more concave curvatures. The high land surface temperatures, ranging from 18 to 28 degrees Celsius, along with a high topographic wetness index of more than 10.87, a low topographic position index of less than -33.19, and a low stream power index of less than -6.30, further contribute to flood risk. These locations also experience high flow accumulation. The presence of soil types such as Calcaric Fluvisols (Jc), Haplic Yermosols (Yh), and Yermosols (Y), which impede water infiltration and can rapidly also change soil saturation levels, exacerbates the situation. Furthermore, the composition of shales, limestones, sandstones, and quartzites in these areas affects their flooding dynamics.

The inherent qualities of these rock types, including porosity, permeability, and structural features, alongside local terrain and climatic conditions, make these regions particularly prone to flooding during rain events (Fig. 6 and Fig. 8).

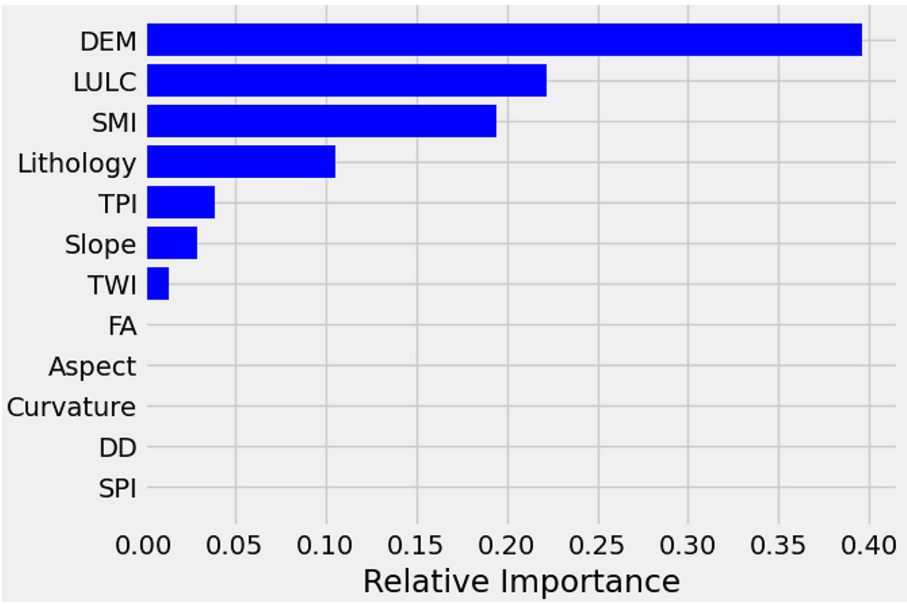


Figure 6. Feature importances

The feature importance analysis from the Random Forest model highlights the critical factors influencing flash flood risk in the Assaka Watershed. Among the evaluated variables, the digital elevation model (DEM) emerges as the most influential factor, with relative importance nearing 40%. This underscores the significant role that terrain elevation plays in determining flood susceptibility. Land use and land cover (LULC) and the soil moisture index (SMI) also contribute substantially to the model, emphasizing the impact of land cover changes and soil moisture variations on flood risk. Lithology, representing the geological characteristics of the area, is another important factor, while topographic indices like TPI (topographic position index), slope, and TWI (topographic wetness index) play more moderate roles. Features such as flow accumulation (FA), aspect, and drainage density (DD), however, have minimal influence on the model's predictions. This analysis provides valuable insights for prioritizing key variables in improving flood risk models, suggesting that elevation and land cover should be given particular attention in predictive frameworks (Fig. 7). In this study, we utilized advanced machine learning techniques, specifically the Random Forest model, to evaluate flood susceptibility in the Assaka watershed. This approach contrasts with and arguably enhances the methodologies applied in previous studies (Khaddari et al., 2023b), such as the one employing analytical hierarchy process (AHP) and fuzzy logic modeling (FLM) techniques detailed in the reviewed article. Although these conventional methods provided substantial outcomes, with AUC values indicative of considerable mapping accuracy, the precision and robustness of the Random Forest model offer improved discrimination and more nuanced classifications of flood-prone zones. The application of Random Forest in our study not only facilitated a refined assessment of the spatial pattern of flood susceptibility but also integrated a wider range of environmental variables effectively. Unlike AHP and FLM, which often rely on subjective weight assignments and simplifications in modeling interactions among multiple factors, the Random Forest approach inherently accommodates complex interdependencies among predictors. This capability results in a more dynamic and accurate depiction of flood risks, which is critical for creating effective flood risk mitigation strategies. Furthermore, our model identified critical zones with high precision, underlining areas that require urgent attention for mitigation efforts. This aligns with the need for advanced predictive tools that can provide local authorities with reliable information to make informed decisions about land use planning and disaster preparedness. The integration of machine learning into flood risk assessment,

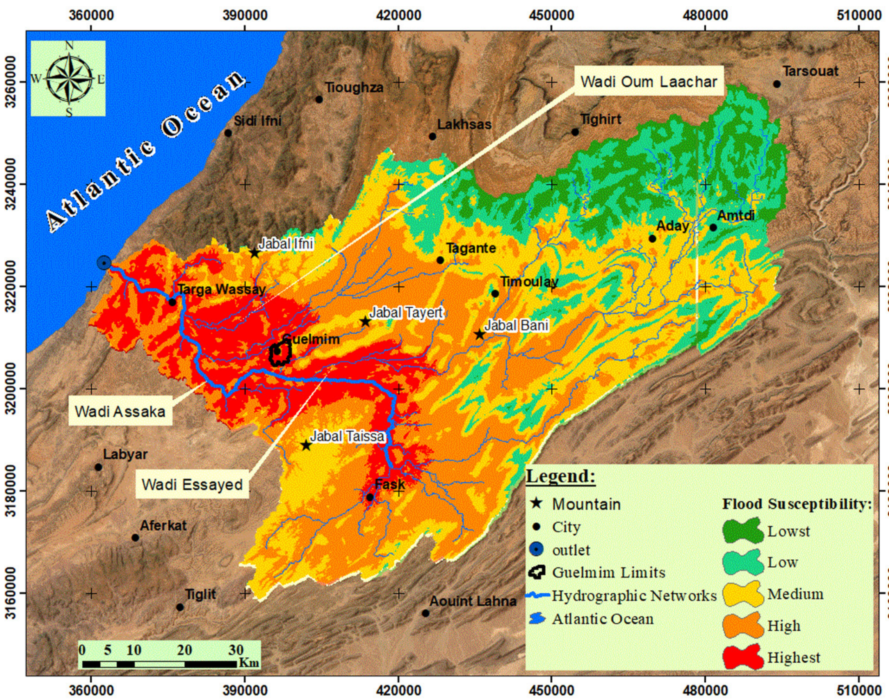
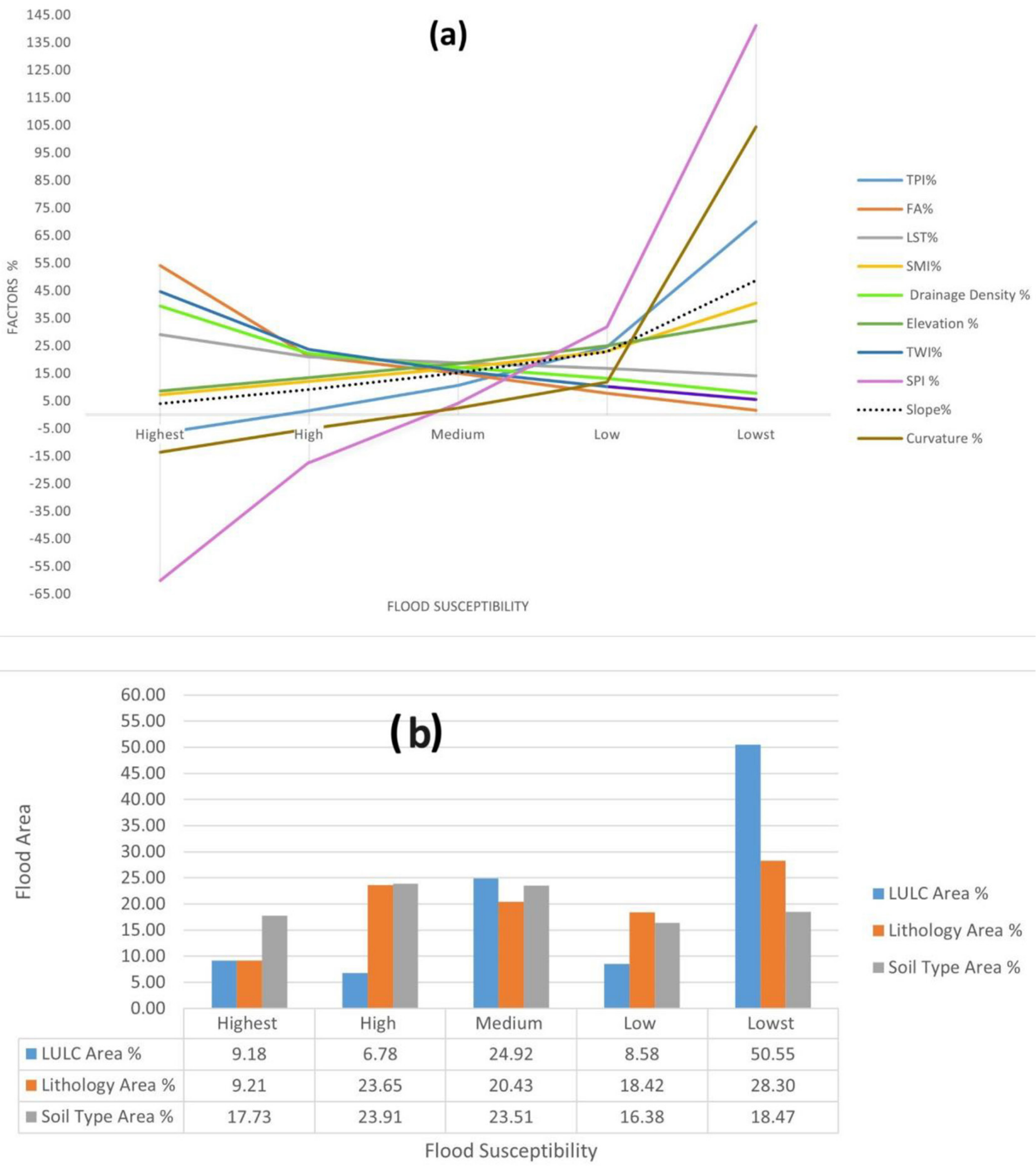


Figure 7. Flood susceptibility maps using RF model



	Highest	High	Medium	Low	Lowst
LULC Area %	9.18	6.78	24.92	8.58	50.55
Lithology Area %	9.21	23.65	20.43	18.42	28.30
Soil Type Area %	17.73	23.91	23.51	16.38	18.47

Figure 8. (a) and (b) flood susceptibility relative to the factors used

as demonstrated by our findings, represents a significant advancement over traditional methods, offering a pathway toward more resilient and adaptive management of flood-prone regions. The influence of climate on flood susceptibility is increasingly significant, particularly under the scenario of global climate change, which alters precipitation patterns and enhances flood risks. Research from the papers reviewed suggests that climate changes, particularly precipitation variability and the impact of the North Atlantic Oscillation (NAO), play a critical role in shaping flood patterns across various regions, including the Assaka watershed. Studies like those by (Said and Ahmed, 2023; Said et al., 2023, n.d.) demonstrate that shifts in precipitation are closely linked to broader atmospheric phenomena such as NAO, which directly affects rainfall distribution and intensity, ultimately influencing flood (Said and Ahmed, 2023; Said et al., 2023, n.d.). This climatic interplay, evidenced by extended periods of both drought and heavy rainfall, underscores the necessity of integrating robust climatic models into flood risk assessments. Utilizing advanced machine learning algorithms, as applied in this study, offers a nuanced understanding of how these climatic factors impact flood susceptibility, offering a more accurate and dynamic framework for predicting and managing flood risks in response to climate variability.

CONCLUSIONS

This study has aimed to enhance flash flood risk prediction within the Assaka watershed in southwestern Morocco by employing the RF machine learning algorithm integrated with GIS technologies. The research sought to address the limitations of conventional flood risk assessment methods and to provide a more accurate and data-driven approach for predicting flood-prone areas. By analyzing over 8 million data points and incorporating fourteen critical flash flood conditioning predictors, such as LST, drainage density, and LULC, the study successfully developed a reliable flood susceptibility map for the region. The results identified five distinct flood-prone zones, ranging from lowest to highest susceptibility, with the highest-risk areas located near the watershed's outlet and along the main tributaries, Essayed and Oum Laachar Wadis. These zones are characterized by specific geological and soil conditions, such as Calcaric Fluvisols and Haplic Yermosols, and factors like low drainage density and high land surface temperatures, which contribute to heightened vulnerability.

The findings of this study confirm that integrating the RF algorithm with GIS provides significantly improved prediction accuracy compared to traditional methods. This new approach offers a scientifically robust and reliable flood susceptibility model, thus achieving the study's goal of enhancing predictive capabilities for flash flood risk. The new scientific contribution of this research lies in its advanced integration of machine learning and geospatial data, which has enabled the identification of specific factors influencing flood susceptibility more accurately than previous studies.

The study fills a critical gap in existing flood management practices by providing a tool that local authorities, urban planners, and policymakers can use to prioritize conservation efforts and implement targeted flood preparedness strategies. Moreover, this methodology demonstrates a novel way to utilize sophisticated data analysis techniques to predict and mitigate natural disaster risks, contributing to enhanced resilience in vulnerable communities.

In conclusion, this research provides valuable insights into the application of ML and GIS in flood risk management, underscoring the scientific novelty of using advanced data-driven approaches to achieve more precise predictions. The framework developed here is not only applicable to the Assaka watershed but could be adapted to other regions facing similar challenges, opening new prospects for flood risk mitigation globally. This proactive approach aims to safeguard vulnerable communities and reduce the economic and human impact of future flooding events, thus contributing to sustainable disaster management strategies in the context of climate change.

Acknowledgments

I wish to convey my heartfelt appreciation to my esteemed professors, Ahmed Akhssas, Abdellatif Aarab, Hassan Rhinane, and Ayoub Aabi, for their invaluable support throughout this research journey. Their expert guidance, practical insights, and unwavering encouragement have been instrumental in the successful completion of this study. Beyond their academic contributions, their psychological support has been a source of great strength, helping me navigate the challenges of this work with confidence and resilience. I am also profoundly grateful to everyone who has supported me in this endeavor.

REFERENCES

1. Aangri, A., Hakkou, M., Krien, Y., Chtioui, T., & Benmohammadi, A. (2024). Risk Assessment of Marine Flooding along the Agadir and Taghazout Coasts (Moroccan Atlantic). *Journal of Coastal Research, 40*(1). https://doi.org/10.2112/jcoastres-d-23-00033.1
2. Abd El-Hamid, H. T., Kaloop, M. R., Abdalla, E. M., Hu, J. W., & Zarzoura, F. (2021). Assessment and prediction of land-use/land-cover change around Blue Nile and White Nile due to flood hazards in Khartoum, Sudan, based on geospatial analysis. *Geomatics, Natural Hazards and Risk, 12*(1), 1258–1286. https://doi.org/10.1080/19475705.2021.1923577
3. ABHSM of Morocco. (n.d.). ABHSM of Morocco.
4. Ait El Haj, F., Ouadif, L., & Akhssas, A. (2023). Simulating and predicting future land-use/land cover trends using CA- Markov and LCM models. *Case Studies in Chemical and Environmental Engineering, 7*. https://doi.org/10.1016/j.cscee.2023.100342
5. Al-Aizari, A. R., Alzahrani, H., AlThuwaynee, O.

F., Al-Masnay, Y. A., Ullah, K., Park, H. J., Al-Areeq, N. M., Rahman, M., Hazaea, B. Y., & Liu, X. (2024). Uncertainty reduction in flood susceptibility mapping using random forest and extreme gradient boosting algorithms in two tropical desert cities, Shibam and Marib, Yemen. *Remote Sensing, 16*(2). https://doi.org/10.3390/rs16020336

6. Albertini, C., Gioia, A., Iacobellis, V., Petropoulos, G. P., & Manfreda, S. (2024). Assessing multi-source random forest classification and robustness of predictor variables in flooded areas mapping. Remote Sensing Applications: *Society and Environment, 35*. https://doi.org/10.1016/j.rsase.2024.101239
7. Allafta, H., & Opp, C. (2021). GIS-based multi-criteria analysis for flood prone areas mapping in the trans-boundary Shatt Al-Arab basin, Iraq-Iran. *Geomatics, Natural Hazards and Risk, 12*(1), 2087–2116. https://doi.org/10.1080/19475705.2021.1955755
8. Atanga, R. A., & Tankpa, V. (2021, August 6). Climate change, flood disaster risk and food security nexus in northern ghana. frontiers in sustainable food systems. Frontiers Media S.A. https://doi.org/10.3389/fsufs.2021.706721
9. Atta, H. A. (2023). Seasonal Flooding Impact on Land Surface Temperature(LST) in Response to Vegetation Index Using Spectral Radiance Model in Kirkuk, Iraq. In *IOP Conference Series: Earth and Environmental Science* 1202. Institute of Physics. https://doi.org/10.1088/1755-1315/1202/1/012019
10. Avand, M., Kuriqi, A., Khazaei, M., & Ghorbanzadeh, O. (2022). DEM resolution effects on machine learning performance for flood probability mapping. *Journal of Hydro-Environment Research, 40*, 1–16. https://doi.org/10.1016/j.jher.2021.10.002
11. Avci, C., Budak, M., Yagmur, N., & Balcik, F. B. (2023). Comparison between random forest and support vector machine algorithms for LULC classification. *International Journal of Engineering and Geosciences, 8*(1), 1–10. https://doi.org/10.26833/ijeg.987605
12. Azari, B., Hassan, K., Pierce, J., & Ebrahimi, S. (2022). Evaluation of machine learning methods application in temperature prediction. *Computational Research Progress In Applied Science & Engineering, 8*(1), 1–12. https://doi.org/10.52547/crpase.8.1.2747
13. Bakhtiari, V., Piadeh, F., Behzadian, K., & Kapelan, Z. (2023, December 1). A critical review for the application of cutting-edge digital visualisation technologies for effective urban flood risk management. Sustainable Cities and Society. *Elsevier Ltd*. https://doi.org/10.1016/j.scs.2023.104958
14. Bannari, A., Bahi, H., & Rhinane, H. (2020). Evaluating land surface moisture conditions before and after flash-flood storm from optical and thermal data: models comparison and validation. In *International Geoscience and Remote Sensing Symposium (IGARSS)* 5038–5041. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/IGARSS39084.2020.9323655
15. Bannari, Abderrazak, Rhinane, H., & Bahi, H. (2019). Synergy between SMOS-MIRAS and Landsat-OLI/TIRS Data for Soil Moisture Mapping before, during, and after Flash-Flood Storm in Southwestern Morocco. In *Topics in Hydrometerology.* IntechOpen. https://doi.org/10.5772/intechopen.79554
16. Bano, S. (2024). Appraisal the Impact of Urban Evolution and Change on Land Use and Land Cover: A Case Study of Abbottabad District. Retrieved from https://www.researchgate.net/publication/382399124
17. Basri, H., Syakur, S., Azmeri, A., & Fatimah, E. (2022). Floods and their problems: Land uses and soil types perspectives. *In IOP Conference Series: Earth and Environmental Science* 951. IOP Publishing Ltd. https://doi.org/10.1088/1755-1315/951/1/012111
18. Bhattacharjee, N. (n.d.). Links among drainage density, frequency and flooding in Nanoi River Basin, Assam. *International Journal of Interdisciplinary Research in Science Society and Culture(IJIRSSC)*, *2*(1), 2016. Retrieved from www.ijirssc.in
19. Boulesteix, A.-L., Janitza, S., Kruppa, J., & König, I. R. (2012). Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics pre-review version of a manuscript accepted for publication in WIREs Data Mining & Knowledge Discovery. Retrieved from http://www.stat.uni-muenchen.de
20. Buttle, J. M., Allen, D. M., Caissie, D., Davison, B., Hayashi, M., Peters, D. L., Pomeroy, J. W., Simonovic, S., St-Hilaire, A., & Whitfield, P. H. (2016). Flood processes in Canada: Regional and special aspects. *Canadian Water Resources Journal, 41*(1–2), 7–30. https://doi.org/10.1080/07011784.2015.1131629
21. Calderón, H., Pérez, J. E., Chow, M., & Reyna, J. B. (2020). Índice de Posición Topográfica (TPI) para identificar áreas inundables y zonificación de especies vegetales en un manglar del Pacífico Sur de Nicaragua. *Revista Torreón Universitario, 9*(26), 124–146. https://doi.org/10.5377/torreon.v9i26.10264
22. Coopera, G. F., Aliferis, C. F., Ambrosinoa, R., Aronisb, J., Buchananb, B. G., Caruana', R., Fined, M. J., Glymour, C., Gordon, G., Hanusad, B. H., Janoskyf, J. E., Meek, C., Mitchell, T., Richardson, T., & Spirtes, P. (1997). An evaluation of machine-learning methods for predicting pneumonia

mortality. *Artificial Intelligence in Medicine 9*.

23. Cotti, D., Harb, M., Hadri, A., Aboufirass, M., Chaham, K. R., Libertino, A., Campo, L., Trasforini, E., Krätzschmar, E., Bellert, F., & Hagenlocher, M. (2022). An integrated multi-risk assessment for floods and drought in the Marrakech-Safi Region (Morocco). *Frontiers in Water, 4*. https://doi.org/10.3389/frwa.2022.886648
24. De Risi, R, Jalayer, F., De Paola, F., & Giugni, M. (2014). Probabilistic delineation of flood-prone areas based on a digital elevation model and the extent of historical flooding: the case of Ouagadougou.
25. De Risi, Raffaele, Jalayer, F., De Paola, F., & Lindley, S. (2018). Delineation of flooding risk hotspots based on digital elevation model, calculated and historical flooding extents: the case of Ouagadougou. *Stochastic Environmental Research and Risk Assessment, 32*(6), 1545–1559. https://doi.org/10.1007/s00477-017-1450-8
26. Diallo, I., Xue, Y., Chen, Q., Ren, X., & Guo, W. (2024). Effects of spring Tibetan Plateau land temperature anomalies on early summer floods/droughts over the monsoon regions of South East Asia. *Climate Dynamics, 62*(4), 2659–2681. https://doi.org/10.1007/s00382-021-06053-8
27. Diallo, I., Xue, Y., Li, Q., De Sales, F., & Li, W. (2019). Dynamical downscaling the impact of spring Western US land surface temperature on the 2015 flood extremes at the Southern Great Plains: effect of domain choice, dynamic cores and land surface parameterization. *Climate Dynamics, 53*(1–2), 1039–1061. https://doi.org/10.1007/s00382-019-04630-6
28. Dietterich, T. G. (1990). Machine Learning 1. Annual Review of Computer Science, 4.
29. Elmahdy, S., Ali, T., & Mohamed, M. (2020). Flash flood susceptibility modeling and magnitude index using machine learning and geohydrological models: A modified hybrid approach. *Remote Sensing, 12*(17). https://doi.org/10.3390/RS12172695
30. FA2. (n.d.).
31. F.Y, O., J.E.T, A., O, A., J. O, H., O, O., & J, A. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology, 48*(3), 128–138. https://doi.org/10.14445/22312803/IJCTT-V48P126
32. Ganjirad, M., & Delavar, M. R. (2023). Flood risk mapping using random forest and support vector machine. In ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 10, 201–208. *Copernicus Publications*. https://doi.org/10.5194/isprs-annals-X-4-W1-2022-201-2023
33. Garosi, Y., Sheklabadi, M., Conoscenti, C., Pourghasemi, H. R., & Van Oost, K. (2019). Assessing the performance of GIS- based machine learning models with different accuracy measures for determining susceptibility to gully erosion. *Science of the Total Environment, 664*, 1117–1132. https://doi.org/10.1016/j.scitotenv.2019.02.093
34. Getirana, A., Kumar, S. V., Konapala, G., & Ndehedehe, C. E. (2021). Impacts of fully coupling land surface and flood models on the simulation of large wetlands' water dynamics: the case of the Inner Niger Delta. *Journal of Advances in Modeling Earth Systems, 13*(5). https://doi.org/10.1029/2021MS002463
35. Ghouldan, A., Benhoussa, A., & Ichen, A. (2024). Evolution of land use/land cover in mediterranean forest areas – a case study of the Maamora in the North-West Morocco. *Ecological Engineering & Environmental Technology, 25*(10), 134–149. https://doi.org/10.12912/27197050/191413
36. Han, C., Ma, T., Xu, G., Chen, S., & Huang, R. (2022). Intelligent decision model of road maintenance based on improved weight random forest algorithm. *International Journal of Pavement Engineering, 23*(4), 985–997. https://doi.org/10.1080/10298436.2020.1784418
37. Hasan, M. M., Mondol Nilay, M. S., Jibon, N. H., & Rahman, R. M. (2023). LULC changes to riverine flooding: A case study on the Jamuna River, Bangladesh using the multilayer perceptron model. *Results in Engineering, 18*. https://doi.org/10.1016/j.rineng.2023.101079
38. Hawker, L., Bates, P., Neal, J., & Rougier, J. (2018). Perspectives on digital elevation model (DEM) simulation for flood modeling in the absence of a high-accuracy open access global DEM. *Frontiers in Earth Science, 6*. https://doi.org/10.3389/feart.2018.00233
39. HCP of Morocco. (2014). statistics of Assaka watershed.
40. He, X., Fang, Y., Wang, B., & Huang, X. (2024a). Towards a verifiable, uncertainty-controlled assessment of the spatiotemporal dynamics of social vulnerability to flash floods. *Ecological Indicators, 166*. https://doi.org/10.1016/j.ecolind.2024.112323
41. He, X., Fang, Y., Wang, B., & Huang, X. (2024b). Towards a verifiable, uncertainty-controlled assessment of the spatiotemporal dynamics of social vulnerability to flash floods. *Ecological Indicators, 166*. https://doi.org/10.1016/j.ecolind.2024.112323
42. Heerdegen, R. G., & Beran, M. A. (1982). Quantifying source areas through land surface curvature and shape. *Journal of Hydrology 57*.
43. Heitmuller, F. T., Hudson, P. F., & Asquith, W. H. (2015). Lithologic and hydrologic controls of mixed alluvial-bedrock channels in flood-prone fluvial systems: Bankfull and macrochannels in the Llano River watershed, central

Texas, USA. *Geomorphology, 232*, 1–19. https://doi.org/10.1016/j.geomorph.2014.12.033

44. Hermawan, W. G., Sunarto, Suryawan, I. W. K., & Suhardono, S. (2024). Hydrogeomorphological assessment of springs on the northern slope of Mount Merbabu. *Ecological Engineering and Environmental Technology, 25*(11), 30–43. https://doi.org/10.12912/27197050/191984
45. Ibrahim, A. M., & Bennett, B. (2014). The assessment of machine learning model performance for predicting alluvial deposits distribution. In *Procedia Computer Science 36*, 637–642. Elsevier B.V. https://doi.org/10.1016/j.procs.2014.09.067
46. Ikirri, M., Faik, F., Echogdali, F. Z., Antunes, I. M. H. R., Abioui, M., Abdelrahman, K., Fnais, M. S., Wanaim, A., Id-Belqas, M., Boutaleb, S., Sajinkumar, K. S., & Quesada-Román, A. (2022). Flood Hazard Index Application in Arid Catchments: Case of the Taguenit Wadi Watershed, Lakhssas, Morocco. *Land, 11*(8). https://doi.org/10.3390/land11081178
47. Ionno, A., Arsenault, R., Troin, M., Martel, J. L., & Brissette, F. (2024). Impacts of climate change on flood volumes over North American catchments. *Journal of Hydrology, 630*. https://doi.org/10.1016/j.jhydrol.2024.130688
48. Iresh, A. D. S., Dilthara, G. W. A. S., & Sugeeshwara, S. P. C. (2024). Estimation of catchment parameters of snyder's synthetic method for water resources planning and flood management in Sri Lanka. Engineer: *Journal of the Institution of Engineers, Sri Lanka, 57*(2), 81–86. https://doi.org/10.4038/engineer.v57i2.7595
49. Isma, F., Kusuma, M. S. B., Nugroho, E. O., & Adityawan, M. B. (2024). Flood hazard assessment in Kuala Langsa village, Langsa city, Aceh Province-Indonesia. *Case Studies in Chemical and Environmental Engineering, 10*. https://doi.org/10.1016/j.cscee.2024.100861
50. Ismail, L., Materwala, H., & Hennebelle, A. (2021). Comparative analysis of machine learning models for students' performance prediction. In *Advances in Intelligent Systems and Computing 1352*, 149–160. Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-71782-7_14
51. Jeong, J. H., Park, R. J., & Kim, S. B. (2003). Flooding Data Measured in Annular Narrow Gaps with Large Diameter of Curvature.
52. Kamal, M., & Bablu, T. A. (n.d.). International Journal of Applied Machine Learning and Computational Intelligence Machine Learning Models for Predicting Click-through Rates on social media: Factors and Performance Analysis.
53. Kang, S., Yeom, J., & Jung, J. (2021). Urban form and natural hazards: Exploring the dual aspect concept of urban forms on flood damage. Sustainability (Switzerland), 13(16). https://doi.org/10.3390/su13169007
54. Khaddari, A., Bouziani, M., Moussa, K., Sammar, C., Chakiri, S., Hadi, H. EL, Jari, A., & Titafi, A. (2022). Evaluation of precipitation spatial interpolation techniques using GIS for better prevention of extreme events: Case of the Assaka Watershed (Southern Morocco). *Ecology, Environment and Conservation, 28*(8), S1–S10. https://doi.org/10.53550/eec.2022.v28i08s.001
55. Khaddari, A., Jari, A., Chakiri, S., El Hadi, H., Labriki, A., Hajaj, S., Goumghar, L., El Harti, A., & Abioui, M. (2023a). A comparative analysis of analytical hierarchy process and fuzzy logic modeling in flood susceptibility mapping in the Assaka Watershed, Morocco. *Journal of Ecological Engineering, 24*(8), 62–83. https://doi.org/10.12911/22998993/165958
56. Khaddari, A., Jari, A., Chakiri, S., El Hadi, H., Labriki, A., Hajaj, S., Goumghar, L., El Harti, A., & Abioui, M. (2023b). A comparative analysis of analytical hierarchy process and fuzzy logic modeling in flood susceptibility mapping in the Assaka Watershed, Morocco. *Journal of Ecological Engineering, 24*(8), 62–83. https://doi.org/10.12911/22998993/165958
57. Kim, Y., Hardisty, R., Torres, E., & Marfurt, K. J. (2018). Seismic facies classification using random forest algorithm. In SEG Technical Program Expanded Abstracts 2161–2165. Society of Exploration Geophysicists. https://doi.org/10.1190/segam2018-2998553.1
58. Krause, J., Gulshan, V., Rahimy, E., Karth, P., Widner, K., Corrado, G. S., Peng, L., & Webster, D. R. (2017). Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. https://doi.org/10.1016/j.ophtha.2018.01.034
59. Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L. M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G. R., Kron, W., Benito, G., Honda, Y., Takahashi, K., & Sherstyukov, B. (2014). Le risque d'inondation et les perspectives de changement climatique mondial et régional. *Hydrological Sciences Journal, 59*(1), 1–28. https://doi.org/10.1080/02626667.2013.857411
60. Langston, A. L., & Temme, A. J. A. M. (2019). Bedrock erosion and changes in bed sediment lithology in response to an extreme flood event: The 2013 Colorado Front Range flood. *Geomorphology, 328*, 1–14. https://doi.org/10.1016/j.geomorph.2018.11.015
61. Lawford, R. G., Prowse, T. D., Hogg, W. D., Warkentin, A. A., & Pilon, P. J. (1995). Hydrometeorological aspects of flood hazards in Canada. *Atmosphere - Ocean, 33*(2), 303–328. https://doi.org/10.1080/0

7055900.1995.9649535

62. Lee, S., & Rezaie, F. (2022). Data used for GIS-based Flood Susceptibility Mapping. *GEO DATA, 4*(1), 1–15. https://doi.org/10.22761/dj2022.4.1.001
63. Loeb, R., Van Daalen, E., Lamers, L. P. M., & Roelofs, J. G. M. (2007). How soil characteristics and water quality influence the biogeochemical response to flooding in riverine wetlands. *Biogeochemistry, 85*(3), 289–302. https://doi.org/10.1007/s10533-007-9135-x
64. Mahesh, B. (2020). Machine Learning Algorithms - A Review. *International Journal of Science and Research (IJSR), 9*(1), 381–386. https://doi.org/10.21275/art20203995
65. Mahmouhi, N. El, Wartiti, M. El, Wissem, S. A., Kemmou, S., & El Bahi, S. (2016). Utilisation des systèmes d'information géographiques et des modèles hydrologiques pour l'extraction des caractéristiques physiques du bassin versant d'Assaka (Guelmim, sud du Maroc) [The use of geographic information system for the extraction of physical characteristics of assaka watershed: sub-basins of sayed and oum laachar wadis (southern Morocco) *]. International Journal of Innovation and Applied Studies, 16*. Retrieved from http://www.ijias.issr-journals.org/
66. Malakouti, S. M. (2023). Utilizing time series data from 1961 to 2019 recorded around the world and machine learning to create a Global Temperature Change Prediction Model. *Case Studies in Chemical and Environmental Engineering, 7*. https://doi.org/10.1016/j.cscee.2023.100312
67. Malik, S., Chandra Pal, S., Chowdhuri, I., Chakrabortty, R., Roy, P., & Das, B. (2020). Prediction of highly flood prone areas by GIS based heuristic and statistical model in a monsoon dominated region of Bengal Basin. *Remote Sensing Applications: Society and Environment, 19*. https://doi.org/10.1016/j.rsase.2020.100343
68. Mao, D., Wang, Z., Du, B., Li, L., Tian, Y., Jia, M., Zeng, Y., Song, K., Jiang, M., & Wang, Y. (2020). National wetland mapping in China: A new product resulting from object-based and hierarchical classification of Landsat 8 OLI images. *ISPRS Journal of Photogrammetry and Remote Sensing, 164*, 11–25. https://doi.org/10.1016/j.isprsjprs.2020.03.020
69. Masmoudi, O., Jaoua, M., Jaoua, A., & Yacout, S. (2021). Data Preparation in Machine Learning for Condition-based Maintenance. *Journal of Computer Science, 17*(6), 525–538. https://doi.org/10.3844/JCSSP.2021.525.538
70. Mathieu, J., Weisrock, A., Wengler, L., Elie Brochier, J., Even, G., Fontugne, M., Mercier, N., Ouammou, A., Senegas, F., Valladas, H., Holocene, al, Sénégas, F., Valladass, H., Vernets, J.-L., & Wahll, L. (2004). Holocene deposits in the lower section of the Assaka wadi, South Morocco : Preliminary results. Holocene deposits in the lower section of the Assaka wadi, south Morocco: preliminary results Holocene deposits in the lower section of the Assaka wadi, south Morocco: preliminary results •. *Quaternaire, 15*(2), 207–218. https://doi.org/10.3406/quate.2004.1768ï
71. Mavromatis, V., Rinder, T., Prokushkin, A. S., Pokrovsky, O. S., Korets, M. A., Chmeleff, J., & Oelkers, E. H. (2016). The effect of permafrost, vegetation, and lithology on Mg and Si isotope composition of the Yenisey River and its tributaries at the end of the spring flood. *Geochimica et Cosmochimica Acta, 191*, 32–46. https://doi.org/10.1016/j.gca.2016.07.003
72. Mawasha, T. S., & Britz, W. (2022a). Hydrological impacts of land use - land cover change on urban flood hazard: A case study of the Jukskei River in Alexandra Township, Johannesburg, South Africa. *South African Journal of Geomatics, 10*(2), 139–162. https://doi.org/10.4314/sajg.v10i2.11
73. Mawasha, T. S., & Britz, W. (2022b). Hydrological impacts of land use - land cover change on urban flood hazard: A case study of the Jukskei River in Alexandra Township, Johannesburg, South Africa. *South African Journal of Geomatics, 10*(2), 139–162. https://doi.org/10.4314/sajg.v10i2.11
74. Mohtasham Moein, M., Saradar, A., Rahmati, K., Ghasemzadeh Mousavinejad, S. H., Bristow, J., Aramali, V., & Karakouzian, M. (2023, January 1). Predictive models for concrete properties using machine learning and deep learning approaches: A review. *Journal of Building Engineering. Elsevier Ltd*. https://doi.org/10.1016/j.jobe.2022.105444
75. Mokarram, M., & Hojati, M. (n.d.). Comparison of Landform Classifications of Elevation, Slope, Relief and Curvature with Topographic Position Index in South of Bojnoord. https://doi.org/10.18869/modares.Ecopersia.4.2.1343
76. Mosavi, A., Ozturk, P., & Chau, K. W. (2018, October 27). Flood prediction using machine learning models: Literature review. Water (Switzerland). MDPI AG. https://doi.org/10.3390/w10111536
77. Muthusamy, M., Casado, M. R., Butler, D., & Leinster, P. (2021). Understanding the effects of Digital Elevation Model resolution in urban fluvial flood modelling. *Journal of Hydrology, 596*. https://doi.org/10.1016/j.jhydrol.2021.126088
78. Naghibi, S. A., Ahmadi, K., & Daneshi, A. (2017). Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. *Water Resources Management, 31*(9), 2761–2775. https://doi.org/10.1007/s11269-017-1660-3
79. Ngai, A. D., Ibrahim, R., & Selintung, M. (2024). Acid mine drainage treatment with organic

waste in constructed wetlands – effluent recirculation. *Ecological Engineering and Environmental Technology, 25*(8), 245–260. https://doi.org/10.12912/27197050/189817

80. Oytun, M., Tinazci, C., Sekeroglu, B., Acikada, C., & Yavuz, H. U. (2020). Performance Prediction and Evaluation in Female Handball Players Using Machine Learning Models. *IEEE Access, 8*, 116321–116335. https://doi.org/10.1109/ACCESS.2020.3004182
81. Papineni, S. L. V., Mallikarjuna Reddy, A., Yarlagadda, S., Yarlagadda, S., & Akkineni, H. (2021). An extensive analytical approach on human resources using random forest algorithm. *International Journal of Engineering Trends and Technology, 69*(5), 119–127. https://doi.org/10.14445/22315381/IJETT-V69I5P217
82. Parvin, M. I., Islam, M. A., Perveen, M. F., Jamal, M. R., Faruque, M. J., Hossen, B., Islam, K., & Sarker, M. M. (2024). Assessment of land use transition and crop intensification using geospatial technology in Bangladesh. *Case Studies in Chemical and Environmental Engineering, 9*. https://doi.org/10.1016/j.cscee.2024.100660
83. Patil, A., & Singh, S. (2014). Differential private random forest. In Proceedings of the 2014 *International Conference on Advances in Computing, Communications and Informatics*, *ICACCI 2014* 2623–2630. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICACCI.2014.6968348
84. Patil, P. P., Jagtap, M. P., Khatri, N., Madan, H., Vadduri, A. A., & Patodia, T. (2024). Exploration and advancement of NDDI leveraging NDVI and NDWI in Indian semi-arid regions: A remote sensing-based study. *Case Studies in Chemical and Environmental Engineering, 9*. https://doi.org/10.1016/j.cscee.2023.100573
85. Probst, P., Wright, M., & Boulesteix, A.-L. (2018). Hyperparameters and Tuning Strategies for Random Forest. https://doi.org/10.1002/widm.1301
86. Rahaman, S. N., & Shermin, N. (2022). Identifying the effect of monsoon floods on vegetation and land surface temperature by using Google Earth Engine. *Urban Climate, 43*, 101162. https://doi.org/10.1016/j.uclim.2022.101162
87. Riadi, B., Barus, B., Widiatmaka, Yanuar, M. J. P., & Pramudya, B. (2018). Identification and delineation of areas flood hazard using high accuracy of DEM data. In *IOP Conference Series: Earth and Environmental Science 149*. Institute of Physics Publishing. https://doi.org/10.1088/1755-1315/149/1/012035
88. Saha, A., Patil, M., Goyal, V. C., & Rathore, D. S. (2018). Assessment and Impact of Soil Moisture Index in Agricultural Drought Estimation Using Remote Sensing and GIS Techniques. 2. MDPI AG. https://doi.org/10.3390/ecws-3-05802
89. Saha, P., Mitra, R., Das, J., & Mandal, D. K. (2024). Urban flash flood prediction modelling using probabilistic and statistical approaches. *Results in Earth Sciences, 2*, 100032. https://doi.org/10.1016/j.rines.2024.100032
90. Said, A., & Ahmed, A. (2023). Drought variability in Agadir's region (Southern Morocco) – recent and future trends. *Ecological Engineering and Environmental Technology, 24*(2), 241–250. https://doi.org/10.12912/27197050/157169
91. Said, A., Ahmed, A., & Kharrim, M. El. (2023). Application of the cross-simulation method to the analysis of the rain-flow relationship from the gr2m-model: case of the Taghjijt Sub-Basin (Guelmim-Morocco), 18(3). Retrieved from www.arpnjournals.com
92. Said, A., Ahmed, A., Lahcen, B., & Rhita, B. (n.d.). Characterization of recent and future climatic trends in the region of Guelmim (Morocco). https://doi.org/10.1051/e3sconf/20
93. Sarker, I. H. (2021, May 1). Machine learning: algorithms, real-world applications and research directions. SN Computer Science. *Springer*. https://doi.org/10.1007/s42979-021-00592-x
94. Savargiv, M., Masoumi, B., & Keyvanpour, M. R. (2021). A new random forest algorithm based on learning automata. Computational Intelligence and Neuroscience, 2021. https://doi.org/10.1155/2021/5572781
95. Shayannejad, M., & Ostad, A.-A.-K. (n.d.). Flood routing in rivers by numerical modeling Preissmann scheme, saint-venant 1 equation and muskingum's method based on kinematic wave model 2 3 2. Retrieved from https://ssrn.com/abstract=4793265
96. Silvestro, F., Gabellani, S., Delogu, F., Rudari, R., & Boni, G. (2013). Exploiting remote sensing land surface temperature in distributed hydrological modelling: The example of the Continuum model. *Hydrology and Earth System Sciences, 17*(1), 39–62. https://doi.org/10.5194/hess-17-39-2013
97. Sivrikaya, F., Günlü, A., Küçük, Ö., & Ürker, O. (2024). Forest fire risk mapping with Landsat 8 OLI images: Evaluation of the potential use of vegetation indices. *Ecological Informatics, 79*. https://doi.org/10.1016/j.ecoinf.2024.102461
98. Sumi, T., Kantoush, S. A., & Saber, M. (n.d.). Natural Disaster Science and Mitigation Engineering: DPRI Reports Wadi Flash Floods Challenges and Advanced Approaches for Disaster Risk Reduction. Retrieved from http://www.springer.com/series/11773
99. Sun, D., Chen, Y., Wang, S., Zhang, H., Qiu, Z., Mao, Z., & He, Y. (2021). Using Landsat 8 OLI data to differentiate Sargassum and Ulva prolifera blooms in the South Yellow Sea. *International Journal of Applied Earth Observation and Geoinformation, 98*. https://doi.org/10.1016/j.jag.2021.102302

100. Talha, S., Maanan, M., Atika, H., & Rhinane, H. (2019). Prediction of flash flood susceptibility using fuzzy analytical hierarchy process (FAHP) algorithms and GIS: A study case of Guelmim Region in Southwestern of Morocco. In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives 42, 407–414. *International Society for Photogrammetry and Remote Sensing*. https://doi.org/10.5194/isprs-archives-XLII-4-W19-407-2019
101. Tan, X. Z., Li, Y., Wu, X. X., Dai, C., Zhang, X. L., & Cai, Y. P. (2024a). Identification of the key driving factors of flash flood based on different feature selection techniques coupled with random forest method. *Journal of Hydrology: Regional Studies, 51*. https://doi.org/10.1016/j.ejrh.2023.101624
102. Tan, X. Z., Li, Y., Wu, X. X., Dai, C., Zhang, X. L., & Cai, Y. P. (2024b). Identification of the key driving factors of flash flood based on different feature selection techniques coupled with random forest method. *Journal of Hydrology: Regional Studies, 51*. https://doi.org/10.1016/j.ejrh.2023.101624
103. Tarekegn, T. H., Haile, A. T., Rientjes, T., Reggiani, P., & Alkema, D. (2010). Assessment of an ASTER-generated DEM for 2D hydrodynamic flood modeling. *International Journal of Applied Earth Observation and Geoinformation, 12*(6), 457–465. https://doi.org/10.1016/j.jag.2010.05.007
104. Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2013). Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *Journal of Hydrology, 504*, 69–79. https://doi.org/10.1016/j.jhydrol.2013.09.034
105. Theilen-Willige, B., Charif, A., El Ouahidi, A., Chaibi, M., Ougougdal, M. A., & AitMalek, H. (2015). Flash floods in the guelmim area/southwest morocco–use of remote sensing and GIS-tools for the detection of flooding-prone areas. *Geosciences (Switzerland), 5*(2), 203–221. https://doi.org/10.3390/geosciences5020203
106. Yao, L., Chen, L., & Wei, W. (2017). Exploring the linkage between urban flood risk and spatial patterns in small urbanized catchments of Beijing, China. *International Journal of Environmental Research and Public Health, 14*(3). https://doi.org/10.3390/ijerph14030239
107. Zhang, D., & Zhou, G. (2016, August 17). Estimation of soil moisture from optical and thermal remote sensing: A review. Sensors (Switzerland). MDPI AG. https://doi.org/10.3390/s16081308
108. Zhu, M., Xia, J., Jin, X., Yan, M., Cai, G., Yan, J., & Ning, G. (2018). Class weights random forest algorithm for processing class imbalanced medical data. *IEEE Access, 6*, 4641–4652. https://doi.org/10.1109/ACCESS.2018.2789428
109. Zingaro, M., Refice, A., D'Addabbo, A., Hostache, R., Chini, M., & Capolongo, D. (2020). Experimental application of sediment flow connectivity index (SCI) in flood monitoring. *Water (Switzerland), 12*(7). https://doi.org/10.3390/w12071857