Journal of Environmental Management 364 (2024) 121291

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Flood susceptibility mapping of Cheongju, South Korea based on the integration of environmental factors using various machine learning approaches

Liadira Kusuma Widya[a,b], Fatemeh Rezaie[c,d], Woojin Lee[e], Chang-Wook Lee[a,**], Nurwatik Nurwatik[f], Saro Lee[c,d,*]

[a] *Department of Science Education, Kangwon National University, 1 Gangwondaehak-gil, Chuncheon-si, Gangwon-do, 24341, Republic of Korea*
[b] *Civil Engineering, Sunan Bonang University, Tuban, East Java, 62315, Indonesia*
[c] *Geoscience Data Center Division, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124, Gwahak-ro, Yuseong-gu, Daejeon, 34132, Republic of Korea*
[d] *Department of Geophysical Exploration, Korea University of Science and Technology, 217 Gajeong-ro, Yuseong-gu, Daejeon, 305-350, Republic of Korea*
[e] *College of AI Convergence, Dongguk University-Seoul, Jung-gu, Seoul, 04620, Republic of Korea*
[f] *Department of Geomatics Engineering, Institut Teknologi Sepuluh Nopember, 60111, Indonesia*

A R T I C L E I N F O

Keywords:
Flood susceptibility
Machine learning
GIS
LSTM
SVR
Boosted tree

A B S T R A C T

Floods are natural occurrences that pose serious risks to human life and the environment, including significant property and infrastructure damage and subsequent socioeconomic challenges. Recent floods in Cheongju County, South Korea have been linked to river overflow. In this study, we created flood susceptibility maps of Cheongju, South Korea using machine learning techniques including support vector regression (SVR), boosted tree (BOOST), and long short-term memory (LSTM) algorithms, based on environmental factors. Potentially influential variables were selected based on flood data gathered through field surveys; these included the slope, aspect, length–slope factor, wind exposition index, terrain wetness index, plan curvature, normalized difference water index, geology, soil drainage, soil depth, soil texture, land use type, and forest density. To improve the robustness of the flood susceptibility model, the most influential factors were identified using the frequency ratio method. Implementing machine learning techniques like SVR and BOOST produced encouraging outcomes, achieving the area under the curve (AUC) of 83.16% and 86.70% for training, and 81.65% and 86.43% for testing, respectively. While, the LSTM algorithm showed superior flood susceptibility mapping performance, with an AUC value of 87.01% for training and 86.91% for testing, demonstrating its robust performance and reliability in accurately assessing flood susceptibility. The results of this study enhance our understanding of flood susceptibility in South Korea and demonstrate the potential of the proposed approach for informing and guiding crucial regional policy decisions, contributing to a more resilient and prepared future.

1. Introduction

Floods have devastating consequences in terms of their destruction to infrastructure and economies (Allaire, 2018) and loss of human life (Jonkman, 2005; Yin and Li, 2001), and these effects are experienced on a global scale. Furthermore, floods are more frequent than natural disasters including drought, aualanches, and landslide (Hussain et al., 2023). Flooding has increased globally because of natural or modified drainage systems failing to manage water after heavy rainfall or long-term precipitation events (Pour et al., 2020). Although some factors that contribute to floods could be uncontrollable, the powerful techniques thoroughly depth studies for flood management and the implementation of land use management rules show potential in reducing flood hazards. The rising frequency and severity of these natural disasters have highlighted the need for disaster management strategies (Rana et al., 2021). Acknowledging the imperatives of proactive risk reduction and heightened disaster preparedness, the global community has increasingly embraced advanced technologies as critical

* Corresponding author. Geoscience Data Center Division, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124, Gwahak-ro, Yuseong-gu, Daejeon, 34132, Republic of Korea.
** Corresponding author.
E-mail address: leesaro@kigam.re.kr (S. Lee).

https://doi.org/10.1016/j.jenvman.2024.121291
Received 30 January 2024; Received in revised form 26 May 2024; Accepted 28 May 2024
Available online 13 June 2024

guidance in these endeavors (Khan et al., 2020). Flood susceptibility mapping provides a sophisticated, data-driven approach for assessing possible flood risks and implementing proper mitigation measures (Chen et al., 2020).

Natural disasters are frequent in Korea. Notably, Cheongju County is prone to flooding, particularly in the monsoon season, which exhibits increased heavy rainfall frequency. In July 2023, an intense rainfall resulted in river overflow and significant flooding. The damage caused by this overflow led to the inundation of a tunnel, which caused the loss of life of several trapped individuals. The Korea Meteorological Administration (KMA) had issued heavy rain warnings during this event, reporting a maximum hourly rainfall of 33.5 mm that lasted for three days, until July 16, 2023.

Developing a comprehensive strategy for understanding flood dynamics necessitates the consideration of various factors influencing flood susceptibility (Tehrany et al., 2015). The ability to predict intricate relationships within extensive datasets offers a promising approach for incorporating diverse parameters, encompassing elements such as topography, land use, rainfall, soil characteristics, and geology (Tehrany et al., 2015). Multidimensional analysis enables a more nuanced understanding of the interactions among these variables and their impact on flood susceptibility.

The integration of geospatial technology with remote sensing (RS), geographic information system (GIS), and Global Positioning System (GPS) data has emerged as an effective approach for flood management (Mondal and Sahoo, 2022) and provides decision makers with essential information supporting a comprehensive approach to resilience and adaptation in the face of evolving challenges (Albano and Sole, 2018). A GIS can be used to gather and analyze spatial data to allow for more informed decision-making regarding disaster mitigation and rapid response, thereby enhancing overall resilience and preparedness (Graymore et al., 2009). GIS technology encompasses various techniques, including statistics (Machiwal et al., 2018) and machine learning (Du et al., 2020). Statistical methods, such as the frequency ratio (FR) technique, expand the analytical potential of GIS (Arshad et al., 2020)). Previous studies of the effectiveness of FR models in groundwater potential mapping using GIS have shown that FR models provide higher accuracy than analytical hierarchy models. Thus, FR models are capable of handling spatial correlations among variables and offer several advantages for spatial data analysis (Lee and Sambath, 2006).

The application of machine learning algorithm techniques further enhances the capability of GISs in complex and adaptive spatial analyses such as flood susceptibility mapping (Islam et al., 2021). Advanced machine learning methods allow for sophisticated analyses of diverse and intricate factors influencing flood risk (Dodangeh et al., 2020). Machine learning algorithms, including Naïve Bayes tree, alternating decision tree, random forest methods (Chen et al., 2020), artificial neural networks (Priscillia et al., 2021), and support vector machine (Dodangeh et al., 2020) algorithms can identify intricate patterns in large datasets, leading to more accurate predictions of disaster susceptibility. Thus, the integration of machine learning algorithms allows for more complex evaluation of the interplay between flood occurrences and various factors such as topography, land use, and rainfall (Islam et al., 2021), leading to the development of highly precise and reliable flood susceptibility maps. For example, dagging and random subspace models coupled with an artificial neural network, random forest, or support vector machine algorithm were applied to assess flood susceptibility, with an area under the receiver operating characteristic curve (AUROC) consistently exceeding 0.80 for all models (Islam et al., 2021).

Previous studies have utilized various machine learning models to assess and generate flood susceptibility maps. SVR is a type of machine learning algorithm categorized under supervised learning, which aims to find the best-fitting line or hyperplane to predict outcomes accurately (Dodangeh et al., 2020). For instance, a previous study employed Support Vector Regression (SVR) as part of supervised machine learning to map flood susceptibility, achieving an AUC of 83% for the training dataset and 82% for the testing dataset (Rezaie et al., 2022). Another study utilized the Boosted Tree algorithm, a form of ensemble learning (Chen et al., 2021), to map flood susceptibility in the Seoul metropolitan city area of South Korea. This analysis yielded validation accuracies of 77.55% and 77.26% for the regression and classification algorithms, respectively (Lee et al., 2017). On the other hand, deep learning, a subset of machine learning, has emerged as a powerful approach for processing complex data efficiently. Within the realm of deep learning, recurrent neural networks (RNNs) play in sequence modeling, with long short-term memory (LSTM) networks being a specialized architecture within RNNs (Konget al., 2019). A previous study applied LSTM for landslide susceptibility mapping, a validated AUC value of 0.832 was obtained (Hamedi et al., 2022).

The objective of this study was to employ machine learning algorithms, specifically support vector regression (SVR), boosted tree (BOOST), and long short-term memory (LSTM) algorithms, to create flood susceptibility maps that will enhance our understanding and prediction of flood susceptibility. In contrast to conventional techniques, this study leverages state-of-the-art technologies to generate flood susceptibility maps of accuracy and reliability. The utilization of machine learning algorithms signifies a progressive step forward in the assessment of flood risk, thereby enhancing disaster preparedness and response initiatives in regions susceptible to flooding. Moreover, contributing factors such as soil characteristics, vegetation, land cover, and topographic, hydrological, and geological factors were considered by the algorithms in flood risk assessments. Through their capacity to discern complex relationships within vast datasets, these algorithms provided a more nuanced and dynamic assessment of flood risk. This information is invaluable for urban planning, infrastructure development, and disaster management strategies, to empower communities and authorities with actionable insights and enable the implementation of targeted measures for flood mitigation in vulnerable regions.

2. Materials and methods

2.1. Study area

This study was conducted within an area of 984 km^2 in Chungju County, Chungcheongbuk-do, South Korea (Fig. 1). Cheongju County is located on a large alluvial plain bounded by two rivers and includes infrastructure such as the Chungju Dam. The landscape in the eastern region tends to be mountainous, featuring the Uam, Gunyeo, and Sandang Mountains. The rainy season in Cheongju City is from June to September, with the peak occurring in July. The annual precipitation exceeds 1500 mm. The climate is temperate, with warm, humid weather during the summer season. In July 2023, the KMA recorded the highest rainfall amount of 33.5 mm over a span of three days in Cheongju County.

Spatial modeling of flood susceptibility necessitates a foundational flood inventory map delineating flood occurrence. In this study, the flood inventory map was created utilizing 3115 flood points, identified through existing reports, and further validated through field surveys (Fig. 1). Subsequently, these points were randomly divided into separate categories for testing and training. Based on the previous studies, it was determined that an equitable distribution, with 50% of the points allocated for model training and an equivalent 50% for model testing (Nurwatik et al., 2022; Arabameri et al., 2020).

Flood occurrence is influenced by numerous factors including topographic (Khosravi et al., 2018), hydrological (Sharma et al., 2014), and geological (Pradhan et al., 2023) factors, as well as soil characteristics, vegetation, and land cover (Niehoff et al., 2002). The selection of variables for susceptibility modeling relies on the specific region under consideration and the available data. Several studies have adopted diverse sets of variables, and numerous review articles offer comprehensive analyses of their selection (Dodangeh et al., 2020; Fang et al., 2021; Pradhan et al., 2023). In this study, several influencing factors

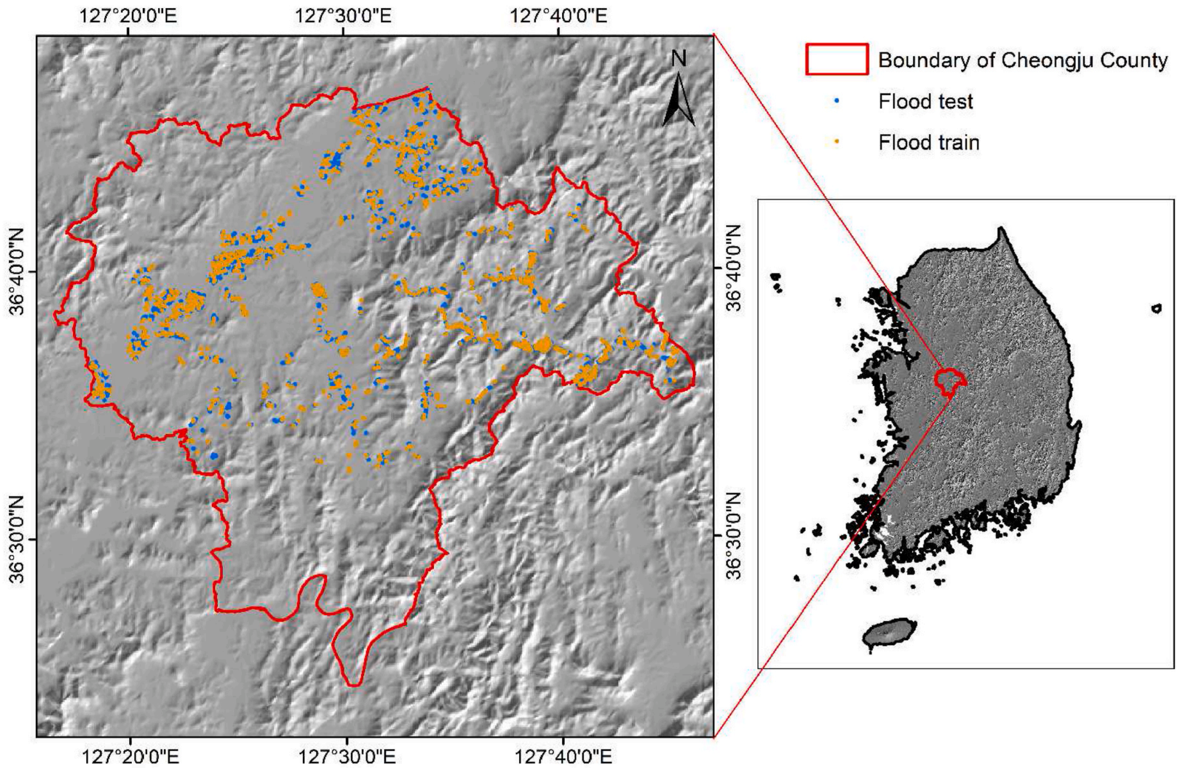

Fig. 1. Study site (red boundary) within Cheongju County, Chungcheongbuk Province, South Korea. The colors of orange and blue indicate past flood occurrences. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

were considered using frequency ratio methods for generating flood susceptibility maps using machine learning algorithms (Table 1).

Topographic factors (Fig. 2 (a)–2 (f)) used in flood susceptibility mapping have included the slope (Bajabaa et al., 2014), aspect, length–slope (LS) factor, wind exposure index, terrain wetness index (TWI), and plan curvature. The slope of a terrain affects the speed and direction of surface water runoff, impacting flood potential. The LS factor, which is derived from both slope length and angle, provides information about terrain steepness (Popa et al., 2019). Aspect refers to the direction faced by a slope; it affects the solar radiation distribution and subsequently impacts surface runoff and water accumulation. Wind exposure plays a role in altering surface water flow patterns and influencing flood susceptibility in specific areas. TWI reflects the wetness of the terrain, affecting water movement. Plan curvature indicates the curvature of the terrain, and contributes to an understanding of how water flows across the landscape.

Table 1
Factors that impact the prediction of flood susceptibility maps in Cheongju County utilizing machine learning algorithms.

Data sources	Type of factors	Factors
Ministry of Land, Infrastructure and Transport (MOLIT), Korea	Topographic Factors	Slope Aspect LS Factor Wind Exposition TWI Plan curvature
Sentinel-2 sattelite imagery	Hydrological Factors	NDWI
Korea Institute of Geoscience and Mineral Resources (KIGAM), Korea	Geological Factors	Material Geology
National Institute of Agricultural Sciences, Korea	Soil Characteristics	Soil drainage Soil depth Deep soil texture Surface soil texture
Ministry of Environment, Korea Forest Service, Korea	Vegetation and Land Cover	Land use type Forest density

Hydrological parameters play a key role in determining flood susceptibility. One of the main hydrological parameters is the normalized difference water index (NDWI) (Sharma et al., 2014). Enhanced indices were related to water bodies were modeled using Google Earth Engine. Various combinations of spectral bands were employed to identify water body features in the study area. The NDWI map of Cheongju County (Fig. 2 (g)). We classified the NDWI values throughout the research region into five groups using quantile classification. This technique was used to enhance the efficiency of interpretation and analysis.

Geological factors (Fig. 2 (h)–2 (i)) such as material and lithology can also influence flood occurrence. The type and composition of geological materials in an area impacts how water interacts with the terrain and affects drainage patterns and flooding susceptibility. For example, some lithological formations may be more prone to water retention or infiltration, influencing the overall hydrological dynamics of the region (Ferreira et al., 2015). Understanding the effects of geological characteristics such as rock type and soil composition on flooding is essential for comprehensive flood susceptibility mapping.

Soil-based parameters (Fig. 2 (j)–2 (m)) such as soil drainage, soil depth, surface soil texture, and deep soil texture were obtained from the National Institute of Agricultural Sciences, South Korea. Soil depth was divided into five classes: <20, 20–50, 50–100, >100 cm, and no soil. Deep soil texture was divided into six classes: sand, sandy loam, silt loam, clay loam, silty clay loam, and clay. Surface soil texture was divided into 10 classes: loamy coarse sand, loamy fine sand, loamy sand, fine sandy loam, sandy loam, loam, clay loam, silt loam, silty clay loam, and other. Different soil types have distinct hydrological characteristics that affect the infiltration rate and surface runoff patterns (Hümann et al., 2011). Soil drainage conditions were categorized as very good, good, slightly good, slightly poor, poor, very poor, or other. Soil drainage is crucial for understanding flood occurrence, as it directly

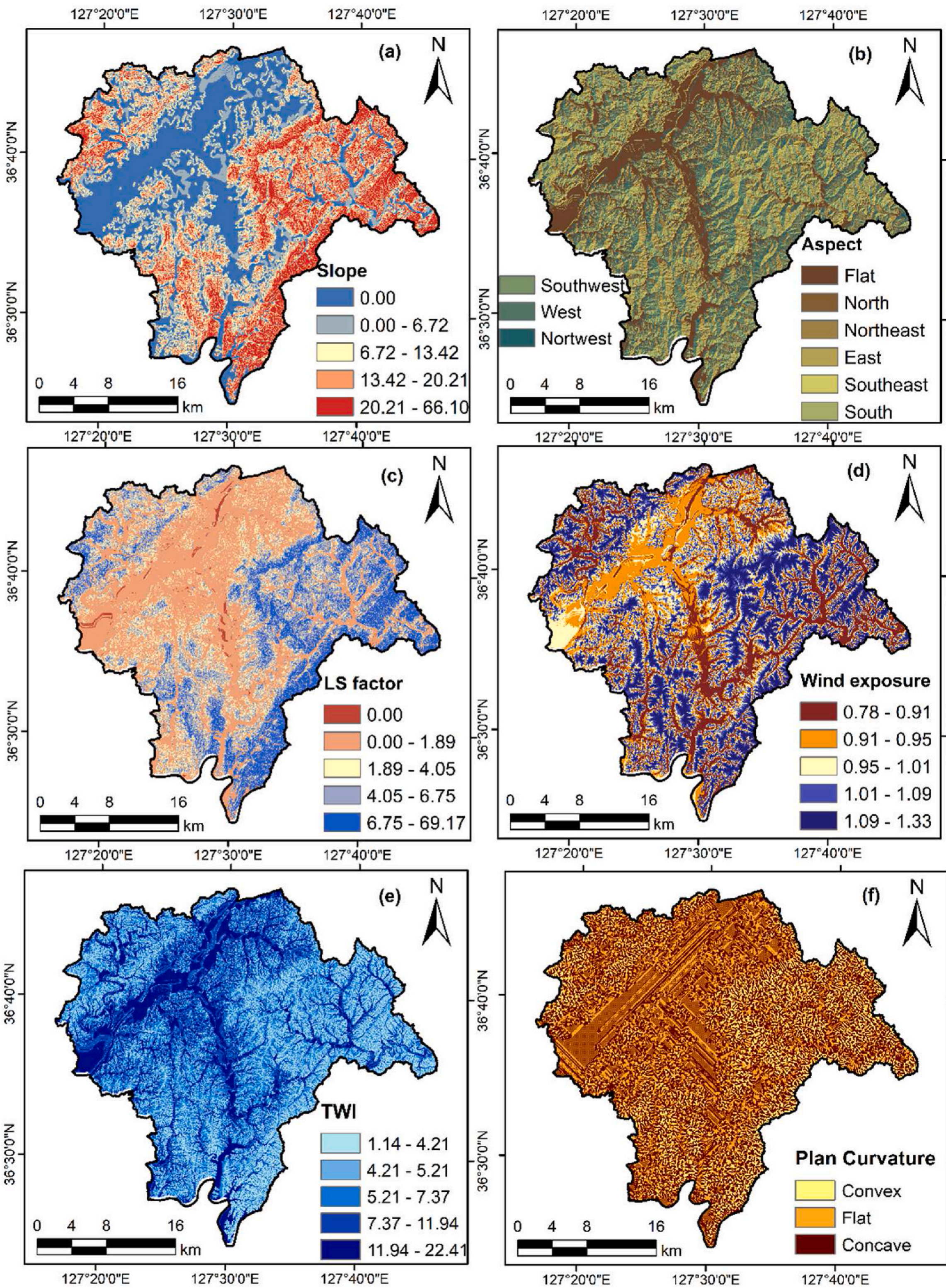

Fig. 2. Flood related factors (a) Slope, (b) Aspect, (c) LS factor (d) Wind exposure, (e) TWI (f) Plan curvature, (g) NDWI, (h) Lithology, (i) Material, (j) Drainage, (k) Soil depth, (l) Deep soil texture, (m) Surface soil texture, Forest densities, and (o) Land cover.

influences water movement and soil absorption capacity (Kaur et al., 2020).

