Journal of Flood Risk Management

WILEY

CIWEM Chartered Institution of Water and Environmental Management | Journal of Flood Risk Management

CRITICAL REVIEW OPEN ACCESS

Analysis of the Utilization of Machine Learning to Map Flood Susceptibility

Ali Pourzangbar | Peter Oberle | Andreas Kron | Mário J. Franca

Institute for Water and Environment (IWU), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Correspondence: Ali Pourzangbar (ali.pourzangbar@kit.edu)

Received: 16 May 2024 | **Revised:** 21 February 2025 | **Accepted:** 16 March 2025

Funding: The authors received no specific funding for this work.

Keywords: artificial intelligence | deep learning | flood management | flood susceptibility mapping | machine learning

ABSTRACT

This article provides an analysis of the utilization of Machine Learning (ML) models in Flood Susceptibility Mapping (FSM), based on selected publications from the past decade (2013–2023). Recognizing the challenge that some stages of ML modeling inherently rely on experience or trial-and-error approaches, this work aims at establishing a clear roadmap for the deployment of ML-based FSM frameworks. The critical aspects of ML-based FSM are identified, including data considerations, the model's development procedure, and employed algorithms. A comparative analysis of different ML models, alongside their practical applications, is made. Findings suggest that despite existing limitations, ML methods, when carefully designed and implemented, can be successfully utilized to determine areas at risk of flooding. We show that the effectiveness of ML-based FSM models is significantly influenced by data preprocessing, feature engineering, and the development of the model using the most impactful parameters, as well as the selection of the appropriate model type and configuration. Additionally, we introduce a structured roadmap for ML-based FSM, identification of overlooked conditioning factors, comparative model analysis, and integration of practical considerations, all aimed at enhancing modeling quality and effectiveness. This comprehensive analysis thereby serves as a critical resource for professionals in the field of FSM.

1 | Introduction

Annually, floods affect approximately 200 million people (Ritchie and Rosado 2024) and are responsible for over half of the damage from natural disasters in the last 50 years (Bates 2021). In Germany, the 2021 extreme flood caused at least 180 deaths and EUR 46 billion of estimated loss (Mohr et al. 2022). The scale, frequency, and impact of flood events are not only increasing globally but also becoming complex due to the interplay of climate change and uncontrolled development (Ludwig et al. 2023; Tellman et al. 2021). To mitigate flood risks, addressing hazard, vulnerability, and exposure is essential. Flood Hazard Mapping (FHM) encompasses various approaches to assessing flood likelihood and severity. As a subsection of FHM, Flood Susceptibility Mapping (FSM) focuses on identifying flood-prone areas based on environmental and topographical factors. This process is supported by Flood Inventory Maps (FIM), which document past flood events and provide historical insights for improved flood risk assessment (Razavi-Termeh et al. 2025). In this study, FSM is a focal point.

The FSM is an important approach in flood management, providing vital information for planning and response (Bentivoglio et al. 2022). Simulation models are the basis of FSM and they can be categorized into four groups: physical, physics-based, empirical, and hybrid models (Mudashiru et al. 2021). Physical models use actual laboratory experiments to simulate flooding (Tomiczek et al. 2020), but they can be costly, are limited

in scale, require specialized equipment, and solid know-how. Physics-based models, such as numerical solvers, use mathematical equations describing principles of fluid motion to simulate the flood dynamics; these require advanced expertise, tedious parameterization, and a comprehensive knowledge of flood-related parameters, making them complex to set up and utilize (Brunner 2016). Empirical models rely on statistics from past floods and other influential parameters such as rainfall, topography, and land use. They can be classified into three types: statistical models, which are used for quantitative data analysis; Machine Learning (ML) models, employed for data-driven predictions; and Multi Criteria Decision Analysis (MCDA) models, which provide experts a key role in assessing criteria and weights (Wang et al. 2019). Finally, Hybrid models are a combination of two or more of the former three.

Rooted in observed data, statistical models are popular for their simplicity and usefulness in data-scarce regions, but they struggle with predicting extreme events and lacked the ability to capture underlying physical processes (Mudashiru et al. 2021). To compensate for the above limitations, a leapfrog is possible via the integration of ML-based models, grounded in the realm of Artificial Intelligence (AI), into FSM. Given flood modeling, an AI-driven model could integrate data-driven (leveraging ML) and knowledge-based (e.g., hydrological models) insights to create a comprehensive model of flood susceptibility and risks for a particular area. ML-based models, however, belong to a subset of AI focused on developing algorithms that can learn from and make predictions or decisions based on data (e.g., historical flood data, real-time data from sensors, and satellite imagery, to name but a few).

ML-based methods have grown in popularity due to their ability to provide accurate and efficient prediction models for both long-term and short-term floods (Mosavi et al. 2018). They also offer simplicity, speed, and reasonable accuracy; however, they are prone to overtraining, limited generalization, and occasionally produce unreliable results. These shortcomings arise because ML-based models are heavily dependent on the quality of input data and contributing parameters, as well as they cannot fully capture the fundamental physics of flood dynamics. To overcome such limitations, the researchers have increasingly focused on integrating ML-based models with other modeling strategies like metaheuristic techniques, optimization algorithms, numerical analyses, and physical models (Razavi-Termeh et al. 2018). These hybrid approaches, categorized as AI-driven hybrid models, have significantly boosted the effectiveness, physical interpretability, soundness, and accuracy of ML-based models.

Many studies have applied ML models in FSM; however, a lack of consensus on critical aspects such as model setup, data processing, and model validation limits their reproducibility and transferability. This paper fills this gap by systematically analyzing the existing literature and presenting a detailed roadmap for developing ML-based FSM models. Our study addresses these gaps by systematically reviewing the current literature and providing a comprehensive framework that emphasizes the often-overlooked stages of data preprocessing, model development, validation, and post-processing. Unlike previous works, this paper not only critically analyzes the state-of-the-art models but also offers guidelines and best practices to improve their application across various geographic and contextual scenarios. The novelty of this paper lies in its provision of a structured roadmap to the development of ML models for FSM, enhanced by a comparative analysis of model performance across various geographic and contextual scenarios. This approach allows for a deeper understanding of ML model effectiveness and improves their applicability for future research and real-world implementations. All abbreviations used in this paper are listed and defined in Appendix 1.

2 | Existing Knowledge on the Application of ML in Flood Management

Here, to identify recent advances, emerging trends, and critical gaps that necessitate further investigation, we analyze recently published literature that explores the role of ML in advancing flood management techniques. Table 1 presents the key distinctions identified in flood management, synthesizing insights from recent reviews on the integration of ML techniques. While current applications of machine learning in flood management provide invaluable insights, several methodological gaps still hinder its broader utilization. Previous studies mentioned important considerations like model explainability and generalization. However, they often lacked an in-depth exploration of the methodologies to undertake these tasks. We present detailed procedures to address these elements, enhancing the application of ML in flood management. We outline comprehensively the phases of model development, from pre-processing to post-processing, detailing suitable methods for each stage.

3 | Research Methodology

To find the most relevant articles to base our analysis, we searched Scopus, Science Direct, Web of Science, and Google Scholar databases for publications dealing with FSM using ML and statistical methodologies. We identified a total number of 521 articles published and indexed in above-mentioned databases between 2008 and 2023, using the following keywords: "machine learning" and ["flood susceptibility mapping" or "flood susceptibility" or "flood susceptibility assessment" or "flood prediction"]. Before 2008, there is no indexed article with these keywords. Figure 1a depicts the distribution of these articles across different years, illustrating both the number and percentage of articles published annually. The selected articles are published across 250 journals. To enhance clarity in representation, Figure 1b highlights only journals contributing more than 2% of the selected articles ($\geq$10 articles). However, the selection process included articles from a broader range of journals, including those contributing less than 2%.

After 2013, there was a notable increase in the number of articles using ML methods for flood management, totaling 503 over the past decade. To select the most relevant publications, we used a "snowballing" strategy (Lechowska 2018), starting with highly cited articles and then exploring those that cite or are cited by these key studies. This approach narrowed the focus to 100 innovative articles, which were analyzed in detail for their contributions to model types, input parameters, and evaluation

TABLE 1 | Assessment of the coverage of several aspects related to the application of ML in flood management in the literature.

		Investigated articles							
Category	**Type**	**Mosavi et al. (2018)**	**Mudashiru et al. (2021)**	**Bentivoglio et al. (2022)**	**Rasheed et al. (2022)**	**Kaya and Derin (2023)**	**Kumar et al. (2023)**	**Karim et al. (2023)**	**Current article**
Studied flood type	Riverine flood	★	★	✓	×	×	★	✓	✓
	Urban/flash flood	★	★	✓	×	×	★	✓	✓
	Coastal flood	—	—	★	×	×	★	×	×
Flood application	FSM	—	—	★	×	★	×	★	✓
	Flood Inundation Mapping (FIM)	—	×	★	×	×	★	★	×
	Flood Hazard Mapping (FHM)	★	★	★	×	×	★	★	×
Model configuration	Supervised learning	✓	★	✓	★	★	★	✓	✓
	DL	×	★	✓	★	★	★	✓	✓
	Hybrid models	✓	★	—	★	★	★	—	✓
Predictive features	Event-based prediction	★	×	✓	★	×	★	★	✓
	Time series prediction	★	×	✓	★	×	★	★	✓
	Probabilistic prediction	★	×	★	★	×	—	×	✓
	Real-time prediction	★	×	★	×	×	—	—	✓
Data considerations	Data collection sources	—	★	★	×	×	✓	★	✓
	Data cleaning	×	×	×	×	×	—	×	✓
	Data preprocessing	★	×	★	×	×	★	×	✓
Pre-processing phase	Influential parameters	×	×	★	×	★	★	—	✓
	Feature engineering	×	×	—	×	×	×	×	✓
	Study area selection	×	×	★	×	×	—	★	✓
Processing phase	Train/test data division	×	×	×	×	×	×	★	✓
	Model selection	★	×	×	×	★	✓	★	✓
	ML-based model architecture	×	×	—	★	×	×	★	✓

(Continues)

TABLE 1 | (Continued)

Category	Type	Investigated articles							
		Mosavi et al. (2018)	Mudashiru et al. (2021)	Bentivoglio et al. (2022)	Rasheed et al. (2022)	Kaya and Derin (2023)	Kumar et al. (2023)	Karim et al. (2023)	Current article
	Hyper parameters	×	×	×	★	×	—	×	✓
	Model validation	×	×	×	×	×	—	★	✓
Post-processing phase	Performance assessment indices	—	—	★	★	×	★	×	✓
	Sensitivity analysis	★	—	×	★	×	—	×	✓
	Uncertainty assessment	★	—	—	★	×	✓	—	✓
	Model generalization	★	×	★	×	×	×	—	✓
	Comparative performance	✓	—	★	★	×	×	—	✓
Gaps on applications	Climate change effects	—	—	×	★	×	★	×	—
	Morphological evolution	×	×	×	×	×	×	×	✓
	Debris transported during floods	×	×	×	×	×	×	×	✓
	Socio-economic vulnerability	×	—	×	★	★	—	×	×
	Social Force Model (SFM)	×	×	×	×	×	×	×	—
	Physics-informed ML models	★	★	★	★	×	★	×	—
	Explainable AI (XAI)	×	×	×	×	×	×	★	✓
	AI-generated decisions reliability	×	×	×	×	×	×	★	✓
	Domain Adaptation	×	×	×	×	×	×	★	✓

Note: Key to the symbols: ★: The article briefly explained the topic but did not cover it in detail; ✓: the article discussed the topic in detail; ×: the article did not discuss the topic; — the article mentioned the topic without providing any supportive discussion.

methods. These articles and their main characteristics are detailed in the appendices of the study (Appendices 2 and 3).

4 | Roadmap for the Construction FSM Based on ML

4.1 | General Steps

Developing the ML-based model for FSM includes pre-processing, processing, and post-processing (Table 2). The pre-processing phase refers to data preparation for the modeling process, which includes data collection, study area selection, contributing parameters determination, feature engineering, data cleaning, and feature scaling (Pourzangbar et al. 2023)—compare Section 4.2.

In the processing phase, the focus is on model development. After determination of the methodology for partitioning the dataset into training and testing subsets, the configuration, architecture, and training algorithm of the ML-based model are selected. Subsequently, hyperparameters of the ML-based models are determined using either conventional approaches or meta-heuristic algorithms. In this phase, the model is developed and ready to be implemented into testing and unseen datasets—compare Section 4.3.

In the post-processing phase, the performance of the developed model in terms of accuracy and precision, model explainability, and generalization is evaluated against testing and unseen datasets. Model explainability is assessed to elucidate the contribution of conditioning factors to the overall predictions (Aydin and Iban 2023). Furthermore, in this phase, the performances of different models are compared (comparative performance), thereby drawing conclusions about their relative superiority based on accuracy and physical interpretability. Some important steps of this phase, as highlighted in Table 2, are elaborated in Section 4.4.

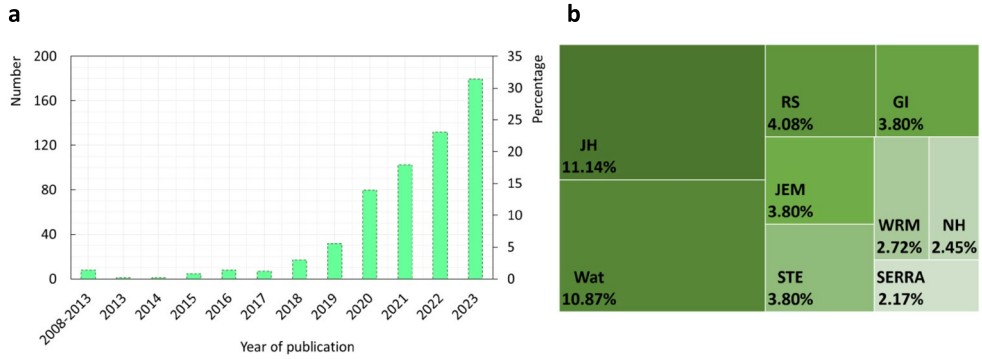

FIGURE 1 | (a) Published articles on the application of ML-based models to FSM in different years, (b) journals with over 2% of published articles on FSM using ML-based models during 2013—2023. GI = Geocarto International; JEM = Journal of Environmental Management; JH = Journal of Hydrology; NH = Natural Hazards; RS = Remote Sensing; SERRA = Stochastic Environmental Research and Risk Assessment; STE = Science of the Total Environment; Wat = Water; WRM = Water Resources Management.

TABLE 2 | Phases for developing robust and physically sound ML-based FSM (the subsections in which each subject is discussed are indicated between square brackets).

Pre-processing [4.2]	Processing [4.3]	Post-processing [4.4]
Data collection [4.2.1]	Data splitting [4.3.1]	Performance assesment [4.4.1]
Study area [4.2.2]	Feature importance[4.3.2]	Model explainability [4.4.2]
Contributing parameters [4.2.3]	Model selection [4.3.3]	Model generalization [4.4.3]
Feature selection [4.2.4]	Model architecture [4.3.4]	Comparative analysis [4.4.4]
Data treatment [4.2.5]	Hyper-parameters [4.3.5]	
	Model validation [4.3.6]	

We incorporated a flowchart in Appendix 3: Figure A1, which illustrates the three phases of developing a ML-based model for FSM. This flowchart clearly outlines each phase, providing a structured overview to guide the reader through the model development process.

4.2 | Pre-Processing Phase

4.2.1 | Data Collection

Existing resources for producing data required for ML-based model development include in situ measurements collected by ground-based sensors; Remote Sensing (RS) such as satellite imagery, Google Earth Engine, and drone-generated data; reconstruction data such as generated by numerical simulations; and finally, crowdsourcing and social media mining (Maspo et al. 2020).

In situ measurements provide highly accurate, detailed data, but are geographically limited and financially costly. Given flood modeling, field observations provide high temporal resolution data for parameters like precipitation and discharge, but they have lower spatial accuracy and are prone to damage during extreme events (Ludwig et al. 2023).

RS, on the other hand, offers global coverage and is cost-effective over time (Pourzangbar et al. 2023) despite high initial costs and varying data quality that require calibration with in situ data. In a majority of the analyzed articles, conditioning factors and flood inventory maps are derived from satellite imagery, including data from Landsat 8 (Operational Land Imager, OLI) and Sentinel-1 (Synthetic Aperture Radar, SAR), among others. SAR images are widely used in flood modeling due to their all-weather, day/night capabilities (Anusha and Bharathi 2020), enabling accurate mapping of flood extents (Hitouri et al. 2024) and improving hydraulic models for better flood simulations, especially in urban areas (Scotti et al. 2020).

Flood inventory maps are either derived from the interpretation of digital satellite images or compiled from historical flood databases (Chen et al. 2019; Yu et al. 2023). Riazi et al. (2023) employed Sentinel-1 SAR data to distinguish between flooded and non-flooded areas. Similarly, Nguyen (2022) utilized SAR imagery, enhanced with the Lee filter to reduce noise, for the detection of flood-affected regions. Topographical data, including Digital Elevation Models (DEMs), slope angle, aspect, and plain curvature, were extracted using satellite imagery. The Shuttle Radar Topography Mission (SRTM) has been widely used in various studies to generate DEM layers (Razavi-Termeh et al. 2023). Land-use and land-cover data, which are also important conditioning factors for flood modeling, were obtained through the OLI onboard Landsat satellites, accessed via Google Earth Engine. This satellite-derived data provides detailed information on human activities and natural landscapes that significantly affect flood susceptibility (Ha et al. 2023; Mahdizadeh Gharakhanlou and Perez 2022).

Furthermore, physics-based models supply spatial and temporal data for training ML models, which are then used for predicting events like water depth and urban floods rapidly. Social media platforms also emerge as vital real-time data sources for flood management, aiding in immediate response actions, although their integration into ML-based FSM models is still in early stages but shows promise for enhancing real-time predictive capabilities. Table 3 summarizes data sources and collection methods, illustrative parameters associated with these sources, and an assessment of the advantages and disadvantages of each method.

4.2.2 | Study Area

Study areas are conditioned by geographical, climatic, human, geopolitical, historical, environmental, and geological factors. Study areas in FSM research cover diverse climate zones from arid regions like Saudi Arabia and Egypt to tropical climates such as Vietnam and Malaysia, and temperate zones including

TABLE 3 | Summary of data collection sources, corresponding tools used, and remarks about their utilization in FSM.

Type of data source	Data collection method	Example parameters	Remarks
In situ (observational data)	Sensor/gauge Field surveys	Precipitation/rainfall Flow properties Lithological maps Historical flood location	High accuracy Detailed data over long periods Limited spatial coverage Vulnerable to extreme events
Remote sensing	Satellite imagery Google Earth Engine Drone Aerial photos	Flood inventory maps Topographical data DEM NDVI	Vast spatial coverage Require in situ data for calibration Expensive and time-consuming Susceptible to weather conditions
Physics-based model	Physics-based model Numerical simulation	Flood maps Flow properties	Accurate and robust Expensive and time consuming
Crowdsourcing	Social media	Flood maps Flow properties Evacuees' behavior	Real-time data for flood maps Difficult to analyze Often of poor quality Temporal and spatial resolution not controlled

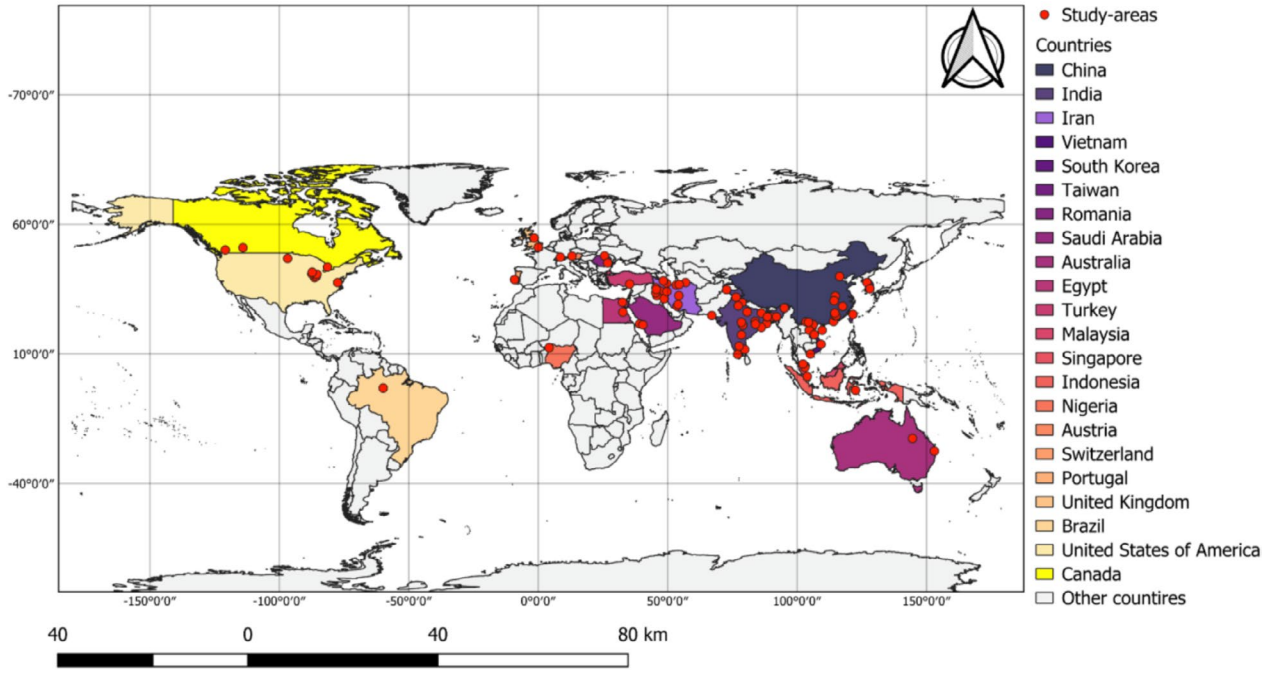

FIGURE 2 | The study areas (marked with red spots) and their associated countries in the analyzed articles.

southern Canada and Switzerland (Figure 2). The spatial scales for flood modeling vary from local (covering small areas like towns), regional (encompassing up to 100,000 km^2 such as provinces), national (involving whole countries), to supra-national (dealing with entire continents or the globe).

Evaluation of the selected study areas has shown that certain intrinsic characteristics of these locations influence the quality of the flood analysis:

- *Geography and topography*: landscapes with varied elevations and landforms significantly condition flood-prone areas, as demonstrated by the diverse terrains of Maneh and Samalqan City (Eslaminezhad et al. 2022), and the Haraz watershed (Chapi et al. 2017).
- *Historical occurrence of floods or natural disasters*: historical flood data is essential for calibrating and validating flood susceptibility models, especially in areas known for severe flooding impacts (Al-Areeq et al. 2022).
- *Climatic patterns and seasonal variation*: areas that experience diverse climate conditions, especially those with seasonal heavy rainfall or extreme weather events, are optimal for analyzing how climate influences flood risk (Luu et al. 2021; Ullah et al. 2022) due to their broad range of data on climatic factors that condition flooding.
- *Urbanization and land use*: population density and land cover are identified as critical factors in choosing study areas for flood susceptibility analysis. Locations experiencing changes in land use and population offer valuable insights into the relationship between human activity and the likelihood of flooding (Aldiansyah and Wardani 2023).
- *Data availability*: adequate and reliable data availability is a crucial factor; areas with extensive data collection allow for ML-based models which are more generalizable and reliable (Mohr et al. 2022). The brief duration of some events, for example, flash floods, makes RS tools like satellites not adequate. Social media mining may be an alternative in these cases (Costa et al. 2023).

4.2.3 | Contributing Parameters

The input parameters for the development of ML-based models can be classified as follows: topographical, hydrological, environmental, and morphological. Each of these is prepared in the form of raster maps with different spatial resolutions, usually the 30 m × 30 m pixel size (Vafakhah et al. 2020). The relationship between different parameters and flood susceptibility, along with the value ranges of these parameters, is detailed in Table 4. The positive and negative correlations, ↑ and ↓ respectively, show that the parameter of interest may have a positive or negative correlation with flood susceptibility depending on its value.