Variables related to vegetation (Fig. 2 (n)–2 (o)) considered in the flood susceptibility models included land cover type and density, which tend to have great influence over flood risk (Tehrany et al., 2015). Forest density was divided into four coverage categories: 50%, 51–70%, ≥71%, and non-vegetation (i.e., no coverage). These categories represent different stages of forest growth, with varying degrees of impact on soil stability and water absorption. The intricate relationship between land use and flood occurrence is rooted in the condition and health of the

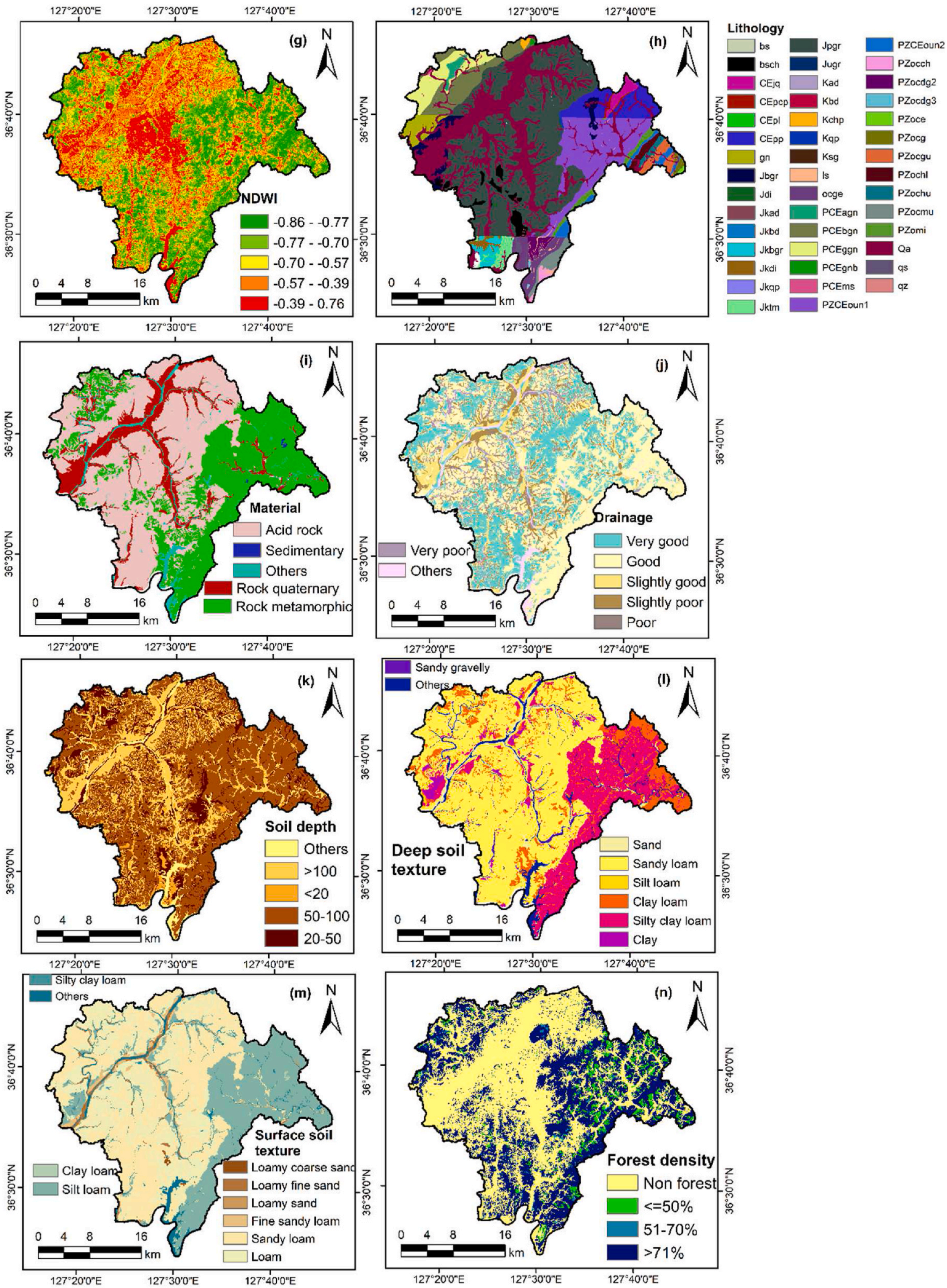

Fig. 2. (*continued*).

natural landscape. Specific land use types, e.g., urbanization, agriculture, and forestry, can have significant impacts on hydrological processes, altering water flow patterns, infiltration rates, and runoff characteristics (Alexakis et al., 2014).

2.2. Methodology

In this study, potentially influential variables were evaluated using the FR method. Flood susceptibility maps were created by applying three machine learning algorithms: SVR, BOOST, and LSTM. The

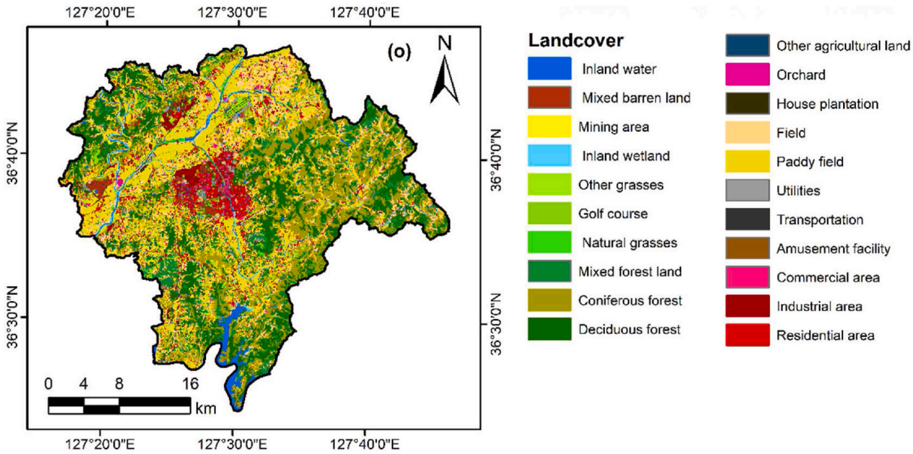

Fig. 2. (*continued*).

accuracy of the model predictions for each algorithm was assessed according to the AUROC of each model.

2.2.1. Impacts of factors influencing flood occurrence

The FR technique is a statistical analysis method for assessing relationships between a phenomenon and its associated variables in a bivariate context (Lee and Talib, 2005) and is widely used in various fields including landslide susceptibility mapping (Ding et al., 2016), flood mapping (Tehrany et al., 2019), and urban planning (Park et al., 2012). In this study, FR values were calculated for each category of factors based on their correlation with flood susceptibility. The FR was calculated by dividing the area of the occurrence related to the flood variable subclass by the total study area within that subclass, as follows (Addis, 2023a):

$$FR = \frac{N_{(F_i)}/N_{(S_i)}}{N_{(F)}/N_{(T)}}, \tag{1}$$

where $N_{(F_i)}$ is the number of flood occurrence pixels in subclass i of the influencing factor; $N_{(S_i)}$ is the total number of pixels in subclass i; $N_{(F)}$ is the overall distribution of flood occurrence for the influencing factor, and $N_{(T)}$ is the number of pixels of the total area. An FR value of 1 indicates an average relationship between the influencing factor and the flood variable, and FR $>$ 1 indicates a strong relationship between environmental conditions and the flood variable, whereas FR $<$ 1 indicates a weak relationship.

2.2.2. Support vector regression algorithm

The SVR technique, which is frequently employed in machine learning applications, was developed based on principles proposed (Cortes and Vapnik, 1995). The SVR method uses structural risk reduction, allowing for adjustment of the training data size to the classifier's potential, thereby enhancing generalization performance. In this study, we adopted a radial basic function (RBF) kernel, also known as a Gaussian kernel, as a crucial component of SVR for handling nonlinear relationships between the input features and the target variable (Dodangeh et al., 2020), by transforming the input features into a higher-dimensional space where a linear relationship can be established. The RBF kernel is particularly effective in capturing complex patterns and relationships that may not be linearly separable in the original feature space (Ruiz and Lopez-de-Teruel, 2001). SVR is widely used in predicting natural occurrences such as landslides (Balogun et al., 2021), flood susceptibility map (Rezaie et al., 2022; Saha et al., 2021) and groundwater potential mapping (Panahi et al., 2020). Previous study in flood susceptibility mapping, SVR uses digital elevation model, sentinel-2 image, rainfall, soil type data to identify flood prone regions (Saha et al., 2021).

2.2.3. Boosted tree algorithm

The BOOST algorithm is a machine learning method that utilizes ensemble learning (Chen et al., 2021), and is widely used for stochastic gradient boosting. The BOOST algorithm was developed based on TreeNet, a main numerical tool of deterministic-gradient boosting (Oh and Lee, 2017). The purpose of this sequential technique is to improve the performance of the model by correcting errors committed in previous iterations. The technique utilizes a binary tree structure, implementing a comprehensive classification and regression tree model. The data are then partitioned into two separate samples at each split node, enabling a thorough examination of the feature space. Each split node has a vital function in determining the best location to divide the data, strategically computing observed deviations and residuals for each partition. Thorough examination at each stage enhances the algorithm's capacity to evaluate intricate connections within the data. The BOOST algorithm has been shown to effectively perform regression tasks and predictive modeling by iteratively improving its predictions and emphasizing the most pertinent features of the dataset. It has been applied in landslide susceptibility (Kim et al., 2018), flood susceptibility (Lee et al., 2017), and erosion susceptibility (Chen et al., 2021).

2.2.4. Long short-term memory algorithm

Another category of deep neural network is the recurrent neural network (RNN), in which the network output is looped back as subsequent input (Kong et al., 2019). LSTM is a widely used RNN that is constructed based on a model or structure designed for sequential data (Wang et al., 2020). The LSTM design integrates specialized memory cells with gating mechanisms, enabling the model to retain and update information over extended sequences without compromising crucial details. The key components of an LSTM cell include input gates (i_t), forget gates (f_t), cells ($\widetilde{C}_t$), output gates (o_t), and cell output state (C_t) (Graves, 2012).

The input gate determines how much new information (input data) should be added to the cell state, as described by Eq. (2). The subsequent methods involve determining the value for the forget gate, which represents the activation of the forget gates of memory cells. Forget gate activation (f_t) is computed using Eq. (3), by taking the product of the concatenated previous hidden state (h_{t-1}) and the current input (x_t). The candidate cell state ($\widetilde{C}_t$) is represented in Eq. (4) as new information that could be added to the cell state (C_t) in Eq. (5). Thus, the following equations define the process by which a layer of memory cells improves at each time step, denoted as t:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \tag{2}$$

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right), \tag{3}$$

$$\widetilde{C}_t = tanh(W_c[h_{t-1}, x_t] + b_c), \tag{4}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \ . \tag{5}$$

The output gate (o_t) in Eq. (6) determines how much of the updated cell state (C_t) should be exposed as the output of the current time step. By considering the memory cells of the output state (C_t), the output gate (h_t) values can be determined using Eq. (7), as follows:

$$o_t = \sigma(W_o\ [h_{t-1}, x_t] + b_o), \tag{6}$$

$$h_t = o_t \odot \tanh(C_t), \tag{7}$$

where W (W_f, W_i, W_c, W_o) are the weighted matrices, b (b_f,b_i,b_c, b_o) are the bias vectors, σ is a sigmoid function, x_t is the input to the memory cell layer at time t, $\odot$ is the operation of element-wise multiplication, and *tanh* is a hyperbolic tangent function. Several investigations have indicated that the LSTM method is widely recognized in ecological and environmental studies, particularly well-suited for modeling sequences and time series data. For instance, previous study employeda a local spatial sequential LSTM to forecast flood susceptibility (Fang et al., 2021), while other studies utilized LSTM method for tasks such as landslide susceptibility mapping (Hamedi et al., 2022), and forest fire prediction (Lin et al., 2023; Natekar et al., 2021).

2.2.5. Model performance validation and comparison

The AUC is an effective indicator for assessing the predictive accuracy of machine learning frameworks (Oh and Lee, 2017). The model results must be validated using data not applied during the training phase to ensure the scientific reliability of the outcomes (Panahi et al., 2020). The AUC is calculated on both training and testing datasets to determine the success and prediction rates, respectively. The success rate indicates the model's ability to faithfully depict the data collected; the prediction rate curve illustrates the model's efficacy in producing accurate predictions (Arora et al., 2021a).

3. Results

3.1. Relationships between flooding occurrence and related factors

The FR model was employed to evaluate the relationship between the incidence of floods and the characteristics of contributing factors. The distribution of flood occurrences was examined through FR methodologies to identify the impact of potentially influential variables. Topographic factors that affected flood occurrence included slope, aspect, LS factor, wind exposure, TWI, and plan curvature (Fig. 3). Additionally, (Table S1) provides a detailed description of the weighting for the FR. Previous study has demonstrated the correlation between

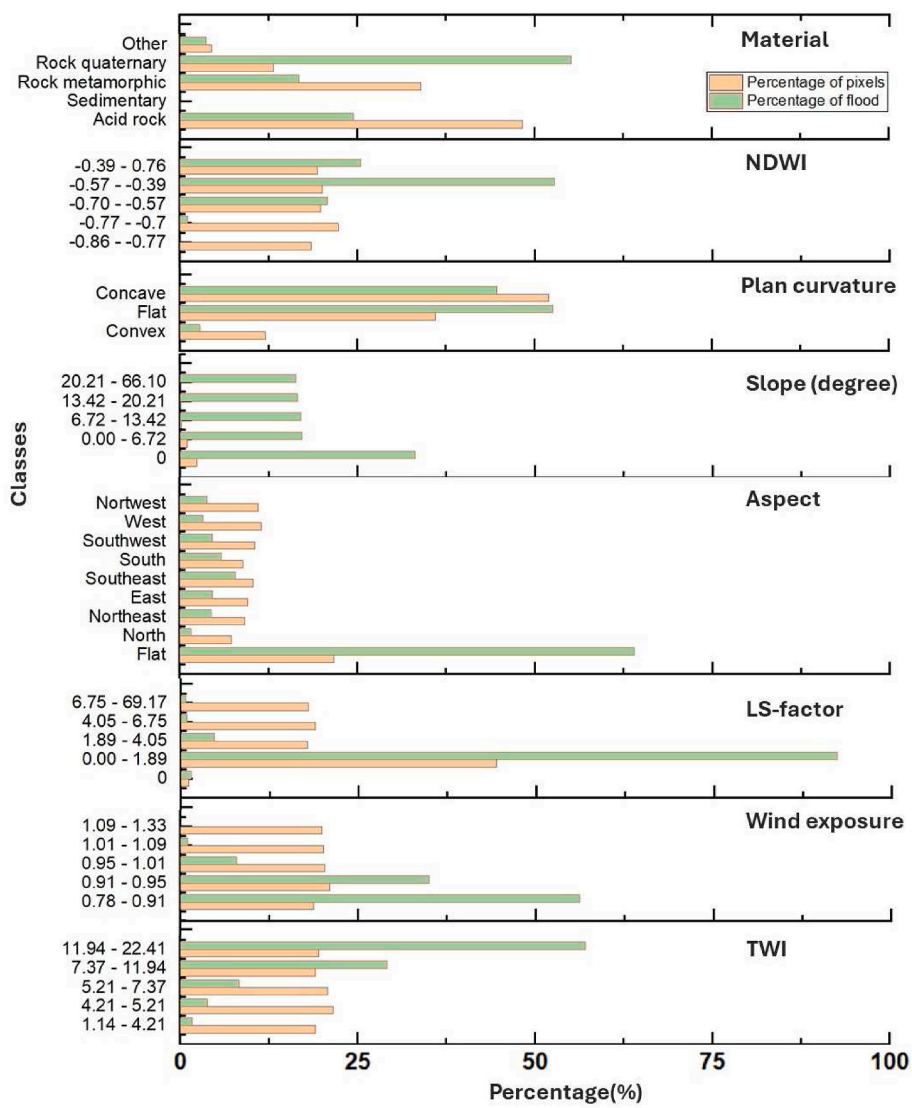

Fig. 3. Percentage of domain and flood occurrence in each subclasses for each factor.

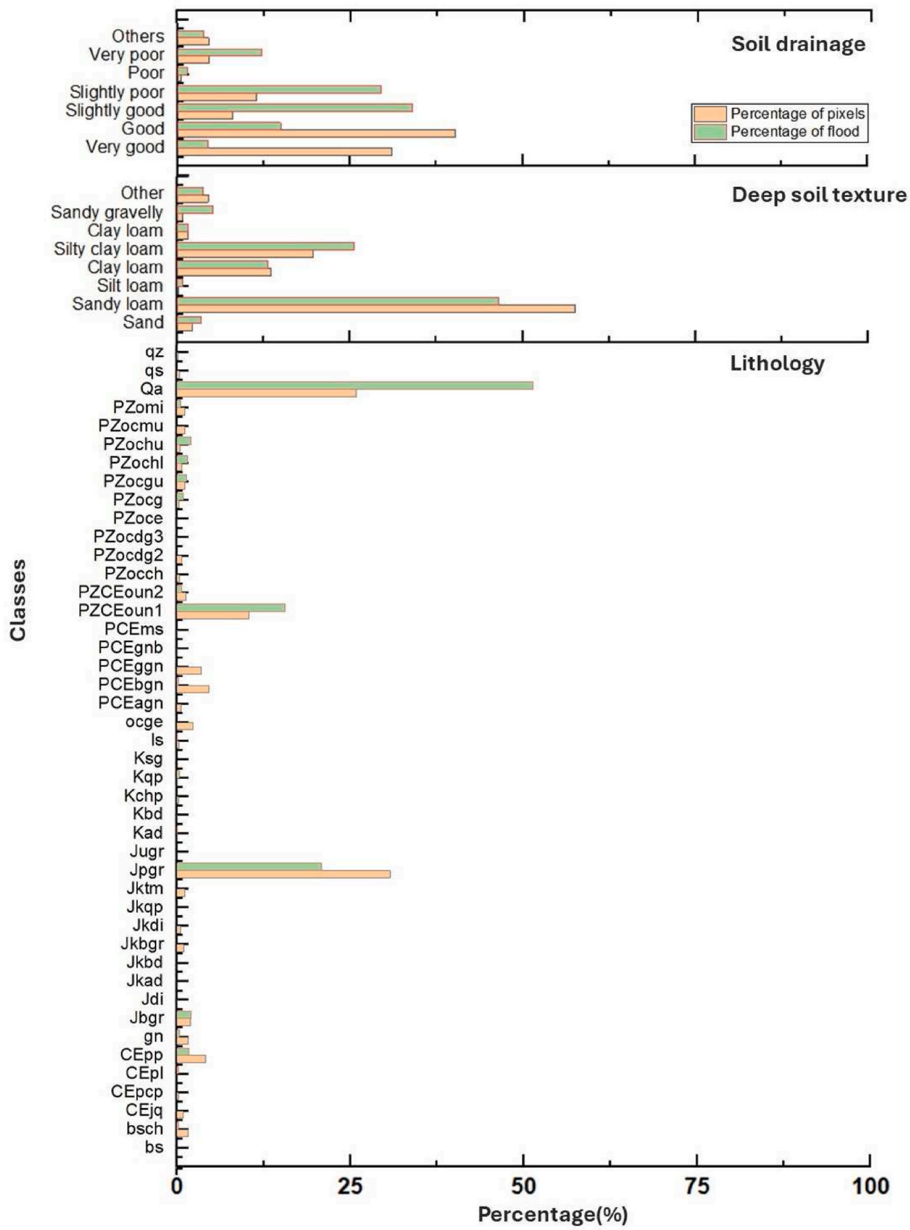

Fig. 3. (*continued*).

slopes and surface water runoff, thereby considerably increasing the risk of floods. Meanwhile, aspect is attributed to hydrodynamics, influencing flow patterns and flood (Yariyan et al., 2020). This study reveals a notable susceptibility in areas categorized as flat-class and low slope. For the slope factor, class 0 had the greatest influence on flood occurrence, with an FR value of 2.356. Specifically, flooding occurred on flat terrain, with FR values for aspect and plan curvature of 2.952 and 1.459, respectively. The LS factor exhibited a notable influence, particularly in class 0–1.89, where it attained an FR value of 2.082. A previous study unveiled the correlation between wind-driven surface currents, and circulation patterns within water bodies (Sakai et al., 2009). Meanwhile, the TWI emerges as a factor in hydrological analyses (Sørensen et al., 2006). Similarly, wind exposure strongly influenced the 0.78–0.91 class, with an FR value of 2.995. TWI played a crucial role in flood occurrence, particularly in class 11.94–22.41, with an FR value of 2.925.

NDWI can effectively differentiate between water and non-water features, informing wet areas (high NDWI) and contributing to the comprehension of hydrological processes, including water availability, distribution, and movement within a landscape (Teng et al., 2021). The hydrological factors influencing flood occurrence (Fig. 3), where the impact was identified using the NDWI. NDWI values in high the range of −0.57 to −0.39 were strongly related to flood occurrence, as indicated by an FR value of 2.625.

The geological factors can influencing flood occurrence (Fig. 3). Notably, quaternary rock exerted an impact on flood occurrence, with an FR value of 4.175. Additionally, flood occurrences in the Cheongju region are most prevalent in the Qa area (Fig. 3), showcasing FR values of 1.986 (Table S1). Geological characteristics and dynamic processes during the Quaternary period could impact flood occurrence, attributable to factors such as loose sediment composition and historical glacial activity (Panin et al., 2020; Zhang et al., 2024).

Soil characteristics emerged as key factors influencing flood occurrence (Fig. 3). Slightly good soil drainage conditions exhibited the highest FR value of 4.291. The FR value for the <20 cm soil depth class was 2.488. Deep soil texture significantly impacted flood susceptibility, particularly in the silt loam class, which had an FR value of 8.994. Fine

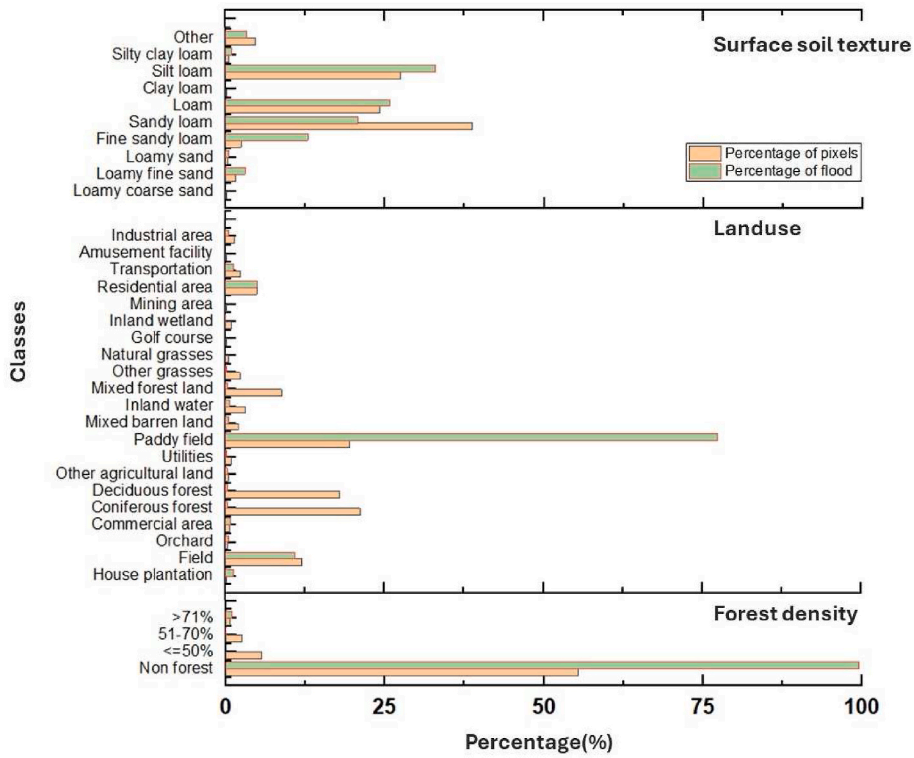

Fig. 3. (*continued*).

sandy loam surface soil had an FR value of 5.325. The impact of land use and forest density on flood occurrence (Fig. 3). Residential plantations strongly influenced flood occurrence, with an FR value of 7.897. Forest density, particularly on non-forested land, also influenced flood occurrence, with an FR value of 1.800. The findings indicate a correlation between flood occurrences and paddy fields, indicating that these areas frequently function as flood-prone reservoirs (Okazawa et al., 2011). While the forested areas demonstrate greater soil stability and water absorption capacity, thereby reducing susceptibility to flooding. In contrast, non-forested regions exhibit inefficient water absorption, resulting in heightened runoff and flooding during flood events (Shah et al., 2022).

Flood susceptibility maps generated using the LSTM, SVR, and BOOST algorithms divided flood susceptibility into five classes: very low, low, moderate, high, and very high, using the quantile method (Fig. 4). The resulting maps demonstrate the efficacy and adaptability of these algorithms in identifying regions susceptible to flooding.

Our thorough assessment of the precision of the machine learning models used in this study was based on the AUROC (Fig. 5). In the training phase, the SVR, BOOST, and LSTM models displayed AUROC values of 83.16%, 86.70%, and 87.01%, respectively. In the testing phase, the models demonstrated robust predictive capability, with AUROC values of 81.65%, 86.43%, and 86.91%, respectively. The predictive efficacy of the models was categorized within five distinct AUROC intervals: failed (0.5–0.6), poor (0.6–0.7), fair (0.7–0.8), good (0.8–0.9), and excellent (0.9–1.0) (Addis, 2023b; Arora et al., 2021a; Fressard et al., 2014).

4. Discussion

Identifying the factors that impact flood events is a crucial initial phase in flood susceptibility mapping. It is essential to utilize correlation tests to assess the multicollinearity among influential factors, as their relationships significantly influence the outcome of the prediction model. Increasing the number of potentially influential factors included in the model increases the intricacy of the data dimensions, resulting in reduced model performance and compromised generalization ability (Surbakti et al., 2020). Therefore, an optimal combination of variables must be determined by eliminating correlated and insignificant factors. In this study, we investigated 20 factors for flood susceptibility mapping based on the available data. The results of Pearson correlation analyses revealed a maximum acceptable linear correlation of 0.7 among the selected factors (Hong et al., 2018; Tehrany et al., 2019), and identified 15 influential factors (Table 1). To generate flood vulnerability maps, we conducted Pearson correlation analysis between flood events and these 15 influencing factors. A weak linear correlation was found among the factors (Fig. S1), indicating a high degree of independence among the variables, contributing to the robustness of our flood susceptibility assessment.

This study used information gain ratio (IGR) approaches to identify factors with significant prediction potential and evaluate linear correlation and multicollinearity among these factors (Chan et al., 2022). The findings of the IGR analysis indicated that forest density, soil depth, LS factor, land use type, and slope had a greater capacity to predict the spatial distribution of flood susceptibility (Fig. 6). Previous studies have highlighted the potential role of forests in mitigating flood risks, attributing their impact to the hydrological cycle through reduced rainfall interception, evapotranspiration, and infiltration, resulting in increased runoff (Shuster et al., 2005). Previous studies evaluating flood susceptibility have also emphasized the essential role of soil properties, including parameters such as soil depth, in controlling water absorption and runoff (Hümann et al., 2011). The results of the current study demonstrate the close relationship between topography and flood susceptibility by considering the potential risk of soil erosion. The LS factor, which indicates vulnerability to soil erosion, is a key consideration in the assessment of flood occurrence (Popa et al., 2019). Our flood susceptibility maps predicted flood occurrence in regions characterized by paddy fields, business zones, orchards, and residential plantations. The integration of land use data into flood susceptibility models provides insights into the hydrological consequences and potential flood hazards

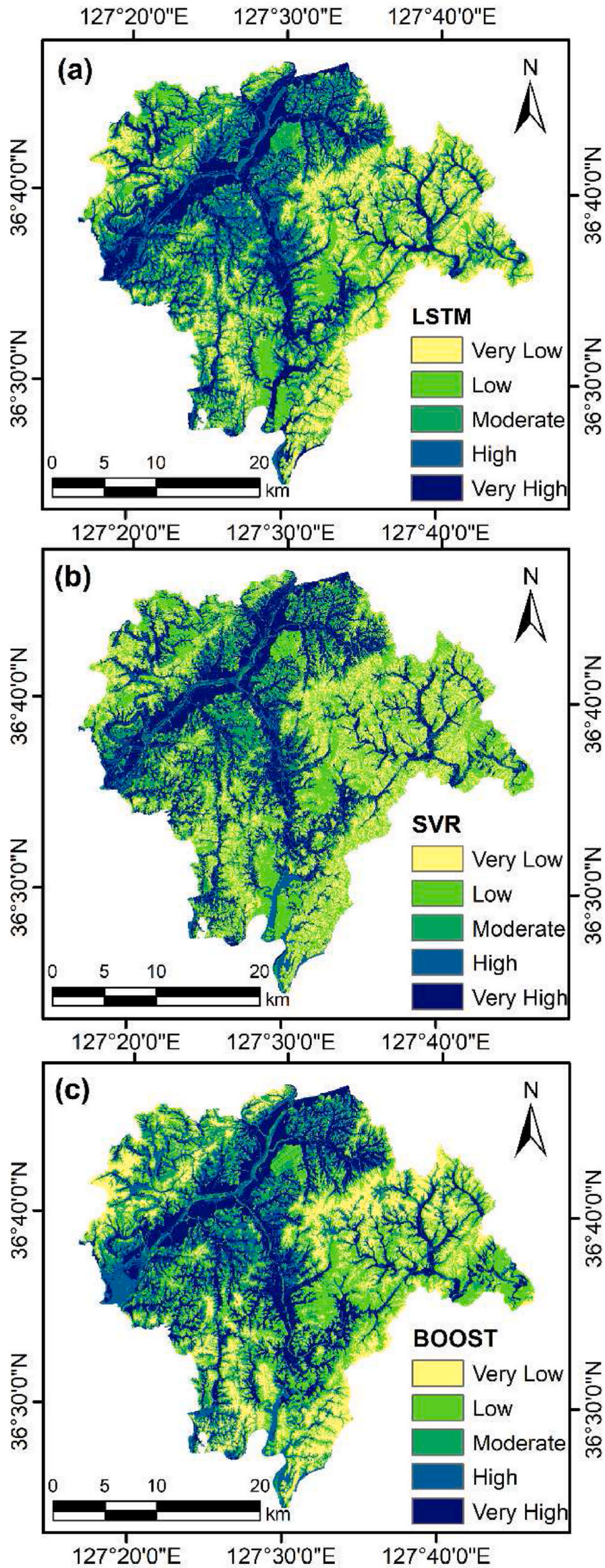

Fig. 4. Flood susceptibility maps created using the (a) long short-term memory (LSTM), (b) support vector regression (SVR), and (c) boosted tree (BOOST) algorithms.

linked to urbanization (Alexakis et al., 2014). The flood susceptibility maps revealed higher flood occurrence risk in areas characterized by low slopes. In a previous study, the consideration of slope as a factor in flood susceptibility revealed that lower slopes tended to favor flood occurrence (Bajabaa et al., 2014).

In this study, we used frequency analysis to examine variables affecting flood susceptibility and their interrelationships, in keeping with prior flood susceptibility analyses (Arora et al., 2021b; Rahmati et al., 2016). Flood susceptibility maps were produced based on SVR, BOOST, and LSTM modeling and showed similar patterns, with AUROC values of 83.16%, 86.70%, and 87.01%, respectively. The accuracy of the three hybrid models introduced in this study indicate that they may be used as references for future research, focusing primarily on evaluating flash flood susceptibility in specific regions. However, some studies have reported higher AUROC values for hybrid models in flood susceptibility mapping. For example, a flood susceptibility model based on an SVR-based genetic algorithm had an AUROC of 75% (Dodangeh et al., 2020). BOOST is an ensemble learning method that can be regarded as a subset within the broader category of decision tree algorithms. A BOOST model supported by 12 influential factors applied by Lee et al. (2017) showed good capability for flood susceptibility mapping of Seoul, Korea, with an AUROC of 77.26%. In another study, flood susceptibility in Shangyou County, China was predicted using an LSTM algorithm with an optimized local spatial sequence algorithm, obtaining an AUROC value of 93.75%, indicating excellent prediction performance (Fang et al., 2021). Optimization algorithms can enhance the performance of machine learning models (Arora et al., 2021a; Fang et al., 2021; Pham et al., 2021). Investigations of machine learning model performance have revealed that their predictive capacity is influenced by the ratio of training and testing data (Nguyen et al., 2021; Nurwatik et al., 2022).

This study demonstrates the utilization of machine learning algorithms for flood susceptibility mapping, highlighting the drawbacks of current approaches. The limitation is a need for additional methods to optimize hyperparameters. The results of the current study contribute to efforts to identify the factors influencing flood susceptibility. Nevertheless, this study had other limitation regarding rainfall distribution data, which was unavailable for our study region. Additionally, there is a lack of high-quality current meteorological maps for the area. These absences may have affected the accuracy of the flood susceptibility model results. In future studies, it is recommended to explore optimization algorithms to identify the flood susceptibility models. Recent developments in machine learning optimization approaches provide potential to improve predicted accuracy and optimize the model (Jaafari et al., 2022; Kaveh et al., 2022). Furthermore, the combination of other geographical data, such as meteorological data and human activity, might improve flood susceptibility maps. Previous studies have highlighted the significance of rainfall in predicting flood occurrence (Khosravi et al., 2018; Priscillia et al., 2021; Tehrany et al., 2015).

5. Conclusion

This study thoroughly examined the mechanisms underlying flooding occurrence in Cheongju County, South Korea using field survey-derived flood occurrence data through in-depth analysis of the correlations between flood occurrence and 15 factors: slope, aspect, LS factor, wind exposition index, TWI, plan curvature, NDWI, material, geology, soil drainage, soil depth, deep soil texture, surface soil texture, land use type, and forest density. The SVR, BOOST, and LSTM algorithms effectively created flood susceptibility maps, among which LSTM produced the most accurate map, as shown by the high AUC values during both training and testing phases. In the training phase, SVR, BOOST, and LSTM achieved AUC values of 83.16%, 86.70%, and 87.01%, respectively. In the testing phase, these models demonstrated robust predictive capabilities, with AUC values of 81.65%, 86.43%, and 86.91%, respectively. The results of this study provide a broad perspective of the factors

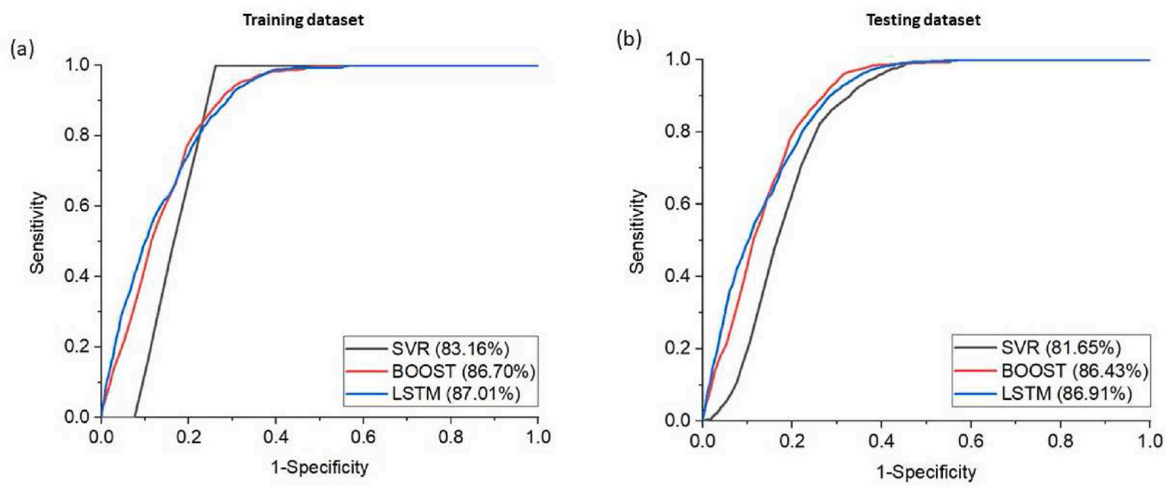

Fig. 5. Comparison of the predictive performances of the models using (a) training and (b) testing datasets.

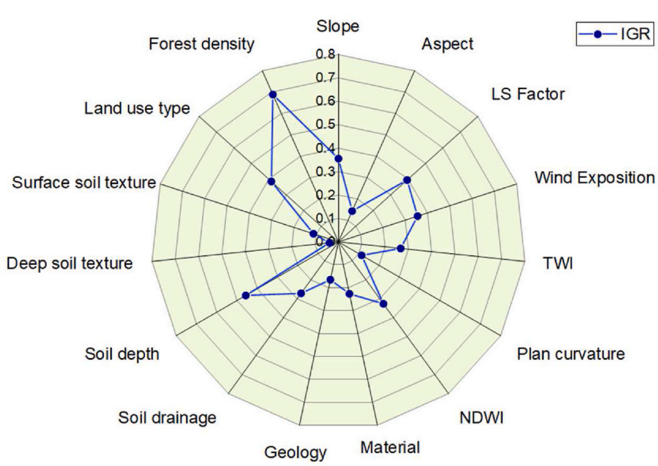

Fig. 6. Performances of influential factors based on information gain ratio values.

influencing flood susceptibility and highlight the importance of using advanced techniques to assess flood risks, while offering helpful insights into the better implementation of appropriate flood risk mitigation measures, allowing decision-makers to create informed strategies for faster, more proactive disaster responses, leading to improved resiliency.

CRediT authorship contribution statement

Liadira Kusuma Widya: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology. **Fatemeh Rezaie:** Writing – review & editing, Visualization, Formal analysis, Data curation. **Woojin Lee:** Writing – review & editing, Validation, Methodology. **Chang-Wook Lee:** Writing – review & editing, Investigation. **Nurwatik Nurwatik:** Writing – review & editing, Validation, Methodology. **Saro Lee:** Supervision, Project administration, Funding acquisition, Data curation.

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.

Acknowledgments

This research was supported by the Basic Research Project of the Korea Institute of Geoscience and Mineral Resources, South Korea (KIGAM, South Korea) and the National Research Foundation of Korea, South Korea (NRF, South Korea) grant funded by Korea government, South Korea (MSIT, South Korea) (No. 2023R1A2C1003095).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2024.121291.

The English in this document has been checked by at least two professional editors, both native speakers of English. For a certificate, please see: http://www.textcheck.com/certificate/9MmK7M.

References

Addis, A., 2023a. GIS-based flood susceptibility mapping using frequency ratio and information value models in upper Abay river basin, Ethiopia. Natural Hazards Research 3, 247–256. https://doi.org/10.1016/j.nhres.2023.02.003.

Addis, A., 2023b. GIS-based landslide susceptibility mapping using frequency ratio and shannon entropy models in Dejen District, Northwestern Ethiopia. J. Eng. 2023, 1–14. https://doi.org/10.1155/2023/1062388.

Albano, R., Sole, A., 2018. Geospatial methods and tools for natural risk management and communications. ISPRS Int. J. Geo-Inf. 7, 470. https://doi.org/10.3390/ijgi7120470.

Alexakis, D.D., Grillakis, M.G., Koutroulis, A.G., Agapiou, A., Themistocleous, K., Tsanis, I.K., Michaelides, S., Pashiardis, S., Demetriou, C., Aristeidou, K., Retalis, A., Tymvios, F., Hadjimitsis, D.G., 2014. GIS and remote sensing techniques for the assessment of land use change impact on flood hydrology: the case study of Yialias basin in Cyprus. Nat. Hazards Earth Syst. Sci. 14, 413–426. https://doi.org/10.5194/nhess-14-413-2014.

Allaire, M., 2018. Socio-economic impacts of flooding: a review of the empirical literature. Water Secur 3, 18–26. https://doi.org/10.1016/j.wasec.2018.09.002.

Arabameri, A., Asadi Nalivan, O., Chandra Pal, S., Chakrabortty, R., Saha, A., Lee, S., Pradhan, B., Tien Bui, D., 2020. Novel machine learning approaches for modelling the gully erosion susceptibility. Rem. Sens. 12, 2833. https://doi.org/10.3390/rs12172833.

Arora, A., Arabameri, A., Pandey, M., Siddiqui, M.A., Shukla, U.K., Bui, D.T., Mishra, V. N., Bhardwaj, A., 2021a. Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India. Sci. Total Environ. 750, 141565 https://doi.org/10.1016/j.scitotenv.2020.141565.

Arora, A., Pandey, M., Siddiqui, M.A., Hong, H., Mishra, V.N., 2021b. Spatial flood susceptibility prediction in Middle Ganga Plain: comparison of frequency ratio and Shannon's entropy models. Geocarto Int. 36, 2085–2116. https://doi.org/10.1080/10106049.2019.1687594.

Arshad, A., Zhang, Z., Zhang, W., Dilawar, A., 2020. Mapping favorable groundwater potential recharge zones using a GIS-based analytical hierarchical process and probability frequency ratio model: a case study from an agro-urban region of Pakistan. Geosci. Front. 11, 1805–1819. https://doi.org/10.1016/j.gsf.2019.12.013.
Bajabaa, S., Masoud, M., Al-Amri, N., 2014. Flash flood hazard mapping based on quantitative hydrology, geomorphology and GIS techniques (case study of Wadi Al Lith, Saudi Arabia). Arabian J. Geosci. 7, 2469–2481. https://doi.org/10.1007/s12517-013-0941-2.
Balogun, A.-L., Rezaie, F., Pham, Q.B., Gigović, L., Drobnjak, S., Aina, Y.A., Panahi, M., Yekeen, S.T., Lee, S., 2021. Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression (SVR) with GWO, BAT and COA algorithms. Geosci. Front. 12, 101104 https://doi.org/10.1016/j.gsf.2020.10.009.
Chan, J.Y.L., Leow, S.M.H., Bea, K.T., Cheng, W.K., Phoong, S.W., Hong, Z.W., Chen, Y. L., 2022. Mitigating the multicollinearity problem and its machine learning approach: a review. Mathematics 10, 1283. https://doi.org/10.3390/math10081283.
Chen, W., Lei, X., Chakrabortty, R., Pal, S.C., Sahana, M., Janizadeh, S., 2021. Evaluation of different boosting ensemble machine learning models and novel deep learning and boosting framework for head-cut gully erosion susceptibility. J. Environ. Manag. 284, 112015 https://doi.org/10.1016/j.jenvman.2021.112015.
Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., Wang, X., Bian, H., Zhang, S., Pradhan, B., Ahmad, B.B., 2020. Modeling flood susceptibility using data-driven approaches of naïve Bayes tree, alternating decision tree, and random forest methods. Sci. Total Environ. 701, 134979 https://doi.org/10.1016/j.scitotenv.2019.134979.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297. https://doi.org/10.1007/BF00994018.
Ding, Q., Chen, W., Hong, H., 2016. Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. Geocarto Int. 1–21. https://doi.org/10.1080/10106049.2016.1165294.
Dodangeh, E., Panahi, M., Rezaie, F., Lee, S., Bui, D.T., 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. (Amst.) 590, 125423. https://doi.org/10.1016/j.jhydrol.2020.125423.
Du, P., Bai, X., Tan, K., Xue, Z., Samat, A., Xia, J., Li, E., Su, H., Liu, W., 2020. Advances of four machine learning methods for spatial data handling: a review. Journal of Geovisualization and Spatial Analysis 4, 13. https://doi.org/10.1007/s41651-020-00048-5.
Fang, Z., Wang, Y., Peng, L., Hong, H., 2021. Predicting flood susceptibility using LSTM neural networks. J. Hydrol. (Amst.) 594, 125734. https://doi.org/10.1016/j.jhydrol.2020.125734.
Ferreira, C.S.S., Walsh, R.P.D., Steenhuis, T.S., Shakesby, R.A., Nunes, J.P.N., Coelho, C. O.A., Ferreira, A.J.D., 2015. Spatiotemporal variability of hydrologic soil properties and the implications for overland flow and land management in a peri-urban Mediterranean catchment. J. Hydrol. (Amst.) 525, 249–263. https://doi.org/10.1016/j.jhydrol.2015.03.039.
Fressard, M., Thiery, Y., Maquaire, O., 2014. Which data for quantitative landslide susceptibility mapping at operational scale? Case study of the Pays d'Auge plateau hillslopes (Normandy, France). Nat. Hazards Earth Syst. Sci. 14, 569–588. https://doi.org/10.5194/nhess-14-569-2014.
Graves, A., 2012. Long short-term memory. In: Supervised Sequence Labelling with Recurrent Neural Networks. Springer, Berlin, Heidelberg, pp. 31–38. https://doi.org/10.1007/978-3-642-24797-2_4.
Graymore, M.L.M., Wallis, A.M., Richards, A.J., 2009. An index of regional sustainability: a GIS-based multiple criteria analysis decision support system for progressing sustainability. Ecol. Complex. 6, 453–462. https://doi.org/10.1016/j.ecocom.2009.08.006.
Hamedi, H., Alesheikh, A.A., Panahi, M., Lee, S., 2022. Landslide susceptibility mapping using deep learning models in Ardabil province, Iran. Stoch. Environ. Res. Risk Assess. 36, 4287–4310. https://doi.org/10.1007/s00477-022-02263-6.
Hong, H., Pradhan, B., Sameen, M.I., Kalantar, B., Zhu, A., Chen, W., 2018. Improving the accuracy of landslide susceptibility model using a novel region-partitioning approach. Landslides 15, 753–772. https://doi.org/10.1007/s10346-017-0906-8.
Hussain, M.A., Shuai, Z., Moawwez, M.A., Umar, T., Iqbal, M.R., Kamran, M., Muneer, M., 2023. A review of spatial variations of multiple natural hazards and risk management strategies in Pakistan. Water (Basel) 15, 407. https://doi.org/10.3390/w15030407.
Hümann, M., Schüler, G., Müller, C., Schneider, R., Johst, M., Caspari, T., 2011. Identification of runoff processes – the impact of different forest types and soil properties on runoff formation and floods. J. Hydrol. (Amst.) 409, 637–649. https://doi.org/10.1016/j.jhydrol.2011.08.067.
Islam, A.R.M.T., Talukdar, S., Mahato, S., Kundu, S., Eibek, K.U., Pham, Q.B., Kuriqi, A., Linh, N.T.T., 2021. Flood susceptibility modelling using advanced ensemble machine learning models. Geosci. Front. 12, 101075 https://doi.org/10.1016/j.gsf.2020.09.006.
Jaafari, A., Panahi, M., Mafi-Gholami, D., Rahmati, O., Shahabi, H., Shirzadi, A., Lee, S., Bui, D.T., Pradhan, B., 2022. Swarm intelligence optimization of the group method of data handling using the cuckoo search and whale optimization algorithms to model and predict landslides. Appl. Soft Comput. 116, 108254 https://doi.org/10.1016/j.asoc.2021.108254.
Jonkman, S.N., 2005. Global perspectives on loss of human life caused by floods. Nat. Hazards 34, 151–175. https://doi.org/10.1007/s11069-004-8891-3.
Kaur, G., Singh, G., Motavalli, P.P., Nelson, K.A., Orlowski, J.M., Golden, B.R., 2020. Impacts and management strategies for crop production in waterlogged or flooded soils: a review. Agron. J. 112, 1475–1501. https://doi.org/10.1002/agj2.20093.
Kaveh, A., Rahmani, P., Eslamlou, A.D., 2022. An efficient hybrid approach based on Harris Hawks optimization and imperialist competitive algorithm for structural optimization. Eng. Comput. 38, 1555–1583. https://doi.org/10.1007/s00366-020-01258-7.
Khan, A., Gupta, S., Gupta, S.K., 2020. Multi-hazard disaster studies: monitoring, detection, recovery, and management, based on emerging technologies and optimal techniques. Int. J. Disaster Risk Reduc. 47, 101642 https://doi.org/10.1016/j.ijdrr.2020.101642.
Khosravi, K., Pham, B.T., Chapi, K., Shirzadi, A., Shahabi, H., Revhaug, I., Prakash, I., Bui, D.T., 2018. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Sci. Total Environ. 627, 744–755. https://doi.org/10.1016/j.scitotenv.2018.01.266.
Kim, J.C., Lee, S., Jung, H.S., Lee, S., 2018. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. Geocarto Int. 33, 1000–1015. https://doi.org/10.1080/10106049.2017.1323964.
Kong, W., Dong, Z.Y., Jia, Y., Hill, D.J., Xu, Y., Zhang, Y., 2019. Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Trans. Smart Grid 10, 841–851. https://doi.org/10.1109/TSG.2017.2753802.
Lee, S., Kim, J.C., Jung, H.S., Lee, M.J., Lee, S., 2017. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics, Nat. Hazards Risk 8, 1185–1203. https://doi.org/10.1080/19475705.2017.1308971Lee.
Lee, S., Sambath, T., 2006. Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. Environ. Geol. (Berl.) 50, 847–855. https://doi.org/10.1007/s00254-006-0256-7.
Lee, S., Talib, J.A., 2005. Probabilistic landslide susceptibility and factor effect analysis. Environ. Geol. (Berl.) 47, 982–990. https://doi.org/10.1007/s00254-005-1228-z.
Lin, X., Li, Z., Chen, W., Sun, X., Gao, D., 2023. Forest fire prediction based on long- and short-term time-series network. Forests 14, 778. https://doi.org/10.3390/f14040778.
Machiwal, D., Cloutier, V., Güler, C., Kazakis, N., 2018. A review of GIS-integrated statistical techniques for groundwater quality evaluation and protection. Environ. Earth Sci. 77, 681. https://doi.org/10.1007/s12665-018-7872-x.
Mondal, B.K., Sahoo, S., 2022. Application of geospatial techniques for urban flood management: a review. https://doi.org/10.1007/978-3-030-94544-2_13.
Natekar, S., Patil, S., Nair, A., Roychowdhury, S., 2021. Forest fire prediction using LSTM. In: 2021 2nd International Conference for Emerging Technology (INCET). IEEE, pp. 1–5. https://doi.org/10.1109/INCET51464.2021.9456113.
Nguyen, Q.H., Ly, H.B., Ho, L.S., Al-Ansari, N., Le, H.V., Tran, V.Q., Prakash, I., Pham, B. T., 2021. Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. Math. Probl Eng. https://doi.org/10.1155/2021/4832864.
Niehoff, D., Fritsch, U., Bronstert, A., 2002. Land-use impacts on storm-runoff generation: scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany. J. Hydrol. (Amst.) 267, 80–93. https://doi.org/10.1016/S0022-1694(02)00142-7.
Nurwatik, N., Ummah, M.H., Cahyono, A.B., Darminto, M.R., Hong, J.H., 2022. A comparison study of landslide susceptibility spatial modeling using machine learning. ISPRS Int. J. Geo-Inf. 11, 602. https://doi.org/10.3390/ijgi11120602.
Oh, H.J., Lee, S., 2017. Shallow landslide susceptibility modeling using the data mining models artificial neural network and boosted tree. Appl. Sci. 7, 1000. https://doi.org/10.3390/app7101000.
Okazawa, Y., Yeh, P.J.-F., Kanae, S., Oki, T., 2011. Development of a global flood risk index based on natural and socio-economic factors. Hydrol. Sci. J. 56, 789–804. https://doi.org/10.1080/02626667.2011.583249.
Panahi, M., Sadhasivam, N., Pourghasemi, H.R., Rezaie, F., Lee, S., 2020. Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR). J. Hydrol. (Amst.) 588, 125033. https://doi.org/10.1016/j.jhydrol.2020.125033.
Panin, A.V., Astakhov, V.I., Lotsari, E., Komatsu, G., Lang, J., Winsemann, J., 2020. Middle and Late Quaternary glacial lake-outburst floods, drainage diversions and reorganization of fluvial systems in northwestern Eurasia. Earth Sci. Rev. 201, 103069 https://doi.org/10.1016/j.earscirev.2019.103069.
Park, S., Jeon, S., Choi, C., 2012. Mapping urban growth probability in South Korea: comparison of frequency ratio, analytic hierarchy process, and logistic regression models and use of the environmental conservation value assessment. Landsc. Ecol. Eng. 8, 17–31. https://doi.org/10.1007/s11355-010-0137-9.
Pham, B.T., Luu, C., Phong, T.V., Trinh, P.T., Shirzadi, A., Renoud, S., Asadi, S., Le, H.V., Meding, J.V., Clague, J.J., 2021. Can deep learning algorithms outperform benchmark machine learning algorithms in flood susceptibility modeling? J. Hydrol. (Amst.) 592, 125615. https://doi.org/10.1016/j.jhydrol.2020.125615.
Popa, M.C., Peptenatu, D., Drăghici, C.C., Diaconu, D.C., 2019. Flood hazard mapping using the flood and flash-flood potential index in the Buzău River Catchment, Romania. Water (Basel) 11, 2116. https://doi.org/10.3390/w11102116.
Pour, S.H., Wahab, A.K.A., Shahid, S., Asaduzzaman, M., Dewan, A., 2020. Low impact development techniques to mitigate the impacts of climate-change-induced urban floods: current trends, issues and challenges. Sustain. Cities Soc. 62, 102373 https://doi.org/10.1016/j.scs.2020.102373.
Pradhan, B., Lee, S., Dikshit, A., Kim, H., 2023. Spatial flood susceptibility mapping using an explainable artificial intelligence (XAI) model. Geosci. Front. 14, 101625 https://doi.org/10.1016/j.gsf.2023.101625.
Priscillia, S., Schillaci, C., Lipani, A., 2021. Flood susceptibility assessment using artificial neural networks in Indonesia. Artificial Intelligence in Geosciences 2, 215–222. https://doi.org/10.1016/j.aiig.2022.03.002.

Rahmati, O., Pourghasemi, H.R., Zeinivand, H., 2016. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. Geocarto Int. 31, 42–70. https://doi.org/10.1080/10106049.2015.1041559.

Rana, I.A., Asim, M., Aslam, A.B., Jamshed, A., 2021. Disaster management cycle and its application for flood risk reduction in urban areas of Pakistan. Urban Clim. 38, 100893 https://doi.org/10.1016/j.uclim.2021.100893.

Rezaie, F., Panahi, M., Bateni, S.M., Jun, C., Neale, C.M.U., Lee, S., 2022. Novel hybrid models by coupling support vector regression (SVR) with meta-heuristic algorithms (WOA and GWO) for flood susceptibility mapping. Nat. Hazards 114, 1247–1283. https://doi.org/10.1007/s11069-022-05424-6.

Ruiz, A., Lopez-de-Teruel, P.E., 2001. Nonlinear kernel-based statistical pattern analysis. IEEE Trans. Neural Network. 12, 16–32. https://doi.org/10.1109/72.896793.

Saha, A., Pal, S., Arabameri, A., Blaschke, T., Panahi, S., Chowdhuri, I., Chakrabortty, R., Costache, R., Arora, A., 2021. Flood susceptibility assessment using novel ensemble of hyperpipes and support vector regression algorithms. Water (Basel) 13, 241. https://doi.org/10.3390/w13020241.

Sakai, A., Nishimura, K., Kadota, T., Takeuchi, N., 2009. Onset of calving at supraglacial lakes on debris-covered glaciers of the Nepal Himalaya. J. Glaciol. 55, 909–917. https://doi.org/10.3189/002214309790152555.

Shah, N.W., Baillie, B.R., Bishop, K., Ferraz, S., Högbom, L., Nettles, J., 2022. The effects of forest management on water quality. For. Ecol. Manage. 522, 120397 https://doi.org/10.1016/j.foreco.2022.120397.

Sharma, C.S., Mishra, A., Panda, S.N., 2014. Assessing impact of flood on river dynamics and susceptible regions: geomorphometric analysis. Water Resour. Manag. 28, 2615–2638. https://doi.org/10.1007/s11269-014-0630-2.

Shuster, W.D., Bonta, J., Thurston, H., Warnemuende, E., Smith, D.R., 2005. Impacts of impervious surface on watershed hydrology: a review. Urban Water J. 2, 263–275. https://doi.org/10.1080/15730620500386529.

Sørensen, R., Zinko, U., Seibert, J., 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrol. Earth Syst. Sci. 10, 101–112. https://doi.org/10.5194/hess-10-101-2006.

Surbakti, F.P.S., Wang, W., Indulska, M., Sadiq, S., 2020. Factors influencing effective use of big data: a research framework. Inf. Manag. 57, 103146 https://doi.org/10.1016/j.im.2019.02.001.

Tehrany, M.S., Jones, S., Shabani, F., 2019. Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. Catena 175, 174–192. https://doi.org/10.1016/j.catena.2018.12.011.

Tehrany, M.S., Pradhan, B., Jebur, M.N., 2015. Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. Stoch. Environ. Res. Risk Assess. 29, 1149–1165. https://doi.org/10.1007/s00477-015-1021-9.

Teng, J., Xia, S., Liu, Y., Yu, X., Duan, H., Xiao, H., Zhao, C., 2021. Assessing habitat suitability for wintering geese by using Normalized Difference Water Index (NDWI) in a large floodplain wetland, China. Ecol. Indicat. 122, 107260 https://doi.org/10.1016/j.ecolind.2020.107260.

Wang, J.Q., Du, Y., Wang, J., 2020. LSTM based long-term energy consumption prediction with periodicity. Energy 197, 117197. https://doi.org/10.1016/j.energy.2020.117197.

Yariyan, P., Avand, M., Abbaspour, R.A., Torabi Haghighi, A., Costache, R., Ghorbanzadeh, O., Janizadeh, S., Blaschke, T., 2020. Flood susceptibility mapping using an improved analytic network process with statistical models. Geomatics, Nat. Hazards Risk 11, 2282–2314. https://doi.org/10.1080/19475705.2020.1836036.

Yin, H., Li, C., 2001. Human impact on floods and flood disasters on the Yangtze River. Geomorphology 41, 105–109. https://doi.org/10.1016/S0169-555X(01)00108-8.

Zhang, J., Khalil, Y.S., Li, T., Wang, X., Guo, W., Wang, L., Hong, J., Zhang, H., 2024. Source, distribution patterns and resources potential of selenium in soil of north Pakistan: revealed by a national-scale geochemical mapping. J. Geochem. Explor. 261, 107470 https://doi.org/10.1016/j.gexplo.2024.107470.