Topographical features play a critical role in the land's capacity to retain water. Steeper slopes can increase runoff velocity, heightening the risk for flash floods (El-Magd et al. 2022). The Digital Terrain Model (DTM) represents land elevations and shows that higher elevations are less flood-prone due to their natural distance from major water bodies, while lower areas are more vulnerable due to water accumulation (Deroliya et al. 2022). TRI and TPI measure the land's unevenness and relative elevation, respectively, influencing how water disperses and potentially reducing flood risk in elevated or rugged terrains (Chapi et al. 2017; Kalantari et al. 2017). Conversely, land curvature affects water flow, with convex areas (positive curvature) often seeing reduced flooding due to water divergence (Ha et al. 2023).

Hydrological parameters assess water movement and soil absorption to predict flood-prone areas during heavy rainfall.

TABLE 4 | Overview of key parameters analyzed in FSM, their typical ranges, and their correlation with flood susceptibility.

Category	Parameters	Unit	Range/attributes of parameter	Correlation with FSM
Topographical	Slope	Degree	0–80	↑
	Elevation	Meter (m)	−22 to 4000	↓
	Terrain Ruggedness Index (TRI)	m	0–871	↑↓
	Topographic Position Index (TPI)	m	−150 to 160	↓
	Plan curvature	Inverse meter (m^{-1})	−142 to 57	↑↓
	Profile curvature	m^{-1}	−50 to 50 (concave, flat, convex)	↑↓
Hydrological	Topographic Wetness Index (TWI)	—	0.50–27	↑
	Rainfall	Millimeter (mm)	180–1025	↑
	River(stream) Density (Riv-Den)	Kilometer per square kilometer (km/km^2)	0–2.574	↑
	Drainage Density (Drain-Den)	km/km^2	0.0017–1.29	↓
	Flow Accumulation (FloAcc)	—	0–1.73e+07	↑
	Flow Direction (FlowDir)	Degree	East-West-North-South-Northeast-Northwest- Southwest-Southeast	↑↓
Environmental	Land Use Land Cover (LULC)	—	Construction-Transportation-Residential areas- Mountainous-Agriculture-Grasslands-Water bodies-Forests-Woodlands-Barren land	↑↓
	Distance to River (DisRiv)	m	0–10,000	↓
	Lithology	—	More than 26 geological units such as Cretaceous, Jurassic, Permian, Quaternary, Cambrian marine, Millennium rocks, to name but a few	↑↓
	Soil type	—	Different soil types are listed, such as Alluvial, Gleisol, Kambisol, Litosol, Mediterranean, Podsolik, to name but a few	↑↓
	NDVI	—	−1—1	↓
	Distance to Road (DisRoa)	m	0–12,806	↑↓
	Distance to Fault (DisFau)	m	0–14,000	↑
Morphological	Stream Power Index (SPI)	Square meter per meter times degree [(m^2/m)degree]	6–26	↑
	Aspect	Degree	−1 (flat) and 0–360	↑↓
	Sediment Transport Index (STI)	(m^2/m)degree	0–22,540	↑

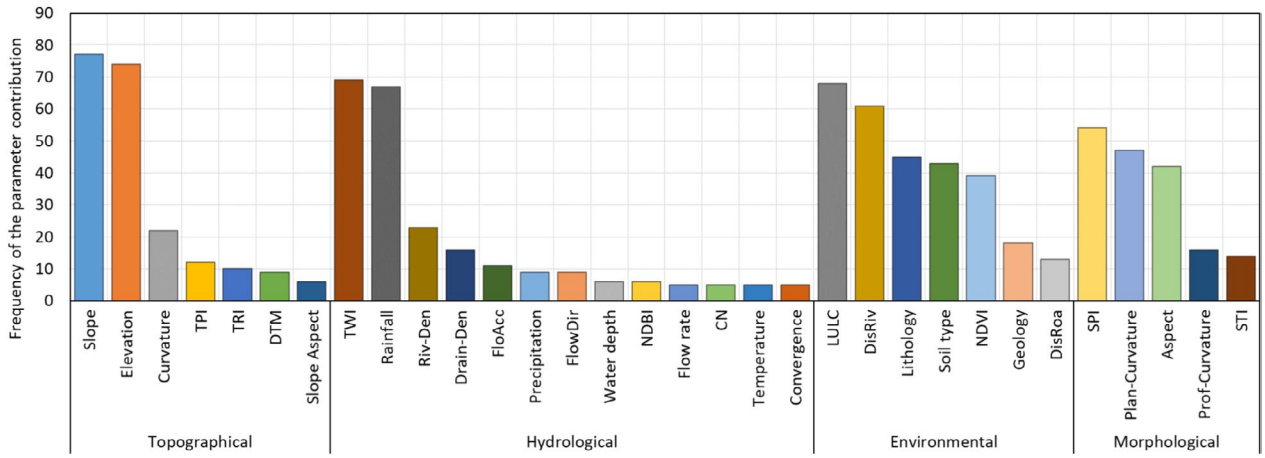

FIGURE 3 | Frequency of the input parameters used in the analyzed articles for FSM (only parameters are reported here whose contribution in the analyzed articles is more than 5%).

Higher TWI values indicate potential water accumulation areas, highlighting flood risks. Similarly, excessive rainfall can surpass the capacity of drainage systems, causing rapid runoff and floods (Park and Lee 2020). River and drainage densities indicate how rivers and their networks can affect water flow and flood likelihood. Higher river density can lead to a more complex network of water paths, potentially increasing flood risk (Horton 1932). Increased drainage density facilitates quicker runoff, possibly heightening flood risk (Sajedi-Hosseini et al. 2018). Flow accumulation (Al-Juaidi et al. 2018; Meliho et al. 2021) and direction (Nguyen 2022; Pham et al. 2021) help determine potential flood paths by tracking how water moves across a landscape.

Environmental parameters influencing floods encompass both natural elements and human activities. LULC notably impacts flood risks by modifying surface runoff and infiltration (Brody et al. 2014), where urbanization increases impermeability, thus enhancing runoff and increasing flood risk (Eslaminezhad et al. 2022). Proximity to rivers increases the likelihood of riverine flooding, particularly with rivers prone to overflow (Luu et al. 2021). Soil types (Gabriels et al. 2020), lithology, and the presence of vegetation (measured by NDVI) (Aldiansyah and Wardani 2023) are significant as they affect water absorption and retention, influencing how floods develop. Additionally, roads can alter flood dynamics by acting as barriers that either block or redirect water flow, thereby affecting flood susceptibility in surrounding areas (Nguyen et al. 2023). Distance to faults is conversely proportional to flooding susceptibility if their activity increases and starts moving (Eslaminezhad et al. 2022).

Morphological parameters, such as SPI and STI, are used to assess the erosion potential and sediment transport in watersheds. SPI correlates with slope angle and watershed area, indicating flood erosion power (Ullah et al. 2022). The aspect factor, which denotes the distribution of different topographical directions, influences slope stability by affecting water flow, solar exposure, and evaporation, which together impact broader hydrological and meteorological processes, potentially leading to landslides and floods (Ullah and Zhang 2020). STI helps estimate erosion and sedimentation rates, highlighting the dynamics of sediment movement influenced by terrain.

Several input parameters have consistently been recognized as crucial for effective FSM. Key conditioning factors in FSM include the topographic wetness index, land slope, land use and land cover, rainfall levels, and proximity to rivers. Notably, DisRiv and TWI are significant factors, being central in 61% and 68% of studies, respectively. These findings highlighted the prevailing importance of these variables in the field of FSM. Figure 3 illustrates the frequency of various parameters commonly employed in the development of ML-based models for flood susceptibility. Only parameters which are used more than five times in the analyzed articles are reported here (i.e., they are used at least in five different articles for the modeling procedure).

Diverse techniques have been employed to assess the relative significance of input parameters in ML-based models—compare Section 4.3.2. The relative importance of the 13 most important parameters is illustrated in Figure 4. A comparison of Figures 3 and 4 shows an alignment between the most commonly incorporated input parameters in ML models and those deemed most critical.

4.2.4 | Feature Selection

Including irrelevant parameters can lead to complex models that are more difficult to interpret and implement than those created using only the most essential parameters (Pourzangbar 2012); the identification of the most relevant parameters for the model development is called feature selection. Multicollinearity testing, which evaluates if two or more conditioning factors are highly correlated with each other, is one of the most common feature selection methods and can play a role. Testing multicollinearity allows the exclusion of redundant flood predictors within the modeling framework. for this, several criteria are commonly employed, such as Pearson's Correlation Coefficient (CC), Variance Inflation Factor (VIF), and condition number. Resolving multicollinearity issues typically involves either removing one of the correlated variables, combining them, or applying methods such as ridge regression or Principal Component Analysis (PCA). It is important to note that none of these tests can definitively prove the presence of multicollinearity, but rather provide evidence

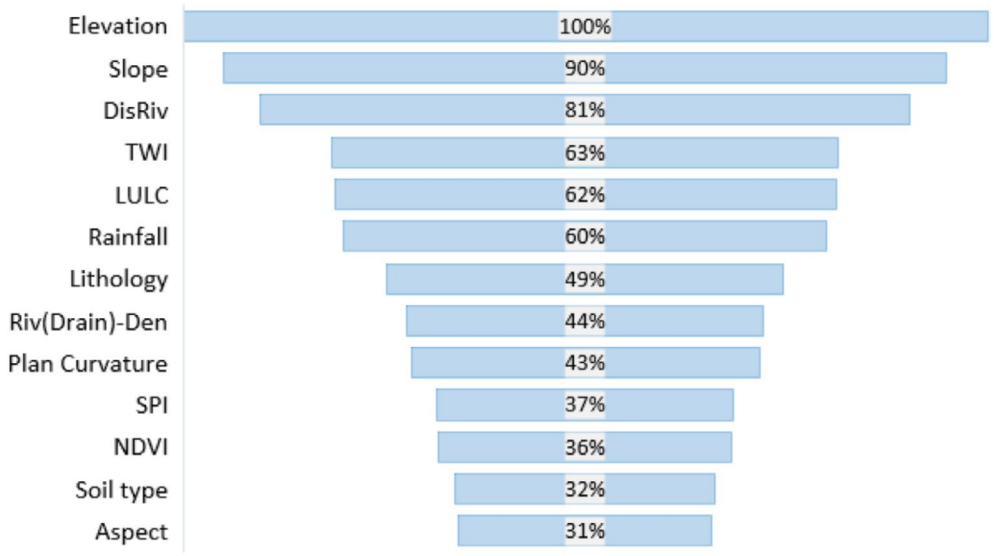

FIGURE 4 | Relative importance of parameters contributed to the development of the ML-based models in the analyzed articles.

that it may be present in a model. Therefore, it is important to use multiple tests and to interpret the results to conclude that there is multicollinearity between contributing factors.

4.2.5 | Data Treatment

Data treatment involves feature scaling techniques such as normalizing or standardizing the raw data, converting data to the appropriate format, and data cleaning (Hastie et al. 2009). Various methods, such as max-min normalization and z-score standardization, are employed to scale data types of differing magnitudes (Tabbussum and Dar 2021; Ding et al. 2020). However, standardization is preferred over min-max scaling for its robustness against outliers. Standardization is especially suitable for algorithms that assume normally distributed data, such as Logistic Regression (LR), linear regression, and Support Vector Machines (SVMs). A common approach involves normalizing the entire dataset before splitting it into training and testing groups, although this may not always yield accurate results. However, data normalization should be done after splitting the dataset into training and testing groups. Normalization parameters from the training data are applied to the test data to ensure consistent transformation. Furthermore, normalization can affect the evaluation metrics. Therefore, during evaluation, outputs are converted back to their original scale (de-normalization) to accurately assess the model's real-world effectiveness.

Data cleaning involves the process of identifying and correcting inaccuracies or inconsistencies in data, such as missing values, duplicates, outliers, or errors, to improve the quality and reliability of the dataset for analysis. Outliers, which are data points significantly deviating from the dataset's norm, can greatly influence the analysis and model performance. Therefore, it is of great importance to handle outliers properly before proceeding with analysis and model development (Khosravi et al. 2023). ML-based models developed for FSM are notably sensitive to outliers, which can influence the model's performance and accuracy. The outliers in flood data are not inherently incorrect; they may represent extreme events relevant to flood modeling. In the case of data with extreme events, one should be careful not to treat extreme flood data as outliers. Outlier detection methods are categorized into statistical methods (based on data distribution properties), distance-based methods (identifying outliers using distance thresholds), density-based methods (detecting anomalies in low-density regions), machine learning-based methods (using supervised, unsupervised, or semi-supervised learning), and ensemble methods (combining multiple techniques for better accuracy) (Pourzangbar et al. 2023). The choice depends on data characteristics and problem requirements. For more details on outlier detection methods, the reader may refer to (Pourzangbar et al. 2023). Outlier detection methods are not typically utilized directly in the preprocessing phase, though they are essential for handling anomalies in data. However, some ML-based models such as Random Forest (RF) (Razavi-Termeh et al. 2023), Boosted Regression Trees (BRT) (Youssef et al. 2022), and Gradient Boosting (GB) (Aydin and Iban 2023) are intrinsically capable of handling outliers. Table 5 summarizes the robustness degree of the ML-based algorithms and statistical indices frequently employed to treat outliers. Regarding the outlier robustness scale, the robustness of different models is categorized into four levels, represented schematically in bars. Each bar has four segments, where green portions indicate the method's capability to handle outliers. A bar with one green segment represents a model highly sensitive to outliers, whereas a fully green bar indicates strong robustness against outliers.

4.3 | Processing Phase

4.3.1 | Data Splitting

Splitting techniques are essential for optimizing model performance through training, testing, and validation. These techniques include cross-validation methods like Leave-one-out and K-fold cross-validation (Xie et al. 2021), sampling methods such as random sampling, random subsampling, and bootstrap resampling (Aldiansyah and Wardani 2023), and split techniques like fixed ratio and proportional splits with a validation set (Khan et al. 2018). These methods help evaluate the robustness of model

TABLE 5 | Robustness degree of the frequently used ML-based models and statistical indices against outliers (Green segments indicate robustness to outliers; Grey segments indicate sensitivity).

Category	Outlier robustness	Reason for robustness against outliers
ML-based models		
RF		Aggregates multiple decision trees to reduce overfitting
SVM		Maximizes the margin which reduces the impact of outliers
ANN		Robust to noise but can be sensitive to outliers
DT		Prone to overfitting, but ensemble methods improve resistance.
LR		Can be significantly affected by outliers
K-NN		Sensitive to outliers due to distance-based reliance
MLP		Generalizes well but may require outlier detection beforehand
AdaBoost		Merges weak learners into a robust, outlier-resistant model.
GB		Sequentially builds models using appropriate loss functions
XGBoost		Has regularization that helps reduce the influence of outliers
LightGBM		Handles large, outlier-prone datasets via gradient-based learning
Optimized tree-based ML		Less sensitive to outliers due to optimization techniques
REPTree		Split data into subsets which can isolate outliers
Maximum entropy		Uses strategies like pre- and post-pruning to handle outliers.
CNN		Uses pooling layers to reduce feature location sensitivity
B-LMT		Resistant via Bagging's ensemble averaging
Statistical indices		
AUROC		Balances true and false positives, minimizing outlier impact
Median		It is not affected much by extreme values
IQR		Measures variability, ignoring extreme outliers
MAD		A robust measure of variability
Accuracy		Accuracy is skewed by outliers, affecting overall correctness.
MAE		MAE is less sensitive to outliers than MSE, which squares errors.
F-Measure		Influenced by outliers affecting the true positive rate
RMSE		Sensitive to outliers because it squares the errors
Kappa Index		Less affected by outliers, depending on their distribution
Sensitivity/Specificity		Affected by outliers causing false negatives and positives.
CC		Outliers can distort CC, exaggerating relationship strength

Abbreviations: AdaBoost: Adaptive Boosting; B-LMT: Bagging-Logistic Model Tree; IQR: Interquartile Range; MAD: Median Absolute Deviation; REPTree: Reduced Error Pruning Tree; XGBoost: Extreme Gradient Boosting..

predictions against data variations, identifying potential areas for improvement in model development. Table 6 provides a summary of different data splitting techniques, along with their definition.

4.3.2 | Feature Importance

Feature importance assesses which factors are most influential in evaluating flood susceptibility. Techniques such as PCA, chi-square evaluation, FR, and Information Gain Ratio (IGR) help determine the relevance of different parameters in flood susceptibility models. IGR assesses the predictive power of factors, removing those with little or no influence (Quinlan 1996). FR measures the correlation between flood occurrences and influencing factors, with values indicating the strength of this correlation. The Jackknife test evaluates a factor's impact on model accuracy by observing changes in AUROC when the factor is removed. PCA and Functional PCA (FPCA) reduce data dimensionality to enhance interpretability (El-Haddad et al. 2021; Youssef et al. 2022), with PCA applied broadly to multivariate data and FPCA suited for data represented as functions or curves. Figure 5 illustrates nine frequently methods utilized in the analyzed articles to determine the parameters' importance. RF is the most popular method since it intrinsically determines the inputs weights and contribution to the FSM.

4.3.3 | Model Selection

A diverse array of ML-based and statistics-based models are employed for FSM. ML-based models include (1) Neural Network

TABLE 6 | Different data splitting techniques utilized in the ML-based FSM models.

Category	Method	Definition
Cross-validation techniques	Leave-one-out	The method excludes one flood event at a time for validation, with the remaining data used for training.
	K-Fold	In K-Fold, cross-validation the dataset is divided into "K" equally sized subsets and "K-1" subsets are used for training and the remaining one for validation. A "K" value of 5 is chosen.
Sampling techniques	Random sampling	The model is trained using a random sample containing 70% of the data, balanced between flood and non-flood samples. The remaining 30% is used for validation.
	Bootstrap resampling	It involves repeatedly selecting small samples from a dataset with replacement, to improve flood prediction models. Train the model on each resampled dataset to assess the variability and stability of the model's predictions.
	Random subsampling	This method, a Monte Carlo technique, involves dividing a dataset randomly into training and testing. This is repeated "B" times, with each iteration creating distinct, non-repeating samples. Unlike Bootstrap, this method ensures unique samples in each iteration, leading to varied data correlations and a different approach to data analysis and sampling.
Fixed and proportional split techniques	Fixed ratio split	The dataset is split into training (70%) and testing (30%) datasets. This technique emphasizes class balance, meaning an equal number of flooded and non-flooded points in each subset, ensuring that the model is trained and tested on a balanced representation of both classes.
	Proportional split with validation set	It divides data into 60% for training, 20% for validation, and 20% for testing. The impact of varying the training data volume is also explored, concluding that a lower fraction (50%) for training could be ideal for allowing a larger test dataset and robust statistical inference.

(NN) based models, for example, ANN, that mimic biological NNs are capable of learning from data for tasks such as classification and regression (Pourzangbar, Losada, et al. 2017; Pourzangbar, Saber, et al. 2017); (2) kernel functions utilized to map input data into a higher-dimensional space, aiding linear algorithms in solving non-linear problems such as SVMs (Pourzangbar, Brocchini, et al. 2017); (3) tree based models such as M5' model tree and Alternating Decision Tree (ADT) that employ decision tree structures to make predictions through a series of binary "if-then" decision thresholds, suitable for both regression and classification tasks (Afsarian et al. 2018; Ong et al. 2022); (4) ensemble models such as RF and Rotation Forest (RoF) that combine multiple base predictive models to improve overall accuracy and robustness of predictions (Obregon and Jung 2022); (5) hybrid models such as ANFIS that merge different types of models or model architectures to leverage strengths of individual models, enhancing performance and explainability (Kurz et al. 2022).

Statistical models rely on a variety of assumptions (e.g., data distribution) and mathematical principles (e.g., regression analysis) to analyze relationships between input and output variables. The implemented statistical models for flood prediction can be categorized into regression-based such as LR, MCDA such as FR,

Bayesian such as Naïve Bayes (NB), instance-based like K-NN, among others. Figure 6 illustrates diverse ML configurations and statistical approaches used in FSM.

The selection of ML models depends on several key factors such as data availability, model generalization, explainability, robustness to outliers, and performance, among others. Statistical models are more efficient in data-scarce regions, but models like RF and SVM are better for complex datasets. Hybrid models tend to generalize better (Section 4.4.3), while RF and SVM also show robustness against outliers, unlike linear regression and K-NN (Table 5). Model explainability is crucial, with tree-based models offering easier interpretation than NNs. Hyperparameter tuning, often through metaheuristic algorithms, plays a key role in optimizing model performance, as shown in Section 4.3.5. Performance varies depending on the dataset and study area, making it important to tailor model selection (Section 4.4.4). Lastly, RF and SVM are the most frequently used models in the literature (Table 9), reflecting their reliability for FSM.

4.3.4 | Model Architecture

Various ML models demonstrate effectiveness in FSM, showing significant diversity and adaptability across architectures. MLP models, often using sigmoidal and linear functions, are efficient with minimal data and optimal with 3 to 10 neurons in hidden layers (Xie et al. 2021). Techniques like Fuzzy Adaptive Resonance Theory (FART) and Self-Organizing Maps (SOM) incorporate complex neuron structures to enhance adaptability (Andaryani et al. 2021), using functions such as commitment and typicality in FART, and larger grids in SOM. CNN models, like the Simple CNN and LeNet-5, vary in complexity with multiple layers, including convolutional and pooling layers (Zhao et al. 2020).

SVM models using the Radial Basis Function (RBF) kernel have shown superior accuracy over other kernels in assessing flood risk. Meanwhile, RF models, augmented with algorithms like Invasive Weed Optimization (IWO) and Slime Mold Algorithm (SMA), have improved prediction accuracy by adjusting the number of trees and their features. Additionally, Gradient

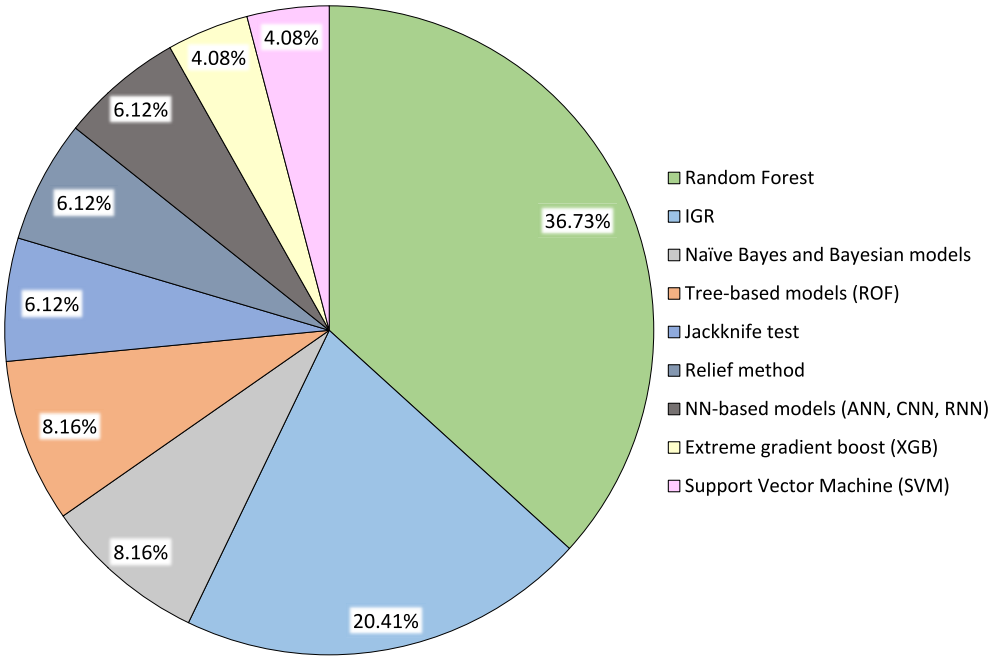

FIGURE 5 | Normalized usage frequencies of various models for assessing feature importance in the analyzed articles.

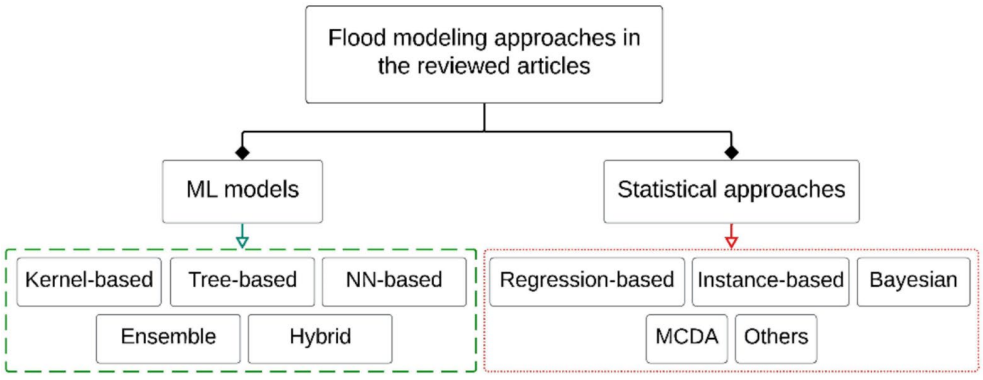

FIGURE 6 | Different ML and statistical models used for FSM.

Boosting Decision Tree (GBDT) models have been effective in boosting the performance of weaker learners through adjustments based on previous errors. These diverse ML approaches underline the potential for tailored and efficient FSM.

4.3.5 | Hyper-Parameters

Hyperparameters such as learning rate, number of layers, and number of hidden neurons are critical for optimizing the performance of ML models (Aydin and Iban 2023). Hyperparameter tuning methods can be categorized into manual and automatic approaches. Manual tuning relies on expert knowledge, trial-and-error, or default ML package values, which, while effective in certain cases, become inefficient as the number of hyperparameters increases (Le Nguyen et al. 2023). It is computationally intensive, relies on intuition, and struggles with complex optimization problems. To address these challenges, automatic search methods have gained prominence (Yang and Shami 2020). These include model-free approaches like Grid Search and Random Search, gradient-based optimization, Bayesian optimization (e.g., Gaussian Processes, Tree-Structured Parzen Estimator), and metaheuristic algorithms inspired by natural processes (Liao et al. 2024). Automatic methods offer improved efficiency and adaptability, making them essential for optimizing ML models, especially in complex applications like FSM. Below, we summarize the methodologies used in the reviewed articles for fine-tuning the hyperparameters of various ML models.

ANNs are optimized using the Levenberg–Marquardt algorithm, adjusting hidden neurons based on predictor count (Xie et al. 2021). MLP fine-tunes learning rates and layer configurations through backpropagation. SVM balances complexity and accuracy by optimizing the cost parameter (Costache et al. 2021). RF performance depends on the number of decision trees, determined via cross-validation. Bagging classifiers are optimized through trial and error, using AUROC as a performance metric (Chapi et al. 2017). Metaheuristic algorithms like GA, Differential Evolution (DE), Particle Swarm Optimization (PSO), and Grasshopper Optimization Algorithm (GOA) are integrated with ML models to improve their predictive performance by fine-tuning hyperparameters (Arora et al. 2021; Sahoo et al. 2021). Given RF models, optimization techniques such as IWO, SBO, and SMA have been utilized to find the most optimal hyperparameters (Razavi-Termeh et al. 2023). Similarly, the GridSearchCV method combined with cross-validation is extensively used across various tree-based classifiers and CNNs to identify the best hyperparameter settings for maximum accuracy (Liao et al. 2023; Lyu and Yin 2023). Based on the analyzed articles, the architecture and parameters of some main models are summarized in Table 7.

4.3.6 | Model Validation

Model validation involves three essential aspects including replicative validity, structural validity, and predictive validity (Xie et al. 2021). These aspects evaluate a model's ability to reproduce input–output relationships in different geographical regions (replicative), its accuracy in representing the underlying real-world (physical) process it is intended to simulate (structural), and its performance on unseen datasets (predictive).

To evaluate the replicative validity, techniques such as K-fold cross-validation and spatial cross-validation are utilized (Al-Areeq et al. 2022). In K-fold cross-validation, the data are split into K folds, with one fold used for validation and the remaining K-1 folds used for training (Darji et al. 2023; Witten et al. 2016). However, spatial cross-validation evaluates the model performance using data from a different spatial region to assess how well it generalizes geographically (Wang et al. 2023). Section 4.4.3 provides a detailed discussion on model generalization. Given structural validity, XAI techniques like SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) can provide insights into how the model's input features influence the model predictions (see Section 4.4.2 for detailed information). By providing local explanations, SHAP identifies key contributing factors in ML models, demonstrating high accuracy and revealing top contributing factors in flood susceptibility. Several studies highlight the value of XAI in improving the interpretation of model results and fostering trust among stakeholders in flood-related decision-making processes (Pradhan et al. 2023). Almost all the analyzed articles address predictive validity. Model accuracy and precision are evaluated using statistical metrics such as CC, RMSE, and MAE, while classification performance is measured using AUROC, accuracy, precision, recall, and the F1-score (Lyu and Yin 2023).

4.4 | Post-Processing Phase

4.4.1 | Performance Assessment

Performance evaluation of ML-based flood models employs various metrics specific to the nature of the task. Regression models utilize RMSE, CC, and MAE, which are sensitive to error magnitudes, while classification models, such as those used in flood susceptibility, favor accuracy, recall (sensitivity), precision, *F*-Score, and the Kappa index. The use of recall is crucial for highlighting false negatives, which are critical in flood modeling, but it can exaggerate model accuracy if not balanced with measures for false positives (Bentivoglio et al. 2022). The F1-score is effective for balancing recall and precision, assessing both false negatives and positives. Additionally, the Receiver Operating Characteristic (ROC) curve helps in evaluating model effectiveness across different thresholds, ensuring a robust assessment of model performance in differentiating flood and non-flood areas.

4.4.2 | Model Explainability

Model explainability, an important topic in XAI, deals with understanding how ML-based models make decisions and determines the contribution level of conditioning factors to the overall predicted results (Aydin and Iban 2023). There are two well-known XAI techniques including global and local methods. Global explanations (see Section 4.3.2) give a broad understanding of the model's behavior. They answer questions like: "on average, which factors are most important across all predictions a model makes?" Some sample methods for the global explanation are shown in Figure 7. Local explanations

TABLE 7 | Various ML-based models and their associated architecture and parameters used in the analyzed articles.

Category	Variants	Structure	Parameters
ANN	MLP, Deep Learning Neural Network (DLNN)	More than 3 layers (input-hidden-output)	Activation functions: sigmoid (frequently used in MLPs), ReLU (often in DLNNs), and softmax (typically in classification layers)
CNN	LeNet, VGG, ResNet	Convolutional layers, pooling layers, and fully connected layers	Activation functions: ReLU is widely used in convolutional layers; softmax often appears in the output layer for classification
RNN	Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU)	RNNs in these studies often include several LSTM or GRU layers, sometimes combined with dense layers	Activation functions: tanh and sigmoid in LSTM gates; ReLU and variants in some newer models
RF	Traditional RF, and extra trees classifier	Tree depth can vary, often unpruned to capture complex patterns; Feature sampling: random subset of features for each tree to reduce correlation among trees	Typically uses hundreds of trees; the optimal number of trees (estimators) was 200, the maximum features to consider at each split was 5, and the best criterion to measure the quality of a split was "Gini"
GBDT	Standard GBDT, and adaptations like LightGBM for larger datasets	Includes L1 and L2 regularization to prevent overfitting; Includes subsample ratio of the training instance and column sampling by tree	Tree count: sequentially added trees between 100 to 500; Tree depth often between 3 to 10 levels; Learning rate: often set between 0.01 and 0.1, balancing speed and accuracy
XGBoost	LightGBM, XGBoost with Linear Classifier, GPU-Accelerated XGBoost	Objective function: often set to "reg:linear" for regression tasks or "binary: logistic" for classification tasks; Subsample and Colsample_bytree parameters control the fraction of the dataset and features used for each tree, helping in reducing overfitting	Tree depth and count: similar ranges to GBDT; learning rate typically ranges between 0.01 and 0.3 A smaller learning rate requires more trees but can lead to a more robust model; a gamma value of 0.5; the regularization parameter alpha was set at 0.01, and it used 1 feature (column sample) per tree
SVM	RBF, linear SVM, polynomial SVM, sigmoid SVM	Kernel types: Linear, polynomial, RBF, and sigmoid; Margin optimization: Central to SVM, involves maximizing the margin between decision boundary and data points; The C parameter controls the trade-off between smooth decision boundary and classifying training points correctly	RBF kernel width (γ) and the regularization parameter (C) range from 10 to 20 and γ values like 0.1, 0.3, 0.2; In Linear SVM, regularization parameter C values ranged from 10 to 15; In Sigmoid SVM C values ranged from 10 to 20, and kernel coefficients ranged from 1 to 3.

provide insight into individual predictions. They help to understand why a model made a specific decision for a single instance (at a sample-wise scale). As illustrated in Figure 7, two popular methods for locally explaining model decisions are SHAP and LIME. SHAP assigns an importance value to each feature for a particular prediction (Shapley 1953). LIME is used to explain individual predictions regardless of the ML-based model used (Ribeiro et al. 2016). It works by approximating the model locally and explaining why the model made a certain decision.

4.4.3 | Model Generalization

Despite its importance, few studies have evaluated model generalization (Riazi et al. 2023). Given generalization, semi-supervised models often outperform supervised ones in flood susceptibility due to less dependence on data quality (Yu et al. 2023). Furthermore, hybrid models demonstrate better generalization compared to standalone models.

Given standalone models, ANNs show good generalization but struggle with long-term predictions (Jain and Indurthy 2003). MLP models are more efficient and generalize better than other ANN types (Senthil Kumar et al. 2005). RBF networks have strong generalization with fewer nodes (Rong et al. 2020), while CNNs benefit from batch normalization and data augmentation to enhance generalization (Zhao et al. 2020). ANFIS models excel in long-term predictions and generalize better than nonlinear regression and ANNs (Shu and Ouarda 2008). SVMs are highly effective in FSM, outperforming ANNs and linear regressions, particularly with radial basis kernels (Mosavi et al. 2018). They generalize well across data dimensions and are robust against overfitting with limited flood data. GBDT and RF models tend to overfit with scarce data, making SVMs more reliable (Yu et al. 2023). Hybrid models, like K-NN with ANNs or ANN with PCA, show improved generalization (Chen et al. 2019). Wavelet Neural Networks (WNNs) outperform ANNs in generalization (Linh et al. 2021). Ensemble Prediction Systems (EPSs), for example, combining ANNs and WNNs with techniques like genetic programming and Bayesian methods, enhance speed, accuracy, and generalization beyond traditional methods.

Challenges in flood modeling generalization persist, especially in DL (Bentivoglio et al. 2022). Incorporating real-world data and pre-training strategies can improve accuracy, but challenges remain in applying these models to new domains without extensive retraining. Advances in mesh-based NNs, particularly geometric and physics-based DL, show promise. Transfer learning enhances model generalization in areas with sparse data by training in regions with sufficient data first (Zhao et al. 2021). Rong et al. (2020) mentioned that an imbalance in the number of flood and non-flood points is another factor impacting the performance and generalization of the models.

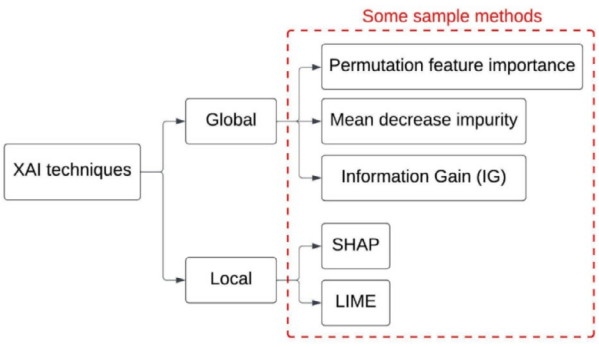

FIGURE 7 | Overview of XAI techniques by scope with examples for each method.

4.4.4 | Comparative Analysis of ML-Based Models Given FSM

In Section 4.3, different structures of ML-based models employed in the analyzed articles are discussed. The focus here is on the frequency of different ML-based and statistical models employed in different studies. Figure 8 illustrates the variation trends in various AI models over the years. It clearly shows that ensemble and hybrid models have gained significant popularity in recent years. It should be noted that the mentioned frequency of a model's usage does not correspond to the same number of separate articles, as a single article may employ a particular model type multiple times.

The deployment of these models across various tasks highlights their specific strengths and applicability. Shallow Neural Networks (Shallow-NN) and Basic Decision Trees (Basic-DTs) are favored for their simplicity and interpretability in less complex scenarios. In contrast, CNNs, LSTMs, and Deep Neural Networks (DNNs) provide the computational power and flexibility needed for handling complex, high-dimensional data. Advanced tree-based and kernel-based models demonstrate the importance of integrating sophisticated techniques to enhance predictive performance and handle more challenging data environments. Ensemble models, whether through bagging, boosting, or hybrid approaches, enhance predictive accuracy by leveraging multiple learning algorithms. Statistical models, with their foundational principles, provide robust and interpretable solutions for various predictive tasks.

The analyzed articles employ various NN-based models 62 times. Shallow-NNs such as ANN are used for classification, regression, and pattern recognition. Their architecture is simple with fewer layers, which makes them suitable for problems with straightforward data patterns. Grid-Based Neural Networks (Grid-NN) such as CNN handle grid-like data, such as images. Sequential Neural Networks (Sequential-NN) like LSTM excel in learning long-term dependencies in sequential data, making them indispensable for time series analysis and natural language processing (NLP). DNNs with multiple hidden layers are used for complex feature extraction and hierarchical data representation.

Tree-based models are employed 36 times in the analyzed articles, focusing on decision-making processes and rule-based learning. Basic-DTs and Classification and Regression Trees (CARTs) are used for classification tasks due to their simplicity and interpretability. Advanced Decision Trees (Advanced-DTs), like LMT and REPTree, enhance basic decision trees with complex rule-based algorithms, improving accuracy and robustness.

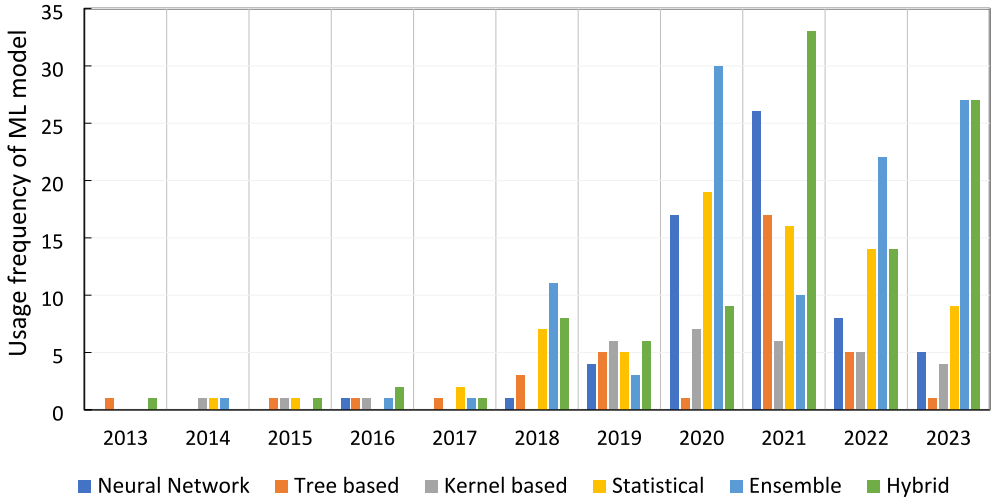

FIGURE 8 | Trends in the utilization of various ML-based models across different years in the analyzed articles.

Kernel-based models utilize different kernel functions to handle non-linear data patterns effectively. SVM-based Models (Kern-SVM) like Weakly Labeled SVM (WELLSVM) and SVM-RBF are used for classification and regression, effectively handling scenarios with partially labeled data. Kernel-based Regression (Kern-Reg) and Kernel-based Classification (Kern-Class) extend the kernel trick to regression and classification tasks, respectively, transforming input data into higher-dimensional spaces to enable linear separation of data.

Ensemble models, used 107 times in various articles, combine multiple models to improve predictive performance. Bagging-based Ensembles (e.g., RF) enhance accuracy and prevent overfitting by using predictions from different data subsets. Boosting-based Ensembles (e.g., AdaBoost, Boosted Generalized Linear Model (BGLM) and Gradient Boosting Machine (GBM)) focus on correcting errors of previous models, reducing bias and variance. Forest-based Ensembles (e.g., RF, RoF) leverage the power of forest algorithms to build a collection of decision trees for improved robustness and prediction accuracy.

Hybrid models, used 102 times in the analyzed articles, integrate various ML techniques to exploit their combined strengths. Fuzzy Logic-based Hybrids (FL-Hyb) like FART and FL-NN enhance robustness by merging fuzzy logic with NNs and evolutionary algorithms. Tree-based Hybrids (Tree-Hyb) such as Naïve Bayes Tree (NBT) combine tree methods with statistical and evolutionary techniques for better prediction reliability. NN-based Hybrids (NN-Hyb) integrate NNs with fuzzy logic, ensemble methods, or evolutionary algorithms, including models like Deep Neural Network-Aquila Optimizer (DNN-AO) and Extreme Learning Machine-Particle Swarm Optimization (ELM-PSO). Weighted Average-based Hybrids (WAve-Hyb) like RF-SVM use weighted averages of predictions from different methods. Kernel-based Hybrids (Kernel-Hyb) combine kernel-based models with statistical, ensemble, and evolutionary techniques.

Statistical models, appear 74 times, employ various principles to provide robust solutions. Regression models, such as LR and GLM, predict numerical outcomes based on linear relationships. Instance-based models like K-NN classify and predict by comparing new data to similar examples. MCDA methods, such as FR, facilitate decision-making by evaluating alternatives across multiple criteria. Bayesian methods, exemplified by NB, use Bayes' theorem for probabilistic classification, updating predictions as new evidence is obtained (Table 8).

Figures 9 and 10 and Table 9 highlight the relative performance of different ML-based models. These comparisons are based on each study's assessment of several ML-based models against a unique dataset, leading to a ranking of models based on their performance. The primary metrics used for this comparative analysis are RMSE and ROC, which are key to determining the effectiveness (accuracy and precision) of the models. Inspection of Figure 9 reveals that in terms of overall performance, the RF and CNN models stand out. Specifically, the RF model excels in 10 different studies, while the CNN model is the best in five analyzed articles. This observation underscores their superior generalization and effectiveness in the field of FSM.

In Table 9, colors differentiate each category, creating a visual hierarchy. Accordingly, RF configuration tops the table, with variations such as RF-BPSO, RF-IWO, and RF-ANN being mentioned, and it has the highest frequency of being reported as superior, with a count of 13. SVM configurations follow closely, with types like AHP-ANP-RF-SVM and FR-SVM among others, showing a frequency of 11. DNN, with variants such as Deep Neural Network-Naked Mole-Rat Algorithm (DNN-NMRA) and Deep Neural Network-Social Spider Optimization (DNN-SSO), and Tree-based models like ADT and BART share a frequency of 8.

To ensure a fair comparison, it is essential to assess the performance of ML models against the dataset used for their development. Figure 10 provides a comparative analysis of different models using different colored symbols to represent the performance of various models on distinct datasets. Each colored symbol (e.g., green circle, blue square, red triangle) corresponds to a specific dataset and the models evaluated on that dataset (details of datasets and models are available in Appendix 3). The size of the symbols reflects model performance; the larger the symbol, the better the performance.

TABLE 8 | Various ML and statistical model types employed in the analyzed articles, each accompanied by examples.

Configuration (no. of utilization)	Model type	Frequency of utilization	Example(s)
Ensemble (107)	Forest-Ens	38	RF, RoF
	Boosting-Ens	37	AdaBoost, GBM, BRT
	Others	22	EMCA, EMMean, RS-GAM, RS-MARS
	Bagging-Ens	10	Bagging Ensembles
Hybrid (102)	FL-Hyb	34	FART, FL-NN, FL-RF, FL-EA, FL-NN-EA
	Tree-Hyb	21	NBT, RF-GA
	NN-Hyb	17	DNN-AO, ELM-PSO
	WAve-Hyb	17	RF-SVM
	Kernel-Hyb	13	AdaBoost-RBF, Bagging-RBF, SVR-BA
Statistical (74)	Regression	23	LR, MARS
	Instance-based	17	KNN
	MCDA	17	FR, AHP
	Others	11	Maximum Entropy
	Bayesian	6	NB
NN-based (62)	Shallow-NN	34	ANN
	Grid-NN	13	CNN
	Sequential-NN	11	LSTM/RNN
	Deep-NN	4	DNN
Tree-based (36)	Basic-DT	20	Classification Tree, J48 DT, CART
	Advanced-DT	16	LMT, REPT, FT, CDT
Kernel-based (31)	Kern-SVM	26	SVM, SVM-RBF, K-SVM, WELLSVM
	Kern-Reg	4	SVR
	Kern-Class	1	SVC

Abbreviations: CDT: Credal Decision Tree; EMCA: Ensemble Model Committee Averaging; EMMean: Ensemble Model to estimate Mean; K-SVM: Kernel Support Vector Machine; MARS: Multivariate Adaptive Regression Splines; RS-GAM: Random Subsampling-Generalized Additive Model; RS-MARS: Random Subsampling-Multivariate Adaptive Regression Splines; SVC: Support Vector Classification.

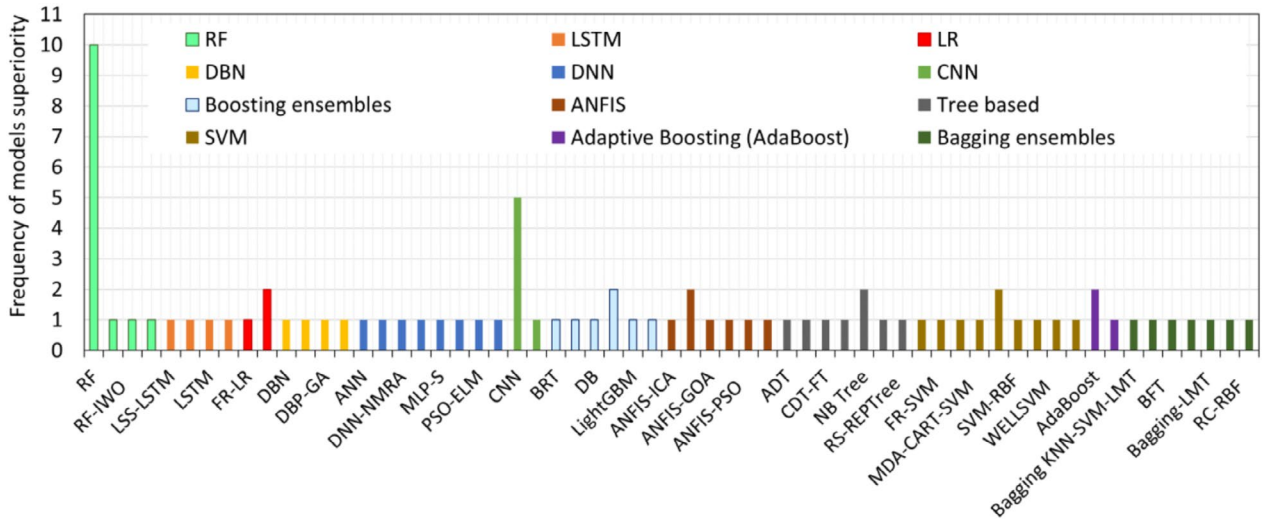

FIGURE 9 | The frequency of Models' superiority in the analyzed articles.

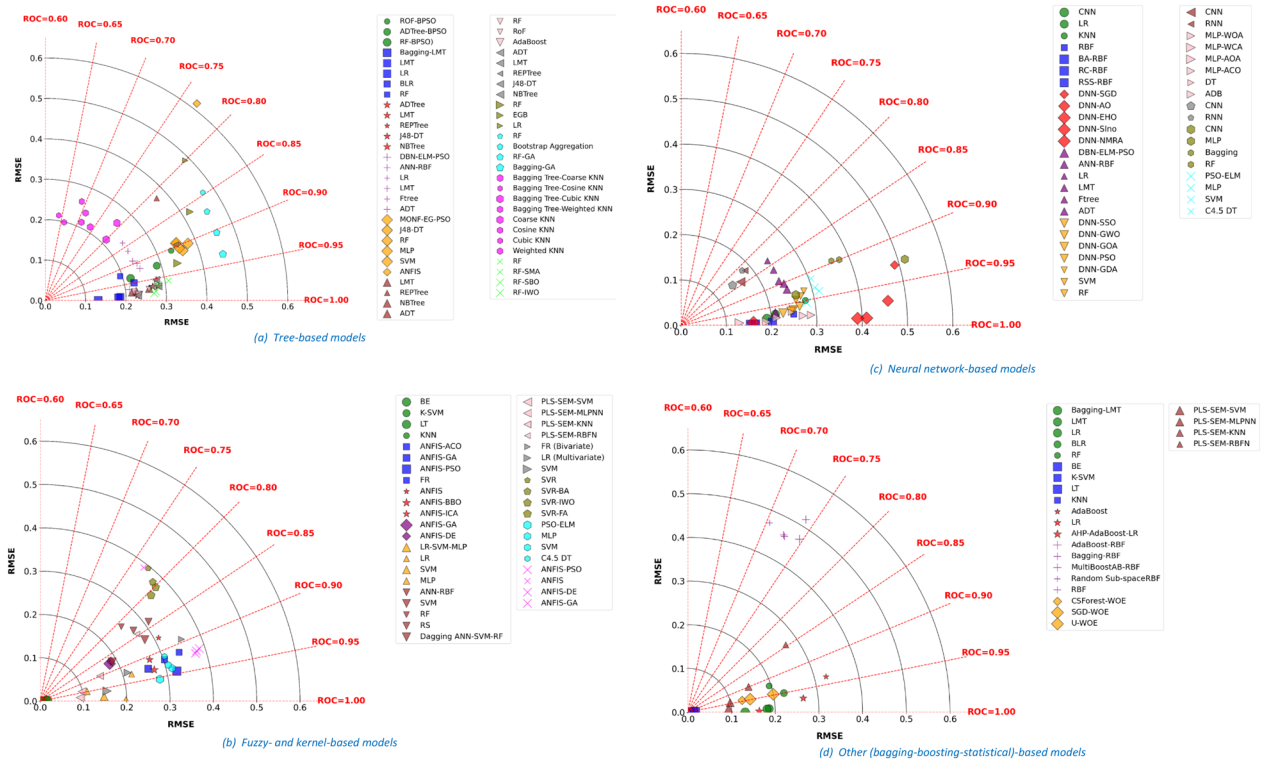

FIGURE 10 | Comparative performance of various ML-based models implemented into various datasets.

TABLE 9 | Superiority frequency of different model configurations.

Models' configuration	Model's types	Frequency
RF	RF, RF-BPSO, RF-IWO, RF-ANN	13
SVM	SVM, AHP-ANP-RF-SVM, FR-SVM, SVM-RBF, SVR-FA, WELLSVM, WoE-SVM, MDA-CART-SVM, PLS-SEM-SVM	11
DNN	DNN, DNN-NMRA, DNN-SSO, WNN, MLP-WCA, PSO-ELM	8
Tree based	ADT, BART, CDT-FT, ERT, QPSO-CDT, RSub-REPTree, NBT	8
ANFIS	ANFIS-GA, ANFIS-ICA, ANFIS-GOA, ANFIS-PSO, MONF-EG-PSO	7
Boosting ensembles	BRT, CatBoost, DB, XGBoost, LightGBM, SGB	7
CNN	CNN, SCNN	6
Bagging ensemble	Bagging KNN-SVM-LMT, Bagging Tree, BFT, Bagging-GA, Bagging-LMT, RC-RBF, Dagging ANN-SVM-RF	6
LSTM	LSS-LSTM, LSTM-ED, LSTM, STA-LSTM	4
DBN	DBN-ELM-PSO, DBP-GA	3
Adaptive Boosting	AHP-AdaBoost, AdaBoost	3
Logistic Regression	LR, FR-LR	3

Note: Shading from green to red indicates model usage frequency, from most to least frequent.
Abbreviations: BFT: Bagging Functional Tree; CDT-FT: Credal Decision Tree-Functional Tree; DB: Deep Boost; DBN: Deep Belief Network; DBP: Deep belief network with Back Propagation algorithm; ERT: Extremely Randomized Tree; ICA: Imperialistic Competitive Algorithm; LSS-LSTM: Local Spatial Sequential Long Short-Term Memory; LSTM-ED: Long Short-Term Memory based Encoder-Decoder; MDA: Multivariate Discriminant Analysis; MONF-EG-PSO: Metaheuristic Optimization and Neural Fuzzy inference-Evolutionary Genetic-Particle Swarm Optimization; PLS-SEM: Partial Least Square-Structural Equation Model; QPSO-CDT: Quantum Particle Swarm Optimization-Credal Decision Tree; RSub: Random Subsampling; SGB: Stochastic Gradient Boosting; STA-LSTM: Spatio-Temporal Attention Long-Short Term Memory; WCA: Water Cycle Algorithm.

Models closer to the bottom-left corner demonstrate superior performance, with lower RMSE and higher ROC values. Figure 10 shows that combining model types (tree-based, NN-based, Fuzzy, and Kernel-based) with optimization algorithms (PSO, WCA), bagging, or boosting significantly improves performance. Specifically, Bagging-LMT and RF

TABLE 10 | Comparative analysis of ML models based on accuracy, interpretability, computational efficiency, and robustness.

Models' configuration	Accuracy	Interpretability	Computational efficiency	Robustness
Shallow-NN and Basic-DT	Moderate	High	High	Low
Deep-NN	High	Low	Low	High
Advanced-DT	High	Moderate	Moderate	High
Kernel-based	High	Low	Low	Moderate
Ensemble models	Very high	Low to moderate	Moderate	Very high
Hybrid models	Very high	Low	Low	Very high
Statistical models	Moderate	High	High	Moderate

among tree-based models, hybrid SVM-EA and ANFIS-EA for fuzzy and kernel-based models, and hybrid DNN-EA for NN-based models demonstrate superior performance based on RMSE and ROC.

Table 10 presents a comparative analysis of ML models based on accuracy, interpretability, computational efficiency, and robustness. Shallow NNs and basic decision trees offer high interpretability and computational efficiency but have moderate accuracy, making them suitable for simple patterns, while their low robustness makes them prone to overfitting. Deep learning models (e.g., CNNs, LSTMs) provide high accuracy and robustness, excelling in complex, high-dimensional data, but require significant computational resources and lack interpretability. Advanced tree-based models balance accuracy and efficiency, performing well on structured data and resisting noise with pruning, though their interpretability depends on tree depth and complexity. Kernel-based models are highly accurate for classification and regression, capturing complex relationships, but are computationally expensive, sensitive to hyperparameters, and difficult to interpret. Ensemble and hybrid models improve predictive performance and robustness by combining multiple models, mitigating individual weaknesses. Statistical models are efficient and interpretable but may struggle with high-dimensional data, limiting their application in complex machine-learning tasks.

5 | Practical Considerations

5.1 | Data Considerations

Data scarcity remains a major challenge, particularly in regions with limited computational resources and technical expertise. This issue arises from technical constraints, especially during extreme weather events, such as malfunctioning ground-based instruments, inadequate spatial and temporal coverage, and limitations of RS tools (Mohr et al. 2022). To address these issues, crowdsourcing, data augmentation, and data fusion are suggested. Alternative data sources like social media can provide supplementary insights when traditional methods are lacking (Costa et al. 2023). Data augmentation helps address data scarcity by artificially expanding training datasets, enhancing the performance of predictive models. Techniques such as geometric transformations (rotating, flipping, altering image patches) improve model accuracy, especially when real-world data are limited (Guo et al. 2022; Madhuri et al. 2021). FSM relies on data from multiple sources, leading to discrepancies in spatial–temporal resolutions. For instance, DTMs typically have higher resolution than rainfall data. To manage these discrepancies, data fusion integrates multiple data sources, ensuring consistent spatial and temporal dimensions, which is crucial in the pre-processing phase (Islam et al. 2023).

5.2 | Model Development

5.2.1 | Physical Relevance of the Conditioning Factors

It is crucial that the conditioning factors in the ML models not only contribute to the model's accuracy but also contribute mechanistically to the system being modeled. This means that a parameter should only be excluded if it has no significant impact on the model's accuracy and reliability, or physical relevance. Removing a factor solely for the purpose of enhancing accuracy is generally not advisable, especially if that parameter is physically important to the mechanisms being modeled. This approach is in line with the principles of physics-informed ML, where the integration of data and physical laws is essential to ensure that the model is not only accurate but also physically meaningful (Carleo et al. 2019).

5.2.2 | Contributing Parameters in the Development of ML-Based Models

Inspection of Figures 3 and 4 (Section 4.2.3) reveals that some important parameters are neglected in the FSM procedure in the analyzed articles. Some of these parameters are as follows:

- *Reservoir existence*: reservoirs may(not) play multifaceted roles in flood control by reducing flood peaks and intercepting sediment (Li and Xu 2023). Hence, their existence may affect flood susceptibility and should be considered in the modeling process.
- *Prior condition in river basins*: catchment preconditions, specifically soil wetness measured by Antecedent Precipitation Index (API) values, are important for determining flood occurrence and magnitude (Mohr et al. 2022). Therefore, these preconditions should be considered in FSM.

- *Snow melting*: precipitation, including rainfall and snowfall, is the most important climate parameter for hydrological processes and flood risk assessment (Chapi et al. 2017). In glacial or nival regions like North and Central Europe, flood intensity has decreased due to reduced spring snowmelt from global warming. However, the frequency and intensity of flood events are increasing under a warming climate, with snowmelt floods being a significant contributor (Zhou et al. 2021).
- *Meteorological conditions such as temperature and humidity*: temperature anomalies due to atmospheric blocking and high solar insulation play a significant role in creating extreme precipitation and subsequent flooding (Madhuri et al. 2021). These factors are among the often-neglected parameters in FSM.
- *Infrastructures influence on flood modeling*: recent studies highlight the significant impact of infrastructure on flood risk and resilience. However, many flood models overlook changes in infrastructure, landscape occupation, and flood protection measures, which can directly affect flood risk.
- *Morphological evolution*: traditional hydrological models struggle to capture dynamic river changes caused by erosion and human activities, impacting flood behavior (Chau and Lee 1991). Infrastructure developments and climate change further exacerbate flood risks (Ludwig et al. 2023). Many models use static parameters, though flood conditions evolve over time. AI-based models like LSTM and RNN offer potential solutions but require extensive data. Given limited post-flood data, researchers may rely on social media and historical sources for qualitative insights.
- *Debris and sediment contribution to flood risk*: debris accumulation worsens floods by obstructing rivers and streets (Mohr et al. 2022). Its inclusion in flood risk analysis is crucial, as it affects flood susceptibility. Methods such as analytical models, GIS techniques, and numerical simulations have been used to study debris transport, but lack robust benchmark data (Valero et al. 2024). Integrating debris effects into flood models, with validation through social media data, is needed.

5.2.3 | Model Overfitting

Overfitting occurs when a model, while performing well on training data, shows diminished effectiveness on test and unseen data. If the training data set is too small, it may not represent all possible scenarios, leading the model to overfit to this limited data. The complex structure of a ML-based model often enhances its performance; however, this can lead to overfitting (Bentivoglio et al. 2022). To avoid overfitting, it is important to feed the ML-based models with processed data and select the most appropriate configuration for the model of interest.

5.2.4 | Trade-Off Analysis

Trade-off analysis focuses on finding an optimal balance between model complexity and factors like computational efficiency, accuracy, and data availability. While complex models offer higher accuracy, they demand more resources and time, whereas simpler models are more efficient but less accurate. This balance remains a central consideration in the evolving field of ML (Andaryani et al. 2021). Only a few of the analyzed articles included discussions on trade-off analysis. While CNNs are known for their high accuracy and efficient predictions, they are resource-heavy, demanding significant computational power for training (Zhao et al. 2020). Liao et al. (2023) proposed the utilization of clustering methods in CNNs to enhance training efficiency and overall performance. Andaryani et al. (2021) found that MLP with a sigmoidal activation offers better accuracy than a linear one, though the latter is faster. Models like FART and SOM need large datasets due to their complexity.

5.3 | Value of the Decisions Produced by AI

The integration of digitalization and AI in knowledge production and management raises concerns about the value of truth generated by these tools. Ethical issues persist, particularly regarding responsibility gaps in AI-driven decisions (Matthias 2004). This is especially critical in AI-based flood emergency management, necessitating collaboration with ethics and philosophy experts to address human-centered challenges.

5.4 | Limitations of the Current Study and Future Directions

This contribution reviews ML models for FSM, focusing on riverine and urban flooding. However, the review does not address other flooding types, such as coastal flooding, or areas like flood inundation mapping, vulnerability modeling, and risk assessments. Additionally, climate change effects, vulnerability types, force-induced models, evacuee behavior, and physics-informed machine learning models were not part of our analysis.

While this review focuses on ML models for FSM, future research should explore hybrid approaches integrating ML with hydrological models (e.g., HEC-RAS) to enhance predictive accuracy and generalization. Combining physics-based simulations with data-driven ML techniques can address challenges such as data scarcity, regional adaptability, and process interpretability. Additionally, while we analyze the performance of ML-hybrid models, further studies should compare their performance against ML-hydrological hybrids to assess robustness across diverse conditions.

6 | Summary and Conclusions

This article investigates the utilization of machine learning models in FSM by examining a 100 articles from the last decade, revealing the potential of machine learning in predicting flood-prone areas. There is no universal consensus on ML model development for FSM, leading to variations in model selection, hyperparameters, and data preprocessing. The optimal approach depends on data quality, generalizability, and interpretability. This study systematically reviews successful ML applications, highlighting

their strengths, limitations, and applicability. We compare ML models in terms of performance, interpretability, and usage trends. To address overlooked aspects, a Practical Considerations section outlines common limitations and solutions. While this study does not introduce new algorithms, it provides novel insights by evaluating existing techniques, identifying gaps, and recommending best practices for feature selection, model optimization, and a structured roadmap for the utilization of ML-based models to map flood susceptibility. These contributions serve as a foundation for future research and practical applications.

We recognized the need to bridge the gap between technical sophistication and real-world applicability. To address this, we have included a discussion in Section 5 on the practical limitations of FSM, emphasizing resource constraints and limitations in available resources, data accessibility, model generalizability, and ethical AI considerations. These factors are crucial for ensuring that ML models remain interpretable, actionable, and adaptable to diverse contexts, including resource-constrained settings.

The main conclusions drawn from this analysis are as follows:

1. Various conventional data sources, including in situ measurements, RS, numerical data, along with emerging sources like social media, may be utilized to refine FSM models, each with its benefits and limitations.
2. The quality and efficacy of ML-based FSM models are profoundly affected by the specific characteristics of the study area, such as geography, historical floods, climate, urbanization, and data availability.
3. The parameters identified as critical for FSM are topographic wetness index, slope, land use and cover, rainfall, and distance to rivers, with distance to river and topographic wetness index being particularly significant. These findings underscore the importance of these variables in flood susceptibility.
4. Techniques like IGR and random forest assess factors strongly correlated with flood susceptibility. Elevation and slope are highlighted as both highly crucial and commonly used parameters in ML-based FSM models.
5. The research reveals diverse ML and statistical models for FSM, such as NNs, SVM, tree-based, ensemble, and hybrid models, each with unique strengths for complex flood prediction tasks. The dominant trend is toward the use of ensemble and hybrid models due to their accuracy and versatility in FSM, with NNs being the most common. The integration of optimization algorithms, bagging, or boosting with other model types enhances their performance.
6. Data collection challenges during extreme weather include equipment failures and data coverage gaps. Solutions like crowdsourcing, data augmentation, and data fusion improve data quality and model training.
7. In ML models, especially for physical systems like FSM, factors should enhance accuracy and have a mechanistic role. Excluding parameters should consider both their impact on accuracy and physical relevance, aligning with physics-informed ML principles for meaningful models.
8. Traditional hydrological models fail to account for dynamic river changes due to erosion and human activities. The ML-based flood models should capture temporal morphological changes despite data collection challenges.
9. Current flood models need to integrate debris dynamics for accurate risk assessment. The gap in utilizing ML models for large debris transport studies highlights the need for incorporating debris effects into flood models, with social media data offering a novel validation approach.
10. Digitalization and AI's synergy raises challenges in knowledge value and ethics, highlighting a responsibility gap in AI-driven decisions, necessitating ethical specialist engagement.

Acknowledgements

Open Access funding enabled and organized by Projekt DEAL.

Data Availability Statement

The data that supports the findings of this study are provided in the Appendices of this article.

References

Afsarian, F., A. Saber, A. Pourzangbar, A. G. Olabi, and M. A. Khanmohammadi. 2018. "Analysis of Recycled Aggregates Effect on Energy Conservation Using M5″ Model Tree Algorithm." *Energy* 156: 264–277. https://doi.org/10.1016/j.energy.2018.05.099.

Al-Areeq, A. M., S. I. Abba, M. A. Yassin, M. Benaafi, M. Ghaleb, and I. H. Aljundi. 2022. "Computational Machine Learning Approach for Flood Susceptibility Assessment Integrated With Remote Sensing and GIS Techniques From Jeddah, Saudi Arabia." *Remote Sensing* 14: 5515. https://doi.org/10.3390/rs14215515.

Aldiansyah, S., and F. Wardani. 2023. "Evaluation of Flood Susceptibility Prediction Based on a Resampling Method Using Machine Learning." *Journal of Water and Climate Change* 14, no. 3: 937–961. https://doi.org/10.2166/wcc.2023.494.

Al-Juaidi, A. E. M., A. M. Nassar, and O. E. M. Al-Juaidi. 2018. "Evaluation of Flood Susceptibility Mapping Using Logistic Regression and GIS Conditioning Factors." *Arabian Journal of Geosciences* 11: 765. https://doi.org/10.1007/s12517-018-4095-0.

Andaryani, S., V. Nourani, A. T. Haghighi, and S. Keesstra. 2021. "Integration of Hard and Soft Supervised Machine Learning for Flood Susceptibility Mapping." *Journal of Environmental Management* 291: 112731. https://doi.org/10.1016/j.jenvman.2021.112731.

Anusha, N., and B. Bharathi. 2020. "Flood Detection and Flood Mapping Using Multi-Temporal Synthetic Aperture Radar and Optical Data." *Egyptian Journal of Remote Sensing and Space Science* 23, no. 2: 207–219. https://doi.org/10.1016/j.ejrs.2019.01.001.

Arora, A., A. Arabameri, M. Pandey, et al. 2021. "Optimization of State-Of-The-Art Fuzzy-Metaheuristic ANFIS-Based Machine Learning Models for Flood Susceptibility Prediction Mapping in the Middle Ganga Plain, India." *Science of the Total Environment* 750: 141565. https://doi.org/10.1016/j.scitotenv.2020.141565.

Aydin, H. E., and M. C. Iban. 2023. "Predicting and Analyzing Flood Susceptibility Using Boosting-Based Ensemble Machine Learning Algorithms With SHapley Additive exPlanations." *Natural Hazards* 116: 2957–2991. https://doi.org/10.1007/s11069-022-05793-y.

Bates, P. D. 2021. "Flood Inundation Prediction." *Annual Review of Fluid Mechanics* 54: 287–315. https://doi.org/10.1146/annurev-fluid-030121-113138.

Bentivoglio, R., E. Isufi, S. N. Jonkman, and R. Taormina. 2022. "Deep Learning Methods for Flood Mapping: A Review of Existing Applications and Future Research Directions." *Hydrology and Earth System Sciences* 26, no. 16: 4345–4378. https://doi.org/10.5194/hess-26-4345-2022.

Brody, S., R. Blessing, A. Sebastian, and P. Bedient. 2014. "Examining the Impact of Land Use/Land Cover Characteristics on Flood Losses." *Journal of Environmental Planning and Management* 57, no. 8: 1252–1265. https://doi.org/10.1080/09640568.2013.802228.

Brunner, G. W. 2016. "HEC-RAS, River Analysis System Hydraulic Reference Manual."

Carleo, G., I. Cirac, K. Cranmer, et al. 2019. "Machine Learning and the Physical Sciences." *Reviews of Modern Physics* 91, no. 4: 045002. https://doi.org/10.1103/RevModPhys.91.045002.

Chapi, K., V. P. Singh, A. Shirzadi, et al. 2017. "A Novel Hybrid Artificial Intelligence Approach for Flood Susceptibility Assessment." *Environmental Modelling & Software* 95: 229–245. https://doi.org/10.1016/j.envsoft.2017.06.012.

Chau, K. W., and J. H. Lee. 1991. "Mathematical Modelling of Shing Mun River Network." *Advances in Water Resources* 14, no. 3: 106–112. https://doi.org/10.1016/0309-1708(91)90001-5.

Chen, W., H. Hong, S. Li, et al. 2019. "Flood Susceptibility Modeling Using a Novel Hybrid Approach of Reduced-Error Pruning Trees With Bagging and Random Subspace Ensembles." *Journal of Hydrology* 575: 864–873. https://doi.org/10.1016/j.jhydrol.2019.05.089.

Costa, J. P., L. Rei, N. Bezak, et al. 2023. "Towards Improved Knowledge About Water-Related Extremes Based on News Media Information Captured Using Artificial Intelligence." *International Journal of Disaster Risk Reduction* 100: 104172. https://doi.org/10.1016/j.ijdrr.2023.104172.

Costache, R., A. Arabameri, I. Elkhrachy, O. Ghorbanzadeh, and Q. B. Pham. 2021. "Detection of Areas Prone to Flood Risk Using State-of-the-Art Machine Learning Models." *Geomatics, Natural Hazards and Risk* 12, no. 1: 1488–1507. https://doi.org/10.1080/19475705.2021.1920480.

Darji, K., D. Patel, V. Vakharia, et al. 2023. "Watershed Prioritization and Decision-Making Based on Weighted Sum Analysis, Feature Ranking, and Machine Learning Techniques." *Arabian Journal of Geosciences* 16, no. 1: 1–20. https://doi.org/10.1007/s12517-022-11054-w.

Deroliya, P., M. Ghosh, M. P. Mohanty, S. Ghosh, K. H. V. Durga Rao, and S. Karmakar. 2022. "A Novel Flood Risk Mapping Approach With Machine Learning Considering Geomorphic and Socio-Economic Vulnerability Dimensions." *Science of the Total Environment* 851, no. 1: 158002. https://doi.org/10.1016/j.scitotenv.2022.158002.

Ding, Y., Y. Zhu, J. Feng, P. Zhang, and Z. Cheng. 2020. "Interpretable Spatio-Temporal Attention LSTM Model for Flood Forecasting." *Neurocomputing* 403: 348–359. https://doi.org/10.1016/j.neucom.2020.04.110.

El-Haddad, B. A., A. M. Youssef, H. R. Pourghasemi, et al. 2021. "Flood Susceptibility Prediction Using Four Machine Learning Techniques and Comparison of Their Performance at Wadi Qena Basin, Egypt." *Nature Hazards* 105, no. 1: 83–114. https://doi.org/10.1007/s11069-020-04296-y.

El-Magd, S. A., A. Maged, and H. I. Farhat. 2022. "Hybrid-Based Bayesian Algorithm and Hydrologic Indices for Flash Flood Vulnerability Assessment in Coastal Regions: Machine Learning, Risk Prediction, and Environmental Impact." *Environmental Science and Pollution Research* 29: 57345–57356. https://doi.org/10.1007/s11356-022-19903-7.

Eslaminezhad, S. A., M. Eftekhari, A. Azma, R. Kiyanfar, and M. Akbari. 2022. "Assessment of Flood Susceptibility Prediction Based on Optimized Tree-Based Machine Learning Models." *Journal of Water and Climate Change* 13, no. 6: 2353–2385. https://doi.org/10.2166/wcc.2022.435.

Gabriels, K., P. Willems, and J. van Orshoven. 2020. "A Data-Driven Analysis, and Its Limitations, of the Spatial Flood Archive of Flanders, Belgium to Assess the Impact of Soil Sealing on Flood Volume and Extent." *PLoS One* 15, no. 10: e0239583. https://doi.org/10.1371/journal.pone.0239583.

Guo, Z., V. Moosavi, and J. P. Leitão. 2022. "Data-Driven Rapid Flood Prediction Mapping With Catchment Generalizability." *Journal of Hydrology* 609: 127726. https://doi.org/10.1016/j.jhydrol.2022.127726.

Ha, H., Q. D. Bui, H. D. Nguyen, B. T. Pham, T. D. Lai, and C. Luu. 2023. "A Practical Approach to Flood Hazard, Vulnerability, and Risk Assessing and Mapping for Quang Binh Province, Vietnam." *Environment, Development and Sustainability* 25: 1101–1130. https://doi.org/10.1007/s10668-021-02041-4.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. "The Elements of Statistical Learning: Data Mining, Inference and Prediction." In *EAS Publications Series*. 2nd edn. XXII, 745. Springer.

Hitouri, S., M. Mohajane, M. Lahsaini, et al. 2024. "Flood Susceptibility Mapping Using SAR Data and Machine Learning Algorithms in a Small Watershed in Northwestern Morocco." *Remote Sensing* 16, no. 5: 858. https://doi.org/10.3390/rs16050858.

Horton, R. E. 1932. "Drainage-Basin Characteristics." *Eos, Transactions American Geophysical Union* 13, no. 1: 350–361. https://doi.org/10.1029/TR013i001p00350.

Islam, A. R. M. T., M. M. R. Bappi, S. Alqadhi, A. A. Bindajam, J. Mallick, and S. Talukdar. 2023. "Improvement of Flood Susceptibility Mapping by Introducing Hybrid Ensemble Learning Algorithms and High-Resolution Satellite Imageries." *Natural Hazards* 119: 1–37. https://doi.org/10.1007/s11069-023-06106-7.

Jain, A., and S. K. V. P. Indurthy. 2003. "Comparative Analysis of Event-Based Rainfall-Runoff Modeling Techniques—Deterministic, Statistical, and Artificial Neural Networks." *Journal of Hydrologic Engineering* 8, no. 2: 93–98. https://doi.org/10.1061/(asce)1084-0699(2003)8:2(93).

Kalantari, Z., C. S. S. Ferreira, R. P. D. Walsh, A. J. D. Ferreira, and G. Destouni. 2017. "Urbanization Development Under Climate Change: Hydrological Responses in a Peri-Urban Mediterranean Catchment." *Land Degradation and Development* 28, no. 7: 2207–2221. https://doi.org/10.1002/ldr.2747.

Karim, F., M. A. Armin, D. Ahmedt-Aristizabal, L. Tychsen-Smith, and L. Petersson. 2023. "A Review of Hydrodynamic and Machine Learning Approaches for Flood Inundation Modeling." *Water* 15, no. 3: 566. https://doi.org/10.3390/w15030566.

Kaya, C. M., and L. Derin. 2023. "Parameters and Methods Used in Flood Susceptibility Mapping: A Review." *Journal of Water and Climate Change* 14, no. 6: 1935–1960. https://doi.org/10.2166/wcc.2023.035.

Khan, U. T., J. He, and C. Valeo. 2018. "River Flood Prediction Using Fuzzy Neural Networks: An Investigation on Automated Network Architecture." *Water Science and Technology* 2017, no. 1: 238–247. https://doi.org/10.2166/wst.2018.107.

Khosravi, A., Z. Mohammadi, A. Saber, A. Pourzangbar, and D.Neri. 2023. "Anomaly Detection in Real-time Continuous Fruit-based Monitoring of Olive via Extensimeter." http://dx.doi.org/10.2139/ssrn.4652476.

Kumar, V., K. V. Sharma, T. Caloiero, D. J. Mehta, and K. Singh. 2023. "Comprehensive Overview of Flood Modeling Approaches: A Review of Recent Advances." *Hydrology* 10, no. 7: 141. https://doi.org/10.3390/hydrology10070141.

Kurz, S., H. De Gersem, A. Galetzka, et al. 2022. "Hybrid Modeling: Towards the Next Level of Scientific Computing in Engineering." *Journal of Mathematics in Industry* 12, no. 1: 8. https://doi.org/10.1186/s13362-022-00123-0.

Le Nguyen, K., H. Thi Trinh, T. T. Nguyen, and H. D. Nguyen. 2023. "Comparative Study on the Performance of Different Machine Learning Techniques to Predict the Shear Strength of RC Deep Beams: Model Selection and Industry Implications." *Expert Systems With Applications* 230: 120649. https://doi.org/10.1016/j.eswa.2023.120649.

Lechowska, E. 2018. "What Determines Flood Risk Perception? A Review of Factors of Flood Risk Perception and Relations Between Its Basic Elements." *Natural Hazards* 94, no. 3: 1341–1366. https://doi.org/10.1007/s11069-018-3480-z.

Li, R., and G. Xu. 2023. "Assessing the Impacts of Reservoirs on Downstream Hydrological Frequency Based on a General Rainfall-Reservoir Index." *Frontiers in Earth Science* 11: 1–16. https://doi.org/10.3389/feart.2023.1204640.

Liao, M., H. Wen, L. Yang, G. Wang, X. Xiang, and X. Liang. 2024. "Improving the Model Robustness of Flood Hazard Mapping Based on Hyperparameter Optimization of Random Forest." *Expert Systems With Applications* 241: 122682. https://doi.org/10.1016/j.eswa.2023.122682.

Liao, Y., Z. Wang, X. Chen, and C. Lai. 2023. "Fast Simulation and Prediction of Urban Pluvial Floods Using a Deep Convolutional Neural Network Model." *Journal of Hydrology* 624: 129945. https://doi.org/10.1016/j.jhydrol.2023.129945.

Linh, N. T. T., H. Ruigar, S. Golian, et al. 2021. "Flood Prediction Based on Climatic Signals Using Wavelet Neural Network." *Acta Geophysica* 69: 1413–1426. https://doi.org/10.1007/s11600-021-00620-7.

Ludwig, P., F. Ehmele, M. J. Franca, et al. 2023. "A Multi-Disciplinary Analysis of the Exceptional Flood Event of July 2021 in Central Europe–Part 2: Historical Context and Relation to Climate Change." *Natural Hazards and Earth System Sciences* 23, no. 4: 1287–1311. https://doi.org/10.5194/nhess-23-1287-2023.

Luu, C., Q. D. Bui, R. Costache, et al. 2021. "Flood-Prone Area Mapping Using Machine Learning Techniques: A Case Study of Quang Binh Province, Vietnam." *Natural Hazards* 108: 3229–3251. https://doi.org/10.1007/s11069-021-04821-7.

Lyu, H.-M., and Z.-Y. Yin. 2023. "Flood Susceptibility Prediction Using Tree-Based Machine Learning Models in the GBA." *Sustainable Cities and Society* 97: 104744. https://doi.org/10.1016/j.scs.2023.104744.

Madhuri, R., S. Sistla, and K. Srinivasa Raju. 2021. "Application of Machine Learning Algorithms for Flood Susceptibility Assessment and Risk Management." *Journal of Water and Climate Change* 12, no. 6: 2608–2623. https://doi.org/10.2166/wcc.2021.051.

Mahdizadeh Gharakhanlou, N., and L. Perez. 2022. "Spatial Prediction of Current and Future Flood Susceptibility: Examining the Implications of Changing Climates on Flood Susceptibility Using Machine Learning Models." *Entropy* 24: 1630. https://doi.org/10.3390/e24111630.

Maspo, N. A., A. N. Bin Harun, M. Goto, F. Cheros, N. A. Haron, and M. N. Mohd Nawi. 2020. "Evaluation of Machine Learning Approach in Flood Prediction Scenarios and Its Input Parameters: A Systematic Review." *IOP Conference Series: Earth and Environmental Science* 479, no. 1: 012038. https://doi.org/10.1088/1755-1315/479/1/012038.

Matthias, A. 2004. "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata." *Ethics and Information Technology* 6, no. 3: 175–183. https://doi.org/10.1007/s10676-004-3422-1.

Meliho, M., A. Khattabi, and J. Asinyo. 2021. "Spatial Modeling of Flood Susceptibility Using Machine Learning Algorithms." *Arabian Journal of Geosciences* 14: 2243. https://doi.org/10.1007/s12517-021-08610-1.

Mohr, S., U. Ehret, M. Kunz, et al. 2022. "A Multi-Disciplinary Analysis of the Exceptional Flood Event of July 2021 in Central Europe—Part 1: Event Description and Analysis." *Natural Hazards and Earth System Sciences* 23, no. 2: 1–44. https://doi.org/10.5194/nhess-23-525-2023.

Mosavi, A., P. Ozturk, and K. W. Chau. 2018. "Flood Prediction Using Machine Learning Models: Literature Review." *Water* 10, no. 11: 1536. https://doi.org/10.3390/w10111536.

Mudashiru, R. B., N. Sabtu, I. Abustan, and W. Balogun. 2021. "Flood Hazard Mapping Methods: A Review." *Journal of Hydrology* 603: 126846. https://doi.org/10.1016/j.jhydrol.2021.126846.

Nguyen, H. D. 2022. "GIS-Based Hybrid Machine Learning for Flood Susceptibility Prediction in the Nhat Le–Kien Giang Watershed, Vietnam." *Earth Science Informatics* 15: 2369–2386. https://doi.org/10.1007/s12145-022-00825-4.

Nguyen, H. D., C. P. Van, and A. D. Do. 2023. "Application of Hybrid Model-Based Deep Learning and Swarm-Based Optimizers for Flood Susceptibility Prediction in Binh Dinh Province, Vietnam." *Earth Science Informatics* 16: 1173–1193. https://doi.org/10.1007/s12145-023-00954-4.

Obregon, J., and J. Y. Jung. 2022. "Explanation of Ensemble Models." In *Human-Centered Artificial Intelligence: Research and Applications*. Academic Press. https://doi.org/10.1016/B978-0-323-85648-5.00011-6.

Ong, Y. J., N. Baracaldo, and Y. Zhou. 2022. "Tree-Based Models for Federated Learning Systems." In *Federated Learning: A Comprehensive Overview of Methods and Applications*. Springer. https://doi.org/10.1007/978-3-030-96896-0_2.

Park, S.-J., and D.-K. Lee. 2020. "Prediction of Coastal Flooding Risk Under Climate Change Impacts in South Korea Using Machine Learning Algorithms." *Environmental Research Letters* 15, no. 9: 094052. https://doi.org/10.1088/1748-9326/aba5b3.

Pham, B. T., C. Luu, T. V. Phong, et al. 2021. "Can Deep Learning Algorithms Outperform Benchmark Machine Learning Algorithms in Flood Susceptibility Modeling?" *Journal of Hydrology* 592: 125615. https://doi.org/10.1016/j.jhydrol.2020.125615.

Pourzangbar, A. 2012. "Determination of the Most Effective Parameters on Scour Depth at Seawalls Using Genetic Programming (GP)." In *The 10th International Conference on Coasts, Ports and Marine Structures (ICOPMASS 2012)*. Icopmas.

Pourzangbar, A., M. Brocchini, A. Saber, J. Mahjoobi, M. Mirzaaghasi, and M. Barzegar. 2017. "Prediction of Scour Depth at Breakwaters due to Non-Breaking Waves Using Machine Learning Approaches." *Applied Ocean Research* 63: 120–128. https://doi.org/10.1016/j.apor.2017.01.012.

Pourzangbar, A., M. Jalali, and M. Brocchini. 2023. "Machine Learning Application in Modelling Marine and Coastal Phenomena: A Critical Review." *Frontiers in Environmental Engineering* 2: 1–27. https://doi.org/10.3389/fenve.2023.1235557.

Pourzangbar, A., M. A. Losada, A. Saber, et al. 2017. "Prediction of Non-Breaking Wave Induced Scour Depth at the Trunk Section of Breakwaters Using Genetic Programming and Artificial Neural Networks." *Coastal Engineering* 121: 107–118. https://doi.org/10.1016/j.coastaleng.2016.12.008.

Pourzangbar, A., A. Saber, A. Yeganeh-Bakhtiary, and L. R. Ahari. 2017. "Predicting Scour Depth at Seawalls Using GP and ANNs." *Journal of Hydroinformatics* 19, no. 3: 349–363. https://doi.org/10.2166/hydro.2017.125.

Pradhan, B., S. Lee, A. Dikshit, and H. Kim. 2023. "Spatial Flood Susceptibility Mapping Using an Explainable Artificial Intelligence (XAI) Model." *Geoscience Frontiers* 14, no. 6: 101625. https://doi.org/10.1016/j.gsf.2023.101625.

Quinlan, J. R. 1996. "Bagging, Boosting, and C4.5." In *Proceedings of the National Conference on Artificial Intelligence*, 1.

Rasheed, M. A., M. A. Rasheed, H. A. Tanweer, S. J. Yawar, and L. Farhi. 2022. "A Review on Machine Learning-Based Neural Network Techniques for Flood Prediction REVIEW." *VFAST Transactions on*

Software Engineering 10, no. 1: 66–77. https://doi.org/10.21015/vtse.v10i1.835.

Razavi-Termeh, S. V., A. Kornejady, H. R. Pourghasemi, and S. Keesstra. 2018. "Flood Susceptibility Mapping Using Novel Ensembles of Adaptive Neuro Fuzzy Inference System and Metaheuristic Algorithms." *Science of the Total Environment* 615: 438–451. https://doi.org/10.1016/J.SCITOTENV.2017.09.262.

Razavi-Termeh, S. V., A. Pourzangbar, A. Sadeghi-Niaraki, M. J. Franca, and S. M. Choi. 2025. "Metaheuristic-Driven Enhancement of Categorical Boosting Algorithm for Flood-Prone Areas Mapping." *International Journal of Applied Earth Observation and Geoinformation* 136: 104–357. https://doi.org/10.1016/J.JAG.2025.104357.

Razavi-Termeh, S. V., A. Sadeghi-Niaraki, and S. M. Choi. 2023. "A New Approach Based on Biology-Inspired Metaheuristic Algorithms in Combination With Random Forest to Enhance the Flood Susceptibility Mapping." *Journal of Environmental Management* 345: 118790. https://doi.org/10.1016/j.jenvman.2023.118790.

Riazi, M., K. Khosravi, K. Shahedi, et al. 2023. "Enhancing Flood Susceptibility Modeling Using Multi-Temporal SAR Images, CHIRPS Data, and Hybrid Machine Learning Algorithms." *Journal of Hydrology* 604: 126–853. https://doi.org/10.1016/j.jhydrol.2022.126853.

Ribeiro, M. T., S. Singh, and C. Guestrin. 2016. "'Why Should i Trust You?' Explaining the Predictions of Any Classifier." In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 1135–1144. https://doi.org/10.1145/2939672.2939778.

Ritchie, H., and P. Rosado. 2024. "Natural Disasters Where and From Which Disasters Do People Die? What Can We Do to Prevent Deaths From Natural Disasters?" https://ourworldindata.org/natural-disasters.

Rong, G., S. Alu, K. Li, et al. 2020. "Rainfall Induced Landslide Susceptibility Mapping Based on Bayesian Optimized Random Forest and Gradient Boosting Decision Tree Models—A Case Study of Shuicheng County, China." *Water* 12, no. 11: 3066. https://doi.org/10.3390/w12113066.

Sahoo, A., S. Samantaray, and S. Paul. 2021. "Efficacy of ANFIS-GOA Technique in Flood Prediction: A Case Study of Mahanadi River Basin in India." *H2Open Journal* 4, no. 1: 137–156. https://doi.org/10.2166/h2oj.2021.090.

Sajedi-Hosseini, F., B. Choubin, K. Solaimani, A. Cerdà, and A. Kavian. 2018. "Spatial Prediction of Soil Erosion Susceptibility Using a Fuzzy Analytical Network Process: Application of the Fuzzy Decision Making Trial and Evaluation Laboratory Approach." *Land Degradation and Development* 29, no. 9: 3092–3103. https://doi.org/10.1002/ldr.3058.

Scotti, V., M. Giannini, and F. Cioffi. 2020. "Enhanced Flood Mapping Using Synthetic Aperture Radar (SAR) Images, Hydraulic Modelling, and Social Media: A Case Study of Hurricane Harvey (Houston, TX)." *Journal of Flood Risk Management* 13, no. 4: e12647. https://doi.org/10.1111/jfr3.12647.

Senthil Kumar, A. R., K. P. Sudheer, S. K. Jain, and P. K. Agarwal. 2005. "Rainfall-Runoff Modelling Using Artificial Neural Networks: Comparison of Network Types." *Hydrological Processes* 19, no. 6: 1277–1291. https://doi.org/10.1002/hyp.5581.

Shapley, L. S. 1953. "Stochastic Games." *Proceedings of the National Academy of Sciences of the United States of America* 39: 1095–1100. https://doi.org/10.1073/pnas.39.10.109.

Shu, C., and T. B. M. J. Ouarda. 2008. "Regional Flood Frequency Analysis at Ungauged Sites Using the Adaptive Neuro-Fuzzy Inference System." *Journal of Hydrology* 349, no. 1–2: 31–43. https://doi.org/10.1016/j.jhydrol.2007.10.050.

Tabbussum, R., and A. Q. Dar. 2021. "Performance Evaluation of Artificial Intelligence Paradigms—Artificial Neural Networks, Fuzzy Logic, and Adaptive Neuro-Fuzzy Inference System for Flood Prediction." *Environmental Science and Pollution Research* 28: 25265–25282. https://doi.org/10.1007/s11356-021-12410-1.

Tellman, B., J. A. Sullivan, and C. Kuhn. 2021. "Satellite Imaging Reveals Increased Proportion of Population Exposed to Floods." *Nature* 596, no. 7870: 80–86. https://doi.org/10.1038/s41586-021-03695-w.

Tomiczek, T., A. Wargula, P. Lomónaco, et al. 2020. "Physical Model Investigation of Mid-Scale Mangrove Effects on Flow Hydrodynamics and Pressures and Loads in the Built Environment." *Coastal Engineering* 162: 103791. https://doi.org/10.1016/j.coastaleng.2020.103791.

Ullah, K., Y. Wang, Z. Fang, L. Wang, and M. Rahman. 2022. "Multi-Hazard Susceptibility Mapping Based on Convolutional Neural Networks." *Geoscience Frontiers* 13, no. 5: 101–425. https://doi.org/10.1016/j.gsf.2022.101425.

Ullah, K., and J. Zhang. 2020. "GIS-Based Flood Hazard Mapping Using Relative Frequency Ratio Method: A Case Study of Panjkora River Basin, Eastern Hindu Kush, Pakistan." *PLoS One* 15, no. 3: e0229153. https://doi.org/10.1371/journal.pone.0229153.

Vafakhah, M., S. M. H. Loor, and H. Pourghasemi. 2020. "Comparing Performance of Random Forest and Adaptive Neuro-Fuzzy Inference System Data Mining Models for Flood Susceptibility Mapping." *Arabian Journal of Geosciences* 13: 417. https://doi.org/10.1007/s12517-020-05363-1.

Valero, D., A. Bayón, and M. J. Franca. 2024. "Urban Flood Drifters (UFDs): Onset of Movement." *Science of the Total Environment* 927: 171568. https://doi.org/10.1016/j.scitotenv.2024.171568.

Wang, Y., H. Hong, W. Chen, et al. 2019. "Flood Susceptibility Mapping in Dingnan County (China) Using Adaptive Neuro-Fuzzy Inference System With Biogeography Based Optimization and Imperialistic Competitive Algorithm." *Journal of Environmental Management* 247: 712–729. https://doi.org/10.1016/j.jenvman.2019.06.102.

Wang, Y., M. Khodadadzadeh, and R. Zurita-Milla. 2023. "Spatial+: A New Cross-Validation Method to Evaluate Geospatial Machine Learning Models." *International Journal of Applied Earth Observation and Geoinformation* 121: 103364. https://doi.org/10.1016/j.jag.2023.103364.

Witten, I. H., E. Frank, M. A. Hall, and C. J. Pal. 2016. *Data Mining: Practical Machine Learning Tools and Techniques*. 4th edn, 645. Morgan Kaufmann.

Xie, S., W. Wu, S. Mooser, Q. J. Wang, R. Nathan, and Y. Huang. 2021. "Artificial Neural Network Based Hybrid Modeling Approach for Flood Inundation Modeling." *Journal of Hydrology* 592: 125–605. https://doi.org/10.1016/j.jhydrol.2020.125605.

Yang, L., and A. Shami. 2020. "On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice." *Neurocomputing* 415: 295–316. https://doi.org/10.1016/j.neucom.2020.07.061.

Youssef, A. M., H. R. Pourghasemi, and B. A. El-Haddad. 2022. "Advanced Machine Learning Algorithms for Flood Susceptibility Modeling—Performance Comparison: Red Sea, Egypt." *Environmental Science and Pollution Research* 29: 66768–66792. https://doi.org/10.1007/s11356-022-20213-1.

Yu, H., Z. Luo, L. Wang, X. Ding, and S. Wang. 2023. "Improving the Accuracy of Flood Susceptibility Prediction by Combining Machine Learning Models and the Expanded Flood Inventory Data." *Remote Sensing* 15: 3601. https://doi.org/10.3390/rs15143601.

Zhao, G., B. Pang, Z. Xu, et al. 2021. "Improving Urban Flood Susceptibility Mapping Using Transfer Learning." *Journal of Hydrology* 602: 126777. https://doi.org/10.1016/j.jhydrol.2021.126777.

Zhao, G., B. Pang, Z. Xu, D. Peng, and D. Zuo. 2020. "Urban Flood Susceptibility Assessment Based on Convolutional Neural Networks." *Journal of Hydrology* 590: 125–235. https://doi.org/10.1016/j.jhydrol.2020.125235.

Zhou, G., M. Cui, J. Wan, and S. Zhang. 2021. "A Review on Snowmelt Models: Progress and Prospect." *Sustainability (Switzerland)* 13, no. 20: 11485. https://doi.org/10.3390/su132011485.

Appendix 1

Nomenclature

Abbreviation	Definition
Acc	Accuracy
AdaBoost	Adaptive Boosting
ADT	Alternating Decision Tree
Advanced-DT	Advanced Decision Tree
AHP	Analytical Hierarchy Process
AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
API	Antecedent Precipitation Index
AUROC	Area Under Receiver Operating Characteristic
BA	Bat Algorithm
Bagging-Ens	Bagging-based Ensemble
Basic-DT	Basic Decision Tree
BFT	Bagging Functional Tree
BGLM	Boosted Generalized Linear Model
B-LMT	Bagging-Logistic Model Tree
B-LMT	Bagging-Logistic Model Tree
Boosting-Ens	Boosting-based Ensemble
BPNN	Back Propagation Neural Network
BPSO	Binary Particle Swarm Optimization
BRT	Boosted Regression Tree
CART	Classification And Regression Tree
CatBoost	Categorical Boosting
CC	Pearson's Correlation Coefficient
CDT	Credal Decision Tree
CDT-FT	Credal Decision Tree-Functional Tree
CNN	Convolutional Neural Network
DB	Deep Boost
DBN	Deep Belief Network
DBP	Deep belief network with Back Propagation algorithm
DE	Differential Evolution
DEM	Digital Elevation Model
DisFau	Distance to Fault
DisRiv	Distance To River
DisRoa	Distance to Road
DL	Deep Learning
DLNN	Deep Learning Neural Network
DNN	Deep Neural Network
DNN-AO	DNN-Aquila Optimizer
DNN-NMRA	Deep Neural Network-Naked Mole-Rat Algorithm
DNN-SSO	Deep Neural Network-Social Spider Optimization
Drain-Den	Drainage Density
DT	Decision Tree
DTM	Digital Terrain Model
EA	Evolutionary Algorithm
ELM	Extreme Learning Machine
EMCA	Ensemble Model Committee Averaging
EMMean	Ensemble Model to estimate Mean
EPS	Ensemble Prediction System
ERT	Extremely Randomized Tree
FART	Fuzzy Adaptive Resonance Theory
FHM	Flood Hazard Mapping
FID	Flood Inventory Data
FIM	Flood Inundation Mapping
FL	Fuzzy Logic
FL-Hyb	Fuzzy Logic-based Hybrid
FloAcc	Flow Accumulation
FlowDir	Flow Direction
Forest-Ens	Forest-based Ensemble
FPCA	Functional Principal Component Analysis
FR	Frequency Ratio
FSM	Flood Susceptibility Mapping
GA	Genetic Algorithm
GB	Gradient Boosting
GBDT	Gradient Boosting Decision Tree
GBM	Gradient Boosting Machine
GOA	Grasshopper Optimization Algorithm
Grid-NN	Grid-based Neural Networks
GRU	Gated Recurrent Units
ICA	Imperialistic Competitive Algorithm
IGR	Information Gain Ratio
IQR	Interquartile Range
IWO	Invasive Weed Optimization
Kern-Class	Kernel-based Classification
Kernel-Hyb	Kernel-based Hybrid
Kern-Reg	Kernel-based Regression
Kern-SVM	SVM-based model
K-NN	K-Nearest Neighbors
K-SVM	Kernel Support Vector Machine
LightGBM	Light Gradient Boosting Machine

Abbreviation	Definition
LIME	Local Interpretable Model-Agnostic Explanations
LMT	Logistic Model Tree
LR	Logistic Regression
LSS-LSTM	Local Spatial Sequential Long Short-Term Memory
LSTM	Long Short-Term Memory
LSTM-ED	Long Short-Term Memory based Encoder-Decoder
LULC	Land Use Land Cover
MAD	Median Absolute Deviation
MAE	Mean Absolute Error
MARS	Multivariate Adaptive Regression Splines
MCDA	Multi-Criteria Decision Analysis
MDA	Multivariate Discriminant Analysis
ML	Machine Learning
MLP	Multi-Layer Perceptron
MONF-EG-PSO	Metaheuristic Optimization and Neural Fuzzy inference-Evolutionary Genetic-Particle Swarm Optimization
MSE	Mean Squared Error
NB	Naïve Bayes
NBT	Naïve Bayes Tree
NDVI	Normalized Difference Vegetation Index
NGBoost	Natural Gradient Boosting
NN	Neural Network
NN-Hyb	Neural Network-based Hybrid
NPR	Negative Predictive Rate
PCA	Principal Component Analysis
PLS-SEM	Partial Least Square-Structural Equation Model
PPR	Positive Predictive Rate
PSO	Particle Swarm Optimization
QPSO-CDT	Quantum Particle Swarm Optimization-Credal Decision Tree
RBF	Radial Basis Function
REPTree	Reduced Error Pruning Tree
RF	Random Forest
Riv-Den	River (Stream) Density
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
ROF	Rotation Forest
RS	Remote Sensing
RS-GAM	Random Subsampling-Generalized Additive Model
RS-MARS	Random Subsampling-Multivariate Adaptive Regression Splines
RSub	Random Subsampling
SBO	Satin Bowerbird Optimization
SCNN	Simple CNN
Sequential-NN	Sequential Neural Network
SGB	Stochastic Gradient Boosting
SGD-WOE	Stochastic Gradient Descending-Weights Of Evidence
Shallow-NN	Shallow learning NNs
SHAP	Shapley Additive Explanations
SMA	Slime Mold Algorithm
SOM	Self-Organizing Map
SPI	Stream Power Index
STA-LSTM	Spatio-Temporal Attention Long Short-Term Memory
STI	Sediment Transport Index
SVC	Support Vector Classification
SVM	Support Vector Machine
SVR	Support Vector Regression
SWARA	Stepwise Weight Assessment Ratio Analysis
TPI	Topographic Position Index
Tree-Hyb	Tree-based Hybrid
TRI	Terrain Ruggedness Index
TWI	Topographic Wetness Index
VIF	Variance Inflation Factor
WAve-Hyb	Weighted Average-based Hybrid
WCA	Water Cycle Algorithm
WELLSVM	Weakly Labeled SVM
WNN	Wavelet Neural Network
WOE	Weight Of Evidence
XAI	Explainable AI
XGBoost	Extreme Gradient Boosting

Appendix 2

List of Reviewed Articles

[1] Abu El-Magd, S., A. Maged, and H.I. Farhat. 2022. "Hybrid-Based Bayesian Algorithm and Hydrologic Indices for Flash Flood Vulnerability Assessment in Coastal Regions: Machine Learning, Risk Prediction, and Environmental Impact." Environmental Science and Pollution Research 29: 57345–56. 10.1007/s11356-022-19903-7.

[2] Al-Abadi, A.M. 2018. "Mapping Flood Susceptibility in an Arid Region of Southern Iraq Using Ensemble Machine Learning Classifiers: a Comparative Study." Arabian Journal of Geosciences 11: 218. 10.1007/s12517-018-3584-5.

[3] Al-Areeq, A.M., S.I. Abba, M.A. Yassin, M. Benaafi, M. Ghaleb, and I.H. Aljundi. 2022. "Computational Machine Learning Approach for Flood Susceptibility Assessment Integrated With Remote Sensing and GIS Techniques From Jeddah." Saudi Arabia. Remote Sensing 14: 5515. 10.3390/rs14215515.

[4] Aldiansyah, S., and F. Wardani. 2023. "Evaluation of Flood Susceptibility Prediction Based on a Resampling Method Using Machine Learning." Journal of Water and Climate Change 14(3): 937–61. 10.2166/wcc.2023.494.

[5] Al-Juaidi, A.E.M., A.M. Nassar, and O.E.M. Al-Juaidi. 2018. "Evaluation of Flood Susceptibility Mapping Using Logistic Regression and GIS Conditioning Factors." Arabian Journal of Geosciences 11: 765. 10.1007/s12517-018-4095-0.

[6] Andaryani, S., V. Nourani, A.T. Haghighi, and S. Keesstra. 2021. "Integration of Hard and Soft Supervised Machine Learning for Flood Susceptibility Mapping." Journal of Environmental Management 291: 112731. 10.1016/j.jenvman.2021.112731.

[7] Arabameri, A., S. Saha, W. Chen, J. Roy, B. Pradhan, and D.T. Bui. 2020. "Flash Flood Susceptibility Modelling Using Functional Tree and Hybrid Ensemble Techniques." Journal of Hydrology 587: 125007. 10.1016/j.jhydrol.2020.125007.

[8] Arora, A., A. Arabameri, M. Pandey, M.A. Siddiqui, U.K. Shukla, D.T. Bui, V.N. Mishra, and A. Bhardwaj. 2021. "Optimization of State-of-the-Art Fuzzy-Metaheuristic ANFIS-Based Machine Learning Models for Flood Susceptibility Prediction Mapping in the Middle Ganga Plain, India." Science of the Total Environment 750: 141565. 10.1016/j.scitotenv.2020.141565.

[9] Aydin, H.E., and M.C. Iban. 2023. "Predicting and Analyzing Flood Susceptibility Using Boosting-Based Ensemble Machine Learning Algorithms With SHapley Additive exPlanations." Natural Hazards 116: 2957–91. 10.1007/s11069-022-05793-y.

[10] Band, S.S., S. Janizadeh, S.C. Pal, A. Saha, R. Chakrabortty, A.M. Melesse, and A. Mosavi. 2020. "Flash Flood Susceptibility Modeling Using New Approaches of Hybrid and Ensemble Tree-Based Machine Learning Algorithms." Remote Sensing 12(21): 3568. 10.3390/rs12213568.

[11] Bui, D.T., N.-D. Hoang, T.-D. Pham, P.-T.T. Ngo, P.V. Hoa, N.Q. Minh, X.-T. Tran, and P. Samui. 2019. "A New Intelligence Approach Based on GIS-Based Multivariate Adaptive Regression Splines and Metaheuristic Optimization for Predicting Flash Flood Susceptible Areas at High-Frequency Tropical Typhoon Area." Journal of Hydrology 575: 314–26. 10.1016/j.jhydrol.2019.05.046.

[12] Bui, D.T., P.-T.T. Ngo, T.D. Pham, A. Jaafari, N.Q. Minh, P.V. Hoa, and P. Samui. 2019. "A Novel Hybrid Approach Based on a Swarm Intelligence Optimized Extreme Learning Machine for Flash Flood Susceptibility Mapping." CATENA 179: 184–96. 10.1016/j.catena.2019.04.009.

[13] Bui, D.T., B. Pradhan, H. Nampak, Q.T. Bui, Q.A. Tran, and Q.P. Nguyen. 2016. "Hybrid Artificial Intelligence Approach Based on Neural Fuzzy Inference Model and Metaheuristic Optimization for Flood Susceptibility Modeling in a High-Frequency Tropical Cyclone Area Using GIS." Journal of Hydrology 540: 317–30. 10.1016/j.jhydrol.2016.06.027.

[14] Bui, Q.-T., Q.-H. Nguyen, X.L. Nguyen, V.D. Pham, H.D. Nguyen, and V.-M. Pham. 2020. "Verification of Novel Integrations of Swarm Intelligence Algorithms Into Deep Learning Neural Network for Flood Susceptibility Mapping." Journal of Hydrology 581: 124379. 10.1016/j.jhydrol.2019.124379.

[15] Chakrabortty, R., S.C. Pal, S. Janizadeh, et al. 2021. "Impact of Climate Change on Future Flood Susceptibility: An Evaluation Based on Deep Learning Algorithms and GCM Model." Water Resources Management 35: 4251–74. 10.1007/s11269-021-02944-x.

[16] Chapi, K., V.P. Singh, A. Shirzadi, H. Shahabi, D.T. Bui, B.T. Pham, and K. Khosravi. 2017. "A Novel Hybrid Artificial Intelligence Approach for Flood Susceptibility Assessment." Environmental Modelling & Software 95: 229–45. 10.1016/j.envsoft.2017.06.012.

[17] Chen, C., J. Jiang, Z. Liao, Y. Zhou, H. Wang, and Q. Pei. 2022. "A Short-Term Flood Prediction Based on Spatial Deep Learning Network: A Case Study for Xi County, China." Journal of Hydrology 607: 127535. 10.1016/j.jhydrol.2022.127535.

[18] Chen, W., H. Hong, S. Li, H. Shahabi, Y. Wang, X. Wang, and B. Ahmad. 2019. "Flood Susceptibility Modeling Using a Novel Hybrid Approach of Reduced-Error Pruning Trees With Bagging and Random Subspace Ensembles." Journal of Hydrology 575: 864–73. 10.1016/j.jhydrol.2019.05.089.

[19] Choubin, B., E. Moradi, M. Golshan, J. Adamowski, F. Sajedi-Hosseini, and A. Mosavi. 2019. "An Ensemble Prediction of Flood Susceptibility Using Multivariate Discriminant Analysis, Classification and Regression Trees, and Support Vector Machines." Science of the Total Environment 651(Part 2): 2087–96. 10.1016/j.scitotenv.2018.10.064.

[20] Costache, R., A. Arabameri, I. Costache, et al. 2022. "New Machine Learning Ensemble for Flood Susceptibility Estimation." Water Resources Management 36: 4765–83. 10.1007/s11269-022-03276-0.

[21] Costache, R., A. Arabameri, I. Elkhrachy, O. Ghorbanzadeh, and Q.B. Pham. 2021. "Detection of Areas Prone to Flood Risk Using State-of-the-Art Machine Learning Models." Geomatics, Natural Hazards and Risk 12(1): 1488–507. 10.1080/19475705.2021.1920480.

[22] Deroliya, P., M. Ghosh, M.P. Mohanty, S. Ghosh, K.H.V. Durga Rao, and S. Karmakar. 2022. "A Novel Flood Risk Mapping Approach With Machine Learning Considering Geomorphic and Socio-Economic Vulnerability Dimensions." Science of the Total Environment 851(1): 158002. 10.1016/j.scitotenv.2022.158002.

[23] Ding, Y., Y. Zhu, J. Feng, P. Zhang, and Z. Cheng. 2020. "Interpretable Spatio-Temporal Attention LSTM Model for Flood Forecasting." Neurocomputing 403: 348–59. 10.1016/j.neucom.2020.04.110.

[24] Dodangeh, E., B. Choubin, A.N. Eigdir, N. Nabipour, M. Panahi, S. Shamshirband, and A. Mosavi. 2020. "Integrated Machine Learning Methods With Resampling Algorithms for Flood Susceptibility Prediction." Science of the Total Environment 705: 135983. 10.1016/j.scitotenv.2019.135983.

[25] El-Haddad, B.A., A.M. Youssef, H.R. Pourghasemi, et al. 2021. "Flood Susceptibility Prediction Using Four Machine Learning Techniques and Comparison of Their Performance at Wadi Qena Basin, Egypt." Nature Hazards 105(1): 83–114. 10.1007/s11069-020-04296-y.

[26] El-Magd, S.A.A. 2022. "Random Forest and Naïve Bayes Approaches as Tools for Flash Flood Hazard Susceptibility Prediction, South Ras El-Zait, Gulf of Suez Coast, Egypt." Arabian Journal of Geosciences 15: 217. 10.1007/s12517-022-09531-3.

[27] El-Magd, S.A.A., B. Pradhan, and A. Alamri. 2021. "Machine Learning Algorithm for Flash Flood Prediction Mapping in Wadi El-Laqeita and Surroundings, Central Eastern Desert, Egypt." Arabian Journal of Geosciences 14: 323. 10.1007/s12517-021-06466-z.

[28] Eslaminezhad, S.A., M. Eftekhari, A. Azma, R. Kiyanfar, and M. Akbari. 2022. "Assessment of Flood Susceptibility Prediction Based on Optimized Tree-Based Machine Learning Models." Journal of Water and Climate Change 13. 10.2166/wcc.2022.435.

[29] Fang, Z., Y. Wang, L. Peng, and H. Hong. 2021. "Predicting Flood Susceptibility Using LSTM Neural Networks." Journal of Hydrology 594. 10.1016/j.jhydrol.2020.125734.

[30] Gude, V., S. Corns, and S. Long. 2020. "Flood Prediction and Uncertainty Estimation Using Deep Learning." Water 12(3): 884. 10.3390/w12030884.

[31] Guo, Z., V. Moosavi, and J.P. Leitão. 2022. "Data-driven rapid flood prediction mapping with catchment generalizability." Journal of Hydrology 609: 127726. 10.1016/j.jhydrol.2022.127726.

[32] Ha, H., Q.D. Bui, H.D. Nguyen, B.T. Pham, T.D. Lai, and C. Luu. 2023. "A Practical Approach to Flood Hazard, Vulnerability, and Risk Assessing and Mapping for Quang Binh province." Vietnam. Environment, Development and Sustainability 25: 1101–30. 10.1007/s10668-021-02041-4.

[33] Hong, H., Panahi, M., Shirzadi, A., Ma, T., Liu, J., Zhu, A.-X., Chen, W., Kougias, I., & Kazakis, N. (2018). Flood Susceptibility Assessment in Hengfeng Area Coupling Adaptive Neuro-Fuzzy Inference System With Genetic Algorithm and Differential Evolution. Science of the Total Environment, 621, 1124-1141. 10.1016/j.scitotenv.2017.10.114.

[34] Islam, A.R.M.T., M.M.R. Bappi, S. Alqadhi, A.A. Bindajam, J. Mallick, and S. Talukdar. 2023. "Improvement of Flood Susceptibility Mapping by Introducing Hybrid Ensemble Learning Algorithms and High-Resolution Satellite Imageries." Natural Hazards 119: 1–37. 10.1007/s11069-023-06106-7.

[35] Islam, A.R.M.T., S. Talukdar, S. Mahato, S. Kundu, K.U. Eibek, Q.B. Pham, A. Kuriaç, and N.T.T. Linh. 2021. "Flood Susceptibility Modelling Using Advanced Ensemble Machine Learning Models." Geoscience Frontiers 12(3): 101075. 10.1016/j.gsf.2020.09.006.

[36] Janizadeh, S., M. Vafakhah, Z. Kapelan, and N. Mobarghaee Dinan. 2021. "Novel Bayesian Additive Regression Tree Methodology for Flood Susceptibility Modeling." Water Resources Management 35: 4621–46. 10.1007/s11269-021-02972-7.

[37] Kabir, S., S. Patidar, X. Xia, Q. Liang, J. Neal, and G. Pender. 2020. "A Deep Convolutional Neural Network Model for Rapid Prediction of Fluvial Flood Inundation." Journal of Hydrology 590: 125481. 10.1016/j.jhydrol.2020.125481.

[38] Kabir, S., S. Patidar, X. Xia, Q. Liang, J. Neal, and G. Pender. 2020. "A Deep Convolutional Neural Network Model for Rapid Prediction of Fluvial Flood Inundation." Journal of Hydrology 590. 10.1016/j.jhydrol.2020.125481.

[39] Kao, I.-F., Zhou, Y., Chang, L.-C., and Chang, F.-J. (2020). Exploring a Long Short-Term Memory Based Encoder-Decoder Framework for Multi-Step-Ahead Flood Forecasting. Journal of Hydrology, 583, 124631. 10.1016/j.jhydrol.2020.124631.

[40] Khan, U.T., J. He, and C. Valeo. 2018. "River Flood Prediction Using Fuzzy Neural Networks: An Investigation on Automated Network Architecture." Water Science and Technology 2017(1): 238–47. 10.2166/wst.2018.107.

[41] Khosravi, K., M. Panahi, A. Golkarian, S.D. Keesstra, P.M. Saco, D.T. Bui, and S. Lee. 2020. "Convolutional Neural Network Approach for Spatial Prediction of Flood Hazard at National Scale of Iran." Journal of Hydrology 591: 125552. 10.1016/j.jhydrol.2020.125552.

[42] Khosravi, K., B.T. Pham, K. Chapi, A. Shirzadi, H. Shahabi, I. Revhaug, I. Prakash, and D.T. Bui. 2018. "A Comparative Assessment of Decision Trees Algorithms for Flash Flood Susceptibility Modeling at Haraz Watershed, Northern Iran." Science of the Total Environment 627: 744–55. 10.1016/j.scitotenv.2018.01.266.

[43] Lawal, Z.K., Yassin, H., & Zakari, R.Y. (2021). Flood Prediction Using Machine Learning Models: A Case Study of Kebbi State Nigeria. 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), 08-10 December 2021 in Brisbane, Australia. 10.1109/CSDE53843.2021.9718497.

[44] Lei, X., W. Chen, M. Panahi, F. Falah, O. Rahmati, E. Uuema, Z. Kalantari, et al. 2021. "Urban Flood Modeling Using Deep-Learning Approaches in Seoul, South Korea." Journal of Hydrology 601: 126684. 10.1016/j.jhydrol.2021.126684.

[45] Liao, Y., Z. Wang, X. Chen, and C. Lai. 2023. "Fast Simulation and Prediction of Urban Pluvial Floods Using a Deep Convolutional Neural Network Model." Journal of Hydrology 624: 129945. 10.1016/j.jhydrol.2023.129945.

[46] Linh, N.T.T., H. Ruigar, S. Golian, et al. 2021. "Flood Prediction Based on Climatic Signals Using Wavelet Neural Network." Acta Geophysica 69: 1413–26. 10.1007/s11600-021-00620-7.

[47] Luu, C., Q.D. Bui, R. Costache, et al. 2021. "Flood-Prone Area Mapping Using Machine Learning Techniques: A Case Study of Quang Binh Province, Vietnam." Nature Hazards 108: 3229–51. 10.1007/s11069-021-04821-7.

[48] Luu, C., Q.D. Bui, R. Costache, L.T. Nguyen, T.T. Nguyen, T.V. Phong, H.V. Le, and B.T. Pham. 2021. "Flood-Prone Area Mapping Using Machine Learning Techniques: A Case Study of Quang Binh Province, Vietnam." Natural Hazards 108: 3229–51. 10.1007/s11069-021-04821-7.

[49] Luu, C., H. Ha, Q.D. Bui, N.-D. Luong, D.T. Khuc, H. Vu, and D.Q. Nguyen. 2023. "Flash Flood and Landslide Susceptibility Analysis for a Mountainous Roadway in Vietnam Using Spatial Modeling." Quaternary Science Advances 11: 100083. 10.1016/j.qsa.2023.100083.

[50] Lyu, H.-M., and Z.-Y. Yin. 2023. "Flood Susceptibility Prediction Using Tree-Based Machine Learning Models in the GBA." Sustainable Cities and Society 97: 104744. 10.1016/j.scs.2023.104744.

[51] Madhuri, R., S. Sistla, and K. Srinivasa Raju. 2021. "Application of Machine Learning Algorithms for Flood Susceptibility Assessment and Risk Management." Journal of Water and Climate Change 12(6): 2608–23. 10.2166/wcc.2021.051.

[52] Mahdizadeh Gharakhanlou, N., and L. Perez. 2022. "Spatial Prediction of Current and Future Flood Susceptibility: Examining the Implications of Changing Climates on Flood Susceptibility Using Machine Learning Models." Entropy 24: 1630. 10.3390/e24111630.

[53] Mehravar, S., S.V. Razavi-Termeh, A. Moghimi, B. Ranjgar, F. Foroughnia, and M. Amani. 2023. "Flood Susceptibility Mapping Using Multi-Temporal SAR Imagery and Novel Integration of Nature-Inspired Algorithms Into Support Vector Regression." Journal of Hydrology 617(Part C): 129100. 10.1016/j.jhydrol.2023.129100.

[54] Meliho, M., A. Khattabi, and J. Asinyo. 2021. "Spatial Modeling of Flood Susceptibility Using Machine Learning Algorithms." Arabian Journal of Geosciences 14: 2243. 10.1007/s12517-021-08610-1.

[55] Mirzaei, S., M. Vafakhah, B. Pradhan, et al. 2021. "Flood Susceptibility Assessment Using Extreme Gradient Boosting (EGB), Iran." Earth Science Informatics 14: 51–67. 10.1007/s12145-020-00530-0.

[56] Mishra, B.P., D.K. Ghose, and D.P. Satapathy. 2022. "Geospatial Modeling Using Hybrid Machine Learning Approach for Flood Susceptibility." Earth Science Informatics 15: 2619–36. 10.1007/s12145-022-00872-x.

[57] Motta, M., M. de Castro Neto, and P. Sarmento. 2021. "A Mixed Approach for Urban Flood Prediction Using Machine Learning and GIS." International Journal of Disaster Risk Reduction 56: 102154. 10.1016/j.ijdrr.2021.102154.

[58] Nachappa, T.G., S.T. Piralilou, K. Gholamnia, O. Ghorbanzadeh, O. Rahmati, and T. Blaschke. 2020. "Flood Susceptibility Mapping With Machine Learning, Multi-Criteria Decision Analysis and Ensemble Using Dempster Shafer Theory." Journal of Hydrology 590: 125275. 10.1016/j.jhydrol.2020.125275.

[59] Nayak, M., S. Das, and M.R. Senapati. 2022. "Improving Flood Prediction with Deep Learning Methods." Journal of the Institution of Engineers (India): Series B 103: 1189–205. 10.1007/s40031-022-00720-y.

[60] Ngo, P-T.T., Pham, T.D., Nhu, V-H., Le, T.T., Tran, D.A., Phan, D.C., Hoa, P.V., Amaro-Mellado, J.L., & Bui, D.T. (2021). A Novel Hybrid Quantum-PSO and Credal Decision Tree Ensemble for Tropical Cyclone Induced Flash Flood Susceptibility Mapping With Geospatial Data. Journal of Hydrology, 596, 125682. 10.1016/j.jhydrol.2020.125682.

[61] Nguyen, H.D. 2022. "GIS-Based Hybrid Machine Learning for Flood Susceptibility Prediction in the Nhat Le–Kien Giang Watershed, Vietnam." Earth Science Informatics 15: 2369–86. 10.1007/s12145-022-00825-4.

[62] Nguyen, H.D., C.P. Van, and A.D. Do. 2023. "Application of Hybrid Model-Based Deep Learning and Swarm-Based Optimizers for Flood Susceptibility Prediction in Binh Dinh Province, Vietnam." Earth Science Informatics 16: 1173–93. 10.1007/s12145-023-00954-4.

[63] Norallahi, M., and H. Seyed Kaboli. 2021. "Urban Flood Hazard Mapping Using Machine Learning Models: GARP, RF, MaxEnt and NB." Natural Hazards 106: 119–37. 10.1007/s11069-020-04453-3.

[64] Panahi, M., A. Jaafari, A. Shirzadi, H. Shahabi, O. Rahmati, E. Omidvar, S. Lee, and D.T. Bui. 2021. "Deep Learning Neural Networks for Spatially Explicit Prediction of Flash Flood Probability." Geoscience Frontiers 12: 101076. 10.1016/j.gsf.2020.09.007.

[65] Park, S.-J., and D.-K. Lee. 2020. "Prediction of Coastal Flooding Risk Under Climate Change Impacts in South Korea Using Machine Learning Algorithms." Environmental Research Letters 15(9): 94052. 10.1088/1748-9326/aba5b3.

[66] Pham, B.T., C. Luu, T.V. Phong, P.T. Trinh, A. Shirzadi, S. Renoud, S. Asadi, H.V. Le, J. von Meding, and J.J. Clague. 2021. "Can Deep Learning Algorithms Outperform Benchmark Machine Learning Algorithms in Flood Susceptibility Modeling?" Journal of Hydrology 592: 125615. 10.1016/j.jhydrol.2020.125615.

[67] Pradhan, B., S. Lee, A. Dikshit, and H. Kim. 2023. "Spatial Flood Susceptibility Mapping Using an Explainable Artificial Intelligence (XAI) Model." Geoscience Frontiers 14(6): 101625. 10.1016/j.gsf.2023.101625.

[68] Rahman, M., C. Ningsheng, M.M. Islam, A. Dewan, J. Iqbal, and T. Shufeng. 2019. "Flood Susceptibility Assessment in Bangladesh Using Machine Learning and Multi-Criteria Decision Analysis." Earth Systems and Environment 3: 585–601. 10.1007/s41748-019-00123-y.

[69] Razavi Termeh, S.V., A. Kornejady, H.R. Pourghasemi, and S. Keesstra. 2018. "Flood Susceptibility Mapping Using Novel Ensembles of Adaptive Neuro Fuzzy Inference System and Metaheuristic Algorithms." Science of the Total Environment 615: 438–51. 10.1016/J.SCITOTENV.2017.09.262.

[70] Razavi-Termeh, S.V., A. Sadeghi-Niaraki, and S.-M. Choi. 2023. "A New Approach Based on Biology-Inspired Metaheuristic Algorithms in Combination With Random Forest to Enhance the Flood Susceptibility Mapping." Journal of Environmental Management 345: 118790. 10.1016/j.jenvman.2023.118790.

[71] Razavi-Termeh, S.V., A. Sadeghi-Niaraki, M.B. Seo, and S.M. Choi. 2023. "Application of Genetic Algorithm in Optimization Parallel Ensemble-Based Machine Learning Algorithms to Flood Susceptibility Mapping Using Radar Satellite Imagery." Science of the Total Environment 873: 162285. 10.1016/j.scitotenv.2023.162285.

[72] Riazi, M., K. Khosravi, K. Shahedi, S. Ahmad, C. Jun, S.M. Bateni, and N. Kazakis. 2023. "Enhancing Flood Susceptibility Modeling Using Multi-Temporal SAR Images, CHIRPS Data, and Hybrid Machine Learning Algorithms." Science of the Total Environment 871: 162066. 10.1016/j.scitotenv.2023.162066.

[73] Saha, S., A. Gayen, and B. Bayen. 2022. "Deep Learning Algorithms to Develop Flood Susceptibility Map in Data-Scarce and Ungauged River Basin in India." Stochastic Environmental Research and Risk Assessment 36(6): 3295–310. 10.1007/s00477-022-02195-1.

[74] Sahoo, A., S. Samantaray, and S. Paul. 2021. "Efficacy of ANFIS-GOA Technique in Flood Prediction: A Case Study of Mahanadi River Basin in India." H2Open Journal 4(1): 137–56. 10.2166/h2oj.2021.090.

[75] Sajedi Hosseini, F., B. Choubin, A. Mosavi, N. Nabipour, S. Shamshirband, H. Darabi, and A. Torabi Haghighi. 2020. "Flash-Flood Hazard Assessment Using Ensembles and Bayesian-Based Machine Learning Models: Application of the Simulated Annealing Feature Selection Method." Science of the Total Environment 711: 135161. 10.1016/j.scitotenv.2019.135161.

[76] Sankaranarayanan, S., M. Prabhakar, S. Satish, P. Jain, A. Ramprasad, and A. Krishnan. 2020. "Flood Prediction Based on Weather Parameters Using Deep Learning." Journal of Water and Climate Change 11(4): 1766–83. 10.2166/wcc.2019.321.

[77] Saravanan, S., D. Abijith, N.M. Reddy, and KSS, P., Janardhanam, N., Sathiyamurthi, S., & Sivakumar, V. 2023. "Flood Susceptibility Mapping Using Machine Learning Boosting Algorithms Techniques in Idukki District of Kerala India." Urban Climate 49: 101503. 10.1016/j.uclim.2023.101503.

[78] Shafizadeh-Moghadam, H., R. Valavi, H. Shahabi, K. Chapi, and A. Shirzadi. 2018. "Novel Forecasting Approaches Using Combination of Machine Learning and Statistical Models for Flood Susceptibility Mapping." Journal of Environmental Management 217: 1–11. 10.1016/j.jenvman.2018.03.089.

[79] Shahabi, H., A. Shirzadi, K. Ghaderi, E. Omidvar, N. Al-Ansari, J.J. Clague, M. Geertsema, et al. 2020. "Flood Detection and Susceptibility Mapping Using Sentinel-1 Remote Sensing Data and a Machine Learning Approach: Hybrid Intelligence of Bagging Ensemble Based on K-Nearest Neighbor Classifier." Remote Sensing 12(2): 266. 10.3390/rs12020266.

[80] Shahabi, H., A. Shirzadi, S. Ronoud, S. Asadi, B.T. Pham, F. Mansouripour, M. Geertsema, J.J. Clague, and D.T. Bui. 2021. "Flash Flood Susceptibility Mapping Using a Novel Deep Learning Model Based on Deep Belief Network, Back Propagation and Genetic Algorithm." Geoscience Frontiers 12(3): 101100. 10.1016/j.gsf.2020.10.007.

[81] Tabbussum, R., and A.Q. Dar. 2021. "Performance Evaluation of Artificial Intelligence Paradigms—Artificial Neural Networks, Fuzzy Logic, and Adaptive Neuro-Fuzzy Inference System for Flood Prediction." Environmental Science and Pollution Research 28: 25265–82. 10.1007/s11356-021-12410-1.

[82] Tehrany, M.S., S. Jones, and F. Shabani. 2019. "Identifying the Essential Flood Conditioning Factors for Flood Prone Area Mapping Using Machine Learning Techniques." CATENA 175: 174–92. 10.1016/j.catena.2018.12.011.

[83] Tehrany, M.S., B. Pradhan, and M.N. Jebur. 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. 10.1016/j.jhydrol.2013.09.034.

[84] Tehrany, M.S., B. Pradhan, and M.N. Jebur. 2015. "Flood Susceptibility Analysis and its Verification Using A Novel Ensemble Support Vector Machine and Frequency Ratio Method." Stochastic Environmental Research and Risk Assessment 29: 1149–65. 10.1007/s00477-015-1021-9.

[85] Tehrany, M.S., B. Pradhan, and M.N. Jebur. 2014. "Flood Susceptibility Mapping Using a Novel Ensemble Weights-of-Evidence and Support Vector Machine Models in GIS." Journal of Hydrology 512: 332–43. 10.1016/j.jhydrol.2014.03.008.

[86] Tehrany, M.S., B. Pradhan, S. Mansor, and N. Ahmad. 2015. "Flood Susceptibility Assessment Using GIS-Based Support Vector Machine Model With Different Kernel Types." CATENA 125: 91–101. 10.1016/j.catena.2014.10.017.

[87] Ullah, K., Y. Wang, Z. Fang, L. Wang, and M. Rahman. 2022. "Multi-Hazard Susceptibility Mapping Based on Convolutional Neural Networks." Geoscience Frontiers 13(5): 101425. 10.1016/j.gsf.2022.101425.

[88] Vafakhah, M., Mohammad Hasani Loor, et al. 2020. "Comparing Performance of Random Forest and Adaptive Neuro-Fuzzy Inference System Data Mining Models for Flood Susceptibility Mapping." Arabian Journal of Geosciences 13: 417. 10.1007/s12517-020-05363-1.

[89] Wang, Y., H. Hong, W. Chen, S. Li, M. Panahi, K. Khosravi, A. Shirzadi, H. Shahabi, S. Panahi, and R. Costache. 2019. "Flood Susceptibility Mapping in Dingnan County (China) Using Adaptive Neuro-Fuzzy Inference System With Biogeography Based Optimization and Imperialistic Competitive Algorithm." Journal of Environmental Management 247: 712–29. 10.1016/j.jenvman.2019.06.102.

[90] Xie, S., W. Wu, S. Mooser, Q.J. Wang, R. Nathan, and Y. Huang. 2021. "Artificial Neural Network Based Hybrid Modeling Approach for Flood Inundation Modeling." Journal of Hydrology 592: 125605. 10.1016/j.jhydrol.2020.125605.

[91] Xu, K., Han, Z., Xu, H., et al. (2023). Rapid Prediction Model for Urban Floods Based on a Light Gradient Boosting Machine Approach and Hydrological–Hydraulic Model. International Journal of Disaster Risk Science, 14, 79–97. 10.1007/s13753-023-00465-2.

[92] Yang, D., T. Zhang, A. Arabameri, M. Santosh, U.D. Saha, and A. Islam. 2023. "Flash-Flood Susceptibility Mapping: A Novel Credal Decision Tree-Based Ensemble Approaches." Earth Science Informatics. 10.1007/s12145-023-01057-w.

[93] Yaseen, A., J. Lu, and X. Chen. 2022. "Flood Susceptibility Mapping in an Arid Region of Pakistan Through Ensemble Machine Learning Model." Stochastic Environmental Research and Risk Assessment 36: 3041–61. 10.1007/s00477-022-02179-1.

[94] Youssef, A.M., A.M. Mahdi, and H.R. Pourghasemi. 2023. "Optimal Flood Susceptibility Model Based on Performance Comparisons of LR, EGB, and RF Algorithms." Natural Hazards 115: 1071–96. 10.1007/s11069-022-05584-5.

[95] Youssef, A.M., H.R. Pourghasemi, and B.A. El-Haddad. 2022. "Advanced Machine Learning Algorithms for Flood Susceptibility Modeling—Performance Comparison: Red Sea, Egypt." Environmental Science and Pollution Research 29: 66768–92. 10.1007/s11356-022-20213-1.

[96] Youssef, A.M., H.R. Pourghasemi, A.M. Mahdi, et al. 2023. "Flood Vulnerability Mapping and Urban Sprawl Suitability Using FR, LR, and SVM Models." Environmental Science and Pollution Research 30: 16081–105. 10.1007/s11356-022-23140-3.

[97] Yu, H., Z. Luo, L. Wang, X. Ding, and S. Wang. 2023. "Improving the Accuracy of Flood Susceptibility Prediction by Combining Machine Learning Models and the Expanded Flood Inventory Data." Remote Sensing 15: 3601. 10.3390/rs15143601.

[98] Zhang, L., H. Qin, J. Mao, X. Cao, and G. Fu. 2023. "High Temporal Resolution Urban Flood Prediction Using Attention-Based LSTM Models." Journal of Hydrology 620(Part B): 129499. 10.1016/j.jhydrol.2023.129499.

[99] Zhao, G., B. Pang, Z. Xu, D. Peng, and L. Xu. 2019. "Assessment of Urban Flood Susceptibility Using Semi-Supervised Machine Learning Model." Science of the Total Environment 659: 940–9. 10.1016/j.scitotenv.2018.12.217.

[100] Zhao, G., B. Pang, Z. Xu, D. Peng, and D. Zuo. 2020. "Urban Flood Susceptibility Assessment Based on Convolutional Neural Networks." Journal of Hydrology 590: 125235. 10.1016/j.jhydrol.2020.125235.

Appendix 3

Details of the Reviewed Papers

No.	Author (year)	Used ML model		Catchment (size in km²)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
1	Eslaminezhad et al. (2022)	Hybrid	ROF-BPSO ADTree-BPSO RF-BPSO	Maneh and Samalqan watershed, Iran (6053)	CN; HOFD; MFI; VOFD; NDVI; TWI; SPI; TRI; TPI; DEM; Slope; FloAcc; FlowDir; Plan-Curvature; DisFau; DisRoa; DisRiv; Lithology; Soil type; LULC	30 m × 30 m	2016–2019 (370 flood locations)
2	Chapi et al. (2017)	Hybrid, Statistical (Regression), Tree based, Ensemble (Tree)	Bagging-LMT, LMT, LR, BLR, RF	Haraz watershed in northern Iran (4014)	Slope; Elevation; Plan-Curvature; TWI; SPI; Riv-Den; DisRiv; LULC; NDVI; Lithology; Rainfall	30 m × 30 m	N.A. (210 flood locations)
3	Al-Areeq et al. (2022)	Ensemble (Bagging), Kernel based, Instance based, Tree based	BE, LT, k-SVM, KNN	Jeddah City, Saudi Arabia (1821)	Slope; TPI; SPI; Plan-Curvature; TWI; DisRiv; Rainfall; Lithology; LULC; Soil type; Convergence; FloAcc; Elevation; TRI; Aspect	30 m × 30 m	N.A. (282 flood locations)
4	Mahdizadeh Gharakhanlou and Perez (2022)	Ensemble, Tree based, Statistical, Kernel based	GBM, RF, MLP-NN, NB	Loup watershed and Lower Nicola River in Canada (N.A.)	Elevation; Slope; Aspect; Plan-Curvature; Prof-Curvature; Roughness; TWI; LULC; Precipitation; DisRiv; Drain-Den; Lithology; Soil type; SPI; NDVI; NDMI	30 m × 30 m	03 May 2017 and 14 November 2021 (120 flood locations)
5	Ullah et al. (2022)	Kernel based, Statistical (Regression), Instance based	CNN, logistic regression, KNN	Shangla District, eastern Hindu Kush region of Pakistan (1586)	Slope; Elevation; Plan-Curvature; Prof-Curvature; Geology; DisFau; DisRoa; DisRiv; SPI; STI; TWI; NDVI; LULC; Rainfall; Aspect	12.5 m × 12.5 m	2000–2020 (N.A.)
6	Park and Lee (2020)	Instance based	KNN	Coastal areas of South Korea (33°–38° N, 125°–131° E) (N.A.)	Tide; Rainfall; Elevation; Slope; LULC	1 km² × 1 km²	2002–2014 (N.A.)
7	Abu El-Magd et al. (2022)	Hybrid	Naïve Bayes- hydrologic indices	Wadi Ghoweiba and the northern part of EED in Egypt (3258)	STI; SPI; Lithology; TWI; Slope; Riv-Den; Plan-Curvature; Slope Aspect	30 m × 30 m	N.A. (189)
8	Luu et al. (2021)	Tree based, Hybrid	ADT, LMT,R-EPT, J48-DT, NB-T	Quang Binh province, Vietnam (N.A.)	Elevation; Slope; Plan-Curvature; DisRiv; Riv-Den; Rainfall; FlowDir; FloAcc; Geology; LULC	30 m × 30 m	Three floods in 2007, 2010, and 2016 (321 flood locations)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
9	Norallahi and Kaboli (2021)	Ensemble, EA, statistical	GA Rule-Set Production, Maximum Entropy, RF, NB	Kermanshah city in Iran (111)	Precipitation; LULC; CN; Elevation; Slope; DisRiv	12.5 m × 12.5 m	September 2018 to September 2019 (280 locations)
10	Ha et al. (2023)	Ensemble, statistical (Regression), Hybrid	AdaBoost, Logistic Regression, AHP-AdaBoost-Logistic Regression	Quang Binh province (8065.27)	Elevation; Slope; Plan-Curvature; Elevation; FloAcc; Rainfall Intensity; SPI; TWI; Lithology; Riv-Den; DisRiv; LULC; Rainfall; NDVI	12.5 m × 12.5 m	Historical floods in 2007, 2010, 2016, and 2020 (671 flood locations)
11	Madhuri et al. (2021)	Statistical (Regression), Kernel based, Instance based, Ensemble (Adaptive Boosting), Hybrid (Tree based and Ensemble)	Logistic Regression, SVM, KNN, Adaptive Boosting, XGBoost	Greater Hyderabad Municipal orporation (GHMC) in India (625)	Rainfall; Elevation; Slope; DisRiv; Evapotranspiration;LST; NDVI; CN	30 m × 30 m	2000, 2006 and 2016 (295 flood locations)
12	Luu et al. (2023)	Hybrid	AdaBoost-RBF, Bagging-RBF, MultiBoostAB-RBF, Random Sub-spaceRBF	Vietnam's National Highway 6 (NH6) (115 km × 3 km)	Elevation; Slope; Aspect; Plan-Curvature; TWI; SPI; STI; Elevation; Rainfall; Drain-Den; LULC; Lithology; Geomorphology; Weathering crust; Structural zone	12.5 m × 12.5 m	2005–2020 (88 flood locations)
13	Riazi et al. (2023)	Kernel based, Ensemble, Hybrid (Used for feature selection)	RBF, BA-RBF, RC-RBF, RSS-RBF, MI-T	The coastal Goorganrood watershed, Iran (11504)	Elevation; Lithology; Drain-Den; Rainfall; NDVI; Plan-Curvature; Slope; SPI; TWI; Soil type; LULC; DisRiv	30 m × 30 m	Floods in 2012 and 2019 (334 flood locatins)
14	Arabameri et al. (2020)	Ensemble, Tree based	BFT, DFT, RF-FT, FT	Gorgan Basin watershed in North of Iran (11290)	Elevation; Slope; Plan-Curvature; Prof-Curvature; TPI; TRI; TWI; Convergence; SPI; Drain-Den; DisRiv; Rainfall; Lithology; LULC; NDVI; Soil type	12.5 m × 12.5 m	2001 to 2019 (426 flood locations)
15	Nguyen et al. (2023)	Neural Network—Deep Learning	DNN, DNN-AO, DNN-SLnO, DNN-EHO, DNN-NMRA, DNN-SGD	Binh Dinh located in the South Central Coast region of Vietnam (6071)	Elevation; Aspect; Slope; Plan-Curvature; NDVI; NDBI; Riv-Den; Road density; NDWI; FlowDir; LULC; Rainfall; Soil type; Geology	12.5 m × 12.5 m	Floods in 2008, 2010, 2013, 2016, 2017 and 2020 (1061 flood and 822 non-flood points)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
16	Zhang et al. (2023)	Hybrid	ALSTM-DW	Shenzhen, in the south of Guangdong Province, China (N.A.)	Rainfall; Water depth	N.A.	July 2018 to July 2019 (three flooded sites)
17	Aldiansyah et al. (2023)	Statistical (regression), Kernel-based, statistical	GLM, SVM, RF, BRT, MARS, MDA, FDA	Kendari City, Indonesia (274.91)	Aspect; Plan-Curvature; Elevation; FloAcc; FlowDir; Geology; Slope; LULC; NDVI; Rainfall; DisRiv; Soil type; SPI; STI; TRI; TWI; Wind	30 m × 30 m	2013 (23 flood and 28 non-flood locations)
18	Islam et al. (2023)	Neural network, Ensemble (Tree based), Hybrid	ANN, FL, RF, ANN-FL, FL-RF, RF-ANN	The Brahmaputra River Basin (BRB), Bangladesh (583000)	DEM; Slope; FlowDir; Prof-Curvature; Plan-Curvature; FloAcc; SPI; STI; Population; DisRiv; Rainfall; DisRoa; Aspect; NDWI; NDVI; Soil type	10 m × 10 m	Events in 1988, 1998, 2004, 2010, and 2017 (1000 flood and 1000 non-flood points)
19	Yang et al. (2023)	Ensemble (Tree based)	CDT, CDT-ADTree, CDT-REPT, CDT-RF, CDT-FT, CDT-NBTree	The Neka-roud watershed in the Mazandran province, north of Iran (3768.33)	Elevation; Slope; Plan-Curvature; Rainfall; DisRiv; TWI; SPI; NDVI; LULC; Lithology; DisRes; TR; MBR; CC; FF; ER; CCM; Infiltration	12.5 m × 12.5 m	2001–2019 (206 flood locations)
20	Vafakhah et al. (2020)	Statistical (MCDA), Hybrid, Ensemble (tree based)	FR, ANFIS, RF	Gilan Province at the southern margin of the Caspian Sea, Iran (14100)	Slope; Aspect; Elevation; DisRiv; Drain-Den; Lithology; LULC; TWI; SPI	30 m × 30 m	N.A. (220 flood locations)
21	Xie et al. (2021)	Neural network, Hybrid	TPB-ANN, BB-ANN (simple and complex), Hybrid BB-ANN	The Burnett River in Queensland, Australia (7.779 downstream the Paradise Dam)	Water depth	20 m × 20 m	Three historical events (i.e., the 1971, 2010 and 2013) (Time series prediction of flood inundation)
22	Zhao et al. (2020)	Neural Network, Kernel-based, Ensemble (Tree based)	SCNN, LeNet-5, SVM, RF	The Dahongmen catchment in Beijing, China (131)	Precipitation; FP; Elevation; Slope; TWI; DisRiv; DisRoa; Drain-Den; NDBI	N.A.	Between 2004 and 2014 (202 flooded sites and 160 non-flooded sites)
23	Tehrany et al. (2015)	Kernel-based, Statistical model (MCDA)	SVM with 4 kernel functions, FR	Kuala Terengganu basin, Malaysia (1700)	Slope; SPI; TWI; DisRiv; Geology; LULC; Soil type; Elevation; Plan-Curvature; Runoff	15 m × 15 m	27th November 2009 (181 flood locations)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
24	Tehrany et al. (2014)	Statistical (MCDA), Kernel-based, Ensemble	WoE, SVM, WoE-SVM	Terengganu, Malaysia (1804)	Slope; SPI; TWI; Elevation; Curvature; DisRiv; Geology; Rainfall; LULC; Soil type	15 m × 15 m	27th November 2009 (180 flood locations)
25	Pham et al. (2021)	Hybrid, Evolutionary Algorithm, Neural Network, Statistical (Regression), Tree-based	DBN (ELM-DL), PSO, ANNRBF, LR, LMTree, FTree, ADTree	Vu Gia-Thu Bon watershed, central Vietnam (20284)	Slope; Elevation; Aspect; Curvature; Prof-Curvature; Plan-Curvature; TWI; SPI; DisRiv; Riv-Den; FlowDir; FloAcc; Rainfall; Lithology; LULC	30 m × 30 m	Floods in 2007, 2009, and 2013 (847 flood location)
26	Khosravi et al. (2020)	Neural networks	CNN	Iran (1648195)	Slope; Elevation; Aspect; Plan-Curvature; Prof-Curvature; Rainfall; Geology; LULC; DisRoa; DisRiv	30 m × 30 m	N.A. (2769 flood and 2769 non-flood locations)
27	Kabir et al. (2020)	Kernel based, Neural network	CNN, SVR	Eden Catchment in the Northwest England (2500)	Hydrometric (flow hydrographs); DEM; Flow rate (at different time steps)	5 m × 5 m	Two flood events in 2005 and 2015 (whole area)
28	Costache et al. (2022)	Ensemble	U-WOE, SGD-WOE, CSForest-WOE	Putna river basin in Romania (2509)	Slope; TPI; Elevation; Convergence; Plan-Curvature; SPI; Prof-Curvature; TWI; Aspect; LULC; Lithology; DisRiv; Soil type; Rainfall	30 m × 30 m	1980–2015 and 2018 (132 flood and 132 non-flood locations)
29	Chen et al. (2022)	Hybrid	Convolutional LSTM	Xi County basin located to Xinyang City, Henan Province, China (1715.79)	Rainfall; Flow rate	N.A.	January 1, 2010 to December 31, 2018 (time series prediction for whole catchment)
30	Kao et al. (2020)	Hybrid	LSTM-ED, FFNN-ED	The Shihmen Reservoir basin in northern Taiwan (763.4)	Flow rate; Rainfall	N.A.	2007 to 2016 (23 typhoon events-Time series prediction)
31	Bui et al. (2020)	Hybrid, Evolutionary Algorithm, Kernel-based, Ensemble	DLNN-SSOA, DLNN-GWO, DLNN-GOA, DLNN-PSO, DLNN-GD, SVM, RF	Lai Chau, northwest mountainous region of Vietnam (latitudes 21°51′ and 22°49′, and longitudes 102°19′ to 103°59′)	DEM; Aspect; Slope; Curvature; TWI; Rainfall; Riv-Den; SPI; DisRiv; NDVI; NDBI	30 m × 30 m	Severe flash flood in 2018 (2406 flood and 2406 non-flood locations)
32	Razavi Termeh et al. (2018)	Hybrid, statistical	ANFIS-ACO, ANFIS-GA, ANFIS-PSO, FR	The Jahrom Basin in southern Fars Province, Iran (5737)	Slope; Elevation; Plan-Curvature; Rainfall; DisRiv; LULC; Lithology; SPI; TWI	10 m × 10 m	2001–2012 (53 flood locations)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
33	Chen et al. (2019)	Tree based, Ensemble	REPTree, Bag-REPTree, RS-REPTree	Quannan County in Ganzhou City, Jiangxi Province, China (1520)	Elevation; Slope; Aspect; Curvature; SPI; STI; TWI; DisRiv; NDVI; Soil type; LULC; Lithology; Rainfall	30 m × 30 m	N.A. (363 flood and 363 non-flood locations)
34	Ngo et al. (2021)	Hybrid, Kernel based, Tree based, Statistical (Regression)	QPSO, CDT, SVM, CART, LR, BFTree	The Tran Yen district in Vietnam (629)	Slope; Aspect; Elevation; Curvature; TWI; Rainfall; Riv-Den; Soil type; Lithology; LULC	30 m × 30 m	From 2016 to 2019 (1698 flooded locations)
35	Bui et al. (2016)	Hybrid, Tree-Based, Ensemble Model (Tree-Based), Neural Network, Kernel-Based	MONF-(EG + PSO), J48 DT, Random Forest, MLP NN, SVM, ANFIS	Tuong Duong district in Central Vietnam (2803.1)	Slope; Elevation; Curvature; TWI; SPI; DisRiv; Riv-Den; NDVI; Lithology; Rainfall	20 m × 20 m	From 2010 to 2014 (76 historical flood inundated areas)
36	Khosravi et al. (2018)	Hybrid, tree based	LMT, REPT, NBT, ADT	Haraz Watershed in the northern part of Iran (4014)	Slope; Elevation; Curvature; SPI; TWI; LULC; Rainfall; Riv-Den; DisRiv; Lithology; NDVI	30 m × 30 m	From 1995 to 2015 (201 flood and 201 non-flood locations)
37	Choubin et al. (2019)	Statistical model, Tree-based, Kernel-based, Ensemble model	MDA, CART, SVM, Ensemble	Khiyav-Chai watershed in Iran (126)	Elevation; Slope; Aspect; Curvature; DisRiv; TWI; Drain-Den; Soil depth; SHG; LULC; Lithology	30 m × 30 m	2010–2017 (51 flood and 51 non-flood locations)
38	Wang et al. (2019)	Hybrid	ANFIS, ANFIS-BBO, ANFIS-ICA	Dingnan County-China (1318)	Slope; Elevation; Aspect; Curvature; TWI; SPI; STI; DisRiv; LULC; NDVI; Lithology; Rainfall; Soil	30 m × 30 m	2006–2016 (115 flood occurrences in the study area)
39	Shafapour Tehrani et al. (2015)	Hybrid, Tree based	FR-SVM, DT	The upper catchment of Kelantan, Malaysia (923)	Elevation; Curvature; Geology; Riv-Den; SPI; Rainfall; LULC; Soil; TWI; Slope	15 m × 15 m	The flood event of November 2005 (270 flood and non-flood areas)
40	Costache et al. (2021)	Kernel-based, Tree-based, Hybrid (Fuzzy Logic and Neural Network), Ensemble (Tree-based), Neural Network	SVM, J48 DT, ANFIS, RF, ANN, ADT	The river basin of Buzau in Romania (5350)	Slope; Elevation; Aspect; Convergence; LULC; DisRiv; Soil type; Lithology; Plan-Curvature; Rainfall; TPI; TWI	30 m × 30 m	1990–2020 (205 flood and 205 non-flood locations)
41	Hong et al. (2018)	Hybrid	ANFIS-GA, ANFIS-DE	Hengfeng County located in Jiangxi Province, China (655)	Slope; Aspect; Elevation; Curvature; STI; SPI; TWI; Rainfall; DisRiv; Lithology; Soil; LULC; NDVI	30 m × 30 m	The events occurred in Jun 1964, Jun 1971, Aug 1997, Jun 1998 and Jun 1999 (195 total flood events)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
42	Tehrany et al. (2013)	Tree based, Hybrid	Rule-Based DT, FR + LR	Kelantan River Basin in North East part of Peninsular Malaysia (923)	DEM; Curvature; Geology; Riv-Den; SPI; Rainfall; LULC; Soil; TWI; Slope	15 m × 15 m	The flood event of November 2005 (115 flood and 115 non-flooded locations)
43	Gudiyangada Nachappa et al. (2020)	Statistical model (MCDA), Ensemble, Kernel-based, Hybrid	AHP, ANP, RF, SVM, AHP-ANP, RF-SVM, Ensemble of all above models	Salzburg, Austria (7156.03)	Elevation; Slope; Aspect; TWI; SPI; DisRiv; Rainfall; Geology; LULC; NDVI; DisRoa	30 m × 30 m	2002, 2005, 2006, and 2013 (N.A.)
44	Abu El-Magd (2022)	Statistical, Ensemble	NB, RF	Wadi El-Dib on the Gulf of Suez Coast, Egypt (1390)	DEM; Slope; Curvature; Lithology; DisRiv; Riv-Den; TWI	30 m × 30 m	N.A. (1117 flood locations)
45	Abu El-Magd et al. (2021)	Ensemble (Tree-Based), Instance-Based	XGBoost, KNN	Wadi El-Laqeita in the Central Eastern Desert of Egypt (7186)	Elevation; Slope; Curvature; Slope Aspect; Lithology; DisRiv; Riv-Den; TWI	30 m × 30 m	N.A. (355 flood location)
46	Al-Juaidi et al. (2018)	Statistical (Regression based)	LR	The southern Gaza Strip on the eastern coast of the Mediterranean Sea (365)	DEM; Slope; FloAcc; Rainfall; LULC; Soil	20 m × 20 m	The past floods of 1989, 2003, and 2013 (140 flood locations)
47	Aydin and Iban (2022)	Ensemble (tree based)	NGBoost, LightGBM, CatBoost, RF, GB, XGBoost, AdaBoost	Adana province on the Mediterranean coast of Türkiye (14030)	Elevation; Slope; Aspect; Plan-Curvature; Prof-Curvature; TRI; TWI; SPI; Rainfall; Lithology; NDVI; LULC; DisRiv; DisRoa	30 m × 30 m	2000–2022 (227 flood, 227 non-flood locations)
48	Chakrabortty et al. (2021)	Neural network, Ensemble	ANN, DLNN, DB	The Dwarekeswar River basin, west Bengal (4673)	Aspect; Elevation; Plan-Curvature; Drain-Den; Prof-Curvature; Geology; Geomorphology; DisRiv; LULC; LST; NDVI; R factor; Rainfall; SPI; TWI	12.5 m × 12.5 m	N.A.
49	Deroliya et al. (2022)	Tree based, Ensemble	DT, RF, GBDT	Jagatsinghpur, located in the lower Mahanadi River basin, India (1668)	DTM, MRVBF, VDCN, Sl, κh, κp, κt, CI, Al, Ar, hr., hl, D, H, TI, and DI	20 m × 20 m	N.A. (1000 locations)
50	Ding et al. (2020)	Hybrid (RNN + Attention), Statistical, Neural network	STA-LSTM, HA, FCN, CNN, GCN, LSTM, SA-LSTM, TA-LSTM	Three small and medium basins in China: Tunxi basin (2696.76), Changhua basin (3444), and Heihe basin (1350)	Rainfall; Flow rate	N.A. (Time series prediction)	From June 27, 1981 to March 18, 2007 (49,532 samples for Tunxi, 9354 samples for Changhua, 5423 samples for Heihe)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
51	El-Haddad et al. (2020)	Ensemble, statistical (MCDA and Regression)	BRT, FDA, GLM, MDA	Wadi Qena Basin in Egypt (14558)	Slope; Slope Length; Elevation; LULC; Lithological Units; Curvature; Slope Aspect; TWI	30 m × 30 m	2014–2016, and 2018 (342 flood locations)
52	Fang et al. (2021)	Hybrid	LSS-LSTM	Shangyou County, China (1543)	Elevation; Aspect; Curvature; Slope; SPI; STI; TWI; Lithology; LULC; NDVI; Soil; DisRiv; Rainfall	30 m × 30 m	Events that occurred from July 26 to 29, 2006 (108 flood and 108 non-flood locations)
53	Al-Abadi et al. (2018)	Ensemble	Random Forest, Rotation Forest, AdaBoost	The east of Wasit Governorate along the border between Iraq and Iran (13527)	Elevation; Slope; Plan-Curvature; TWI; SPI; DisRiv; Drain-Den; Lithology; Soil; LULC	30 m × 30 m	The flood of 7 May 2013 (3540 flood and non-flood locations)
54	Gude et al. (2020)	Neural network, statistical	LSTM, ARIMA	Meramec River in Valley Park, Missouri (Size not provided)	Flood stage (gauge height data)	N.A. (Time series prediction)	From 15 May 2016 5 PM onward until 1 September 2019 4 PM (113,994 samples)
55	Guo et al. (2022)	Kernel based	CNN	Canton of Zurich, Switzerland (90 × 65) and the cities of Lausanne and Geneva, Switzerland	Elevation; Water depth; Flow velocity	2 m × 2 m	N.A.
56	Yaseen et al. (2022)	Ensemble, statistical (regression), kernel-based, neural network	LR-SVM-MLP, LR, SVM, MLP	Karachi, Pakistan (built-up area increased from 466.5 in 1998 to 666.18 in 2018)	Elevation; Slope; Curvature; SPI; TWI; Rainfall; Lithology; Soil type; DisRiv; Riv-Den; NDVI; LULC	30 m × 30 m	The flood of 2020 (652 flood and 652 nonflood locations)
57	Islam et al. (2021)	Hybrid, Kernel-based, Neural network, Ensemble	DE-ANN, DE-RF, DE-SVM, RS-ANN, RS-RF, RS-SVM, ANN, RF, SVM	The Teesta sub-catchment of the northern region of Bangladesh (2284)	Elevation; Curvature; Aspect; Slope; TRI; TWI; SPI; STI; LULC; DisRiv; Soil type; Rainfall	30 m × 30 m	2009–2021 (167 flood and 167 non-flood locations)
58	Janizadeh et al. (2021)	Hybrid (tree-based and Bayesian), Statistical, Ensemble (tree-based)	BART, NB, RF	The Kan watershed	Elevation; Aspect; Slope; Plan-Curvature; Prof-Curvature; Drain-Den; DisRiv; DisRoa; SPI; TPI; TPI; CN; LULC; Lithology; Rainfall	12.5 m × 12.5 m	1954–1996 and 5–9 October 2019 (118 flood and 118 non-flood locations)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
59	Kabir et al. (2020)	Neural network, Kernel based, Physical model	CNN, SVR, LISFLOOD-FP	The City of Carlisle, UK (14.5)	Water depth	5 m × 5 m (15-min temporal resolution)	The flood events in January 2005 and December 2015 (500 locations)
60	Khan et al. (2018)	Hybrid	Fuzzy Neural Network	Bow River in Calgary, Canada (25123)	Flow rate; Precipitation; Temperature	N.A. (Time series prediction)	1 January 2000–31 December 2010
61	Lawal et al. (2021)	Tree-based, statistical (regression), kernel-based	DT, LR, SVC	Kebbi State Nigeria (36800)	Rainfall	N.A. (Time series prediction)	Historical rainfall data from 1st January 1981 to 31st December 2013
62	Lei et al. (2021)	Neural network (deep learning)	CNN, RNN	Seoul, South Korea (605)	Elevation; Slope; SlopeF; TWI; Curvature; SPI; TPI; DisRiv; TRI; LULC	2 m × 2 m	2018–2020 (295 flooded sites)
63	Linh et al. (2021)	Hybrid, neural network-based, statistical (regression)	WNN, ANN, MLR	The Madarsoo watershed in upstream of Golestan dam, Iran (5155)	SST; SLP	2°	1974–2011 for March and August
64	Luu et al. (2021)	Tree based, Hybrid	AD Tree, LM Tree, REP Tree, J48, NB Tree	Quang Binh province, Vietnam (806526)	Rainfall; Riv-Den; DisRiv; LULC; Geology; Elevation; FlowDir; FloAcc; Slope; Curvature	30 m × 30 m	The three flood events in 2007, 2010, and 2016 (321 flooded locations)
65	Pradhan et al. (2023)	Hybrid	CNN-SHAP	Jinju Province, South Korea (7.6)	LULC; Elevation; Soil depth; Soil drain; Surface soil texture; Forest age class; Deep soil texture); Timber diameter); Tree type; Forest density; Topography; Lithology	N.A.	2018–2019 (582 historical flood events)
66	Meliho et al. (2021)	Ensemble, hybrid (Tree-Boosting), instance-based, neural network	RF, XGB, KNN, ANN	The Souss watershed located in southern Morocco (17000)	Aspect; Curvature; DEM; DisRiv; Drain-Den; FloAcc; FlowDir; Geology; LULC; Rainfall; Slope; Soil type; TWI	30 m × 30 m	1956, 1986–2019 (87 flood and 87 non-flood locations)
67	Mirzaei et al. (2020)	Hybrid, Statistical (MDCA and Regression), Ensemble (Tree based)	EGB, FR, RF, GAM	The Talar watershed, Iran (1765)	Elevation; Slope; Prof-Curvature; TWI; DisRiv; NDVI; Plan-Curvature; Rainfall; LULC; SPI; Lithology	30 m × 30 m	14 February 2018 and 16 July 2019 (243 flood locations)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
68	Mishra et al. (2022)	Hybrid	PLS-SEM-SVM, PLS-SEM-MLPNN, PLS-SEM-KNN, PLS-SEM-RBFN	The Basantpur watershed in the state of Chhattisgarh, India (31165)	Slope; Elevation; Curvature; LULC; Soil; DisRiv; SPI; Aspect; TWI; STI; Riv-Den; TRI; Lithology; NDVI	2.5 m × 2.5 m	2005, 2016, 2018 and 2019 (151 locations)
69	Motta et al. (2021)	Statistical (Regression), Kernel based, Ensemble, Instance based, Neural Network, Statistical	LR, SVM, Gaussian NB, RF, KNN, MLP	The area of Lisbon municipality, Portugal (100)	Temperature; Humidity; Sun exposure; Precipitation; Wind	100 m × 100 m	January 2013 until December 2018 (Whole area)
70	Nayak et al. (2022)	Neural Network, EA	DBN, TLBO	Daya and Bhargavi river, India	Flow velocity; Area; Rainfall; Wind; Pressure; Temperature	N.A. (Time series prediction)	1st July 2012 to 31st October 2013 (a total of 246 samples)
71	Nguyen (2022)	Tree based, Ensemble, Hybrid (NN + EA)	DT, ADB, MLP-WCA, MLP-AOA, MLP-WOA, MLP-ACO	The Nhat Le–Kien Giang watershed in Vietnam (2650)	Aspect; Plan-Curvature; Slope; TWI; NDVI; NDBI; NDWI; Rainfall; LULC; Geology; DisRoa; DisRiv; FlowDir	10 m × 10 m	Flood marks in 2010, 2013, 2016 and 18 October 2020 (1964 flood and non-flood locations)
72	Panahi et al. (2021)	Neural Network	CNN, RNN	The Golestan Province in the northeast of Iran (12000)	Slope; Aspect; Elevation; Plan-Curvature; TWI; DisRiv; Rainfall; LULC; Lithology	30 m × 30 m	2009–2019 (143 flood and 143 non-flood locations)
73	Rahman et al. (2019)	Neutal Network, Statistical (MCDA and Regression), Hybrid	ANN, LR, FR, AHP, Integrated LR-FR	Bangladesh (147570)	Rainfall; DEM; LULC; Geology; Soil; Drain-Den	300 m × 300 m	Flood events in 1988, 1995, and 1998 and on June 24, July 17, August 15, and August 24 of 2017 (475 flood locations)
74	Saha et al. (2022)	Kernel-based, Neural Network, Ensemble, Tree-based	CNN, MLP, Bagging, RF	The Kunur River Basin in Rarh Bengal (646)	Elevation; Slope; Aspect; CN; SPI; TPI; LULC; TWI; Ruggedness; Convergence; NDVI; DisRiv; Rainfall; Geology; Geomorphology; Soil type	12.5 m × 12.5 m	Historical floods in 19th and 20th centuries (170 flood location)
75	Sahoo et al. (2021)	Hybrid	ANFIS, ANFIS-GOA, ANFIS-GWO	The Mahanadi river basin, India (141600)	Precipitation; Temperature; Solar radiation; Humidity; Evapotranspiration; Absorption loss; Percolation Loss	N.A. (Time series prediction)	1970–2019

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
76	Sankaranarayanan et al. (2020)	Neural Network, Kernel-based, Statistical	DNN, SVM, KNN, NB	The states of Bihar and Orissa in India	Rainfall; Humidity; Temperature; Flow velocity; Water depth	N.A. (Time series prediction)	1990–2002
77	Shahabi et al. (2021)	Neural Network (Deep learning), Statistical (Regression), Hybrid, Tree based	DBPGA, LR, LMT, BLR, ADT, NBT, REPTree, ANFIS-BAT, ANFIS-CA, ANFIS-IWO, ANFIS-ICA, ANFIS-FA	The Haraz watershed in Iran (4014)	Slope; Elevation; Plan-Curvature; TWI; SPI; DisRiv; Riv-Den; Rainfall; Lithology; NDVI; LULC	30 m × 30 m	Data from 194 historical (1995–2015) flash floods (194 flood locations)
78	Tabbussum and Dar (2021)	Neural Network, Hybrid Fuzzy-Neural	ANN, Fuzzy Logic (Mamdani, Sugeno), ANFIS	The tributaries of Jhelum river, India (8700)	Slope; Elevation; Evapotranspiration; SPI; Flow rate; STI; LULC; NDVI; Precipitation; Soil type	N.A. (TSP)	1990–2018 (497 flood locations)
79	Xu et al. (2023)	Hybrid (Tree-based, Boosting), Ensemble, Instance based	LightGBM, RF, XGBoost, KNN	Haikou, the capital city of Hainan Province, China(14)	Rainfall; Tide	25 m × 25 m	1990–2018 (a dataset of 490 examples)
80	Youssef et al. (2022)-Advanced	Kernel-based, Ensemble, Statistical (Regression), Hybrid	SVM, RF, MARS, BRT, FDA, GLM, MDA	The study area extends from Safaga to Ras Gharib (10537)	Elevation; Slope Aspect; Lithology; LULC; Slope length; TWI; Slope; Prof-Curvature; Plan-Curvature; SPI; Lithology	30 m × 30 m	2014–2016 (420 flood and 420 non-flood locations)
81	Youssef et al. (2022)-Flood	Statistical (Regression), Kernel based	FR (Bivariate), LR (Multivariate), SVM	The catchment area of Taif, Saudi Arabia (1312)	Elevation; Slope; Aspect; Plan-Curvature; Prof-Curvature; Valley depth; TWI; TPI	5 m × 5 m	Flood events in 2013, 2017, 2018, 2019, and 2020 (135 flood locations)
82	Youssef et al. (2022)-Optimal	Statistical (Regression), Hybrid, Ensemble	LR, EGB, RF	Wadi El-Matulla, located in the eastern desert of Egypt (7231)	Elevation; Slope; Plan-Curvature; TWI; Prof-Curvature; Plan-Curvature; Drain-Den; Geology; NDVI; LULC; DisRoa; DisRiv; Rainfall	12.5 m × 12.5 m	2018–2021 (480 flood and non-flood locations)
83	Mehravar et al. (2023)	Statistical (Regression), Hybrid	SVR, SVR-BA, SVR-IWO, SVR-FA	Ahwaz is a city in the southwest of Iran	LULC; Soil type; TWI; Rainfall; SPI; Plan-Curvature; Slope; Aspect; DisRiv; NDVI; Elevation	30 m × 30 m	11 and 24 March 2019, 7 and 16 April 2019 (300 flood and 300 non-flood locations)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
84	Shafizadeh-Moghadam et al. (2018)	Neural Network, Tree based, Statistical, Ensemble (ML + Statistical)	ANN, CART, FDA, GLM, GAM, BRT, MARS, MaxEnt, EMca, EMciInf, EMciSup, EMcv, EMmean, EMmedian, EMwmean	Haraz watershed, which is located in the north of Iran (4014)	Slope; Plan-Curvature; Elevation; TWI; SPI; DisRiv; Riv-Den; LULC; NDVI; Rainfall; Lithology	20 m × 20 m	1991–2011 (201 flood events and 10,000 randomly selected non-occurrence points)
85	Razavi-Termeh et al. (2023)	Ensemble, Hybrid	RF, Bagging, RF-GA, Bagging-GA	Sulaymaniyah province of Iraq (17023)	Rainfall; NDVI; Aspect; LULC; Elevation; SPI; Plan-Curvature; TWI; Geology	30 m × 30 m	December 2018, February 2019, 27–28 January 2019 (160 flood points and 160 non-flood locations)
86	Saravanan et al. (2023)	Ensemble	Adaboost, GB, XGB, CatBoost, SGB	Idukki district, Kerala, India (4358)	Aspect; Elevation; Slope; SPI; STI; TRI; TWI; Plan-Curvature	30 m × 30 m	From June 1 to August 19, 2018
87	Zhao et al. (2019)	Kernel based, Statistical (Regression), Neural Network	WELLSVM, LR, ANN, SVM	Beijing, the capital of China (2266)	Precipitation; FP; Elevation; Slope; TWI; DisRiv; Drain-Den; DisRoa; NDBI	30 m × 30 m	2004 to 2014 (1001 flood and 1001 non-flood locations)
88	Shahabi et al. (2020)	Hybrid, Instance-Based	Bagging Tree-KNN (Coarse, Cosine, Cubic, Weighted), KNN (Coarse, Cosine, Cubic, Weighted)	Haraz watershed in northern Iran (4015)	Elevation; Slope; Plan-Curvature; TWI; SPI; DisRiv; Drain-Den; Rainfall; LULC; Lithology	30 m × 30 m	Flood events in 2008, 2012, 2016 and 2017 (201 flood and 201 non-flood locations)
89	Band et al. (2020)	Tree based, Ensemble	BRT, RF, PRF, RRF, ERT	Kalvan watershed in Markazi Province, Iran (2056.75)	Elevation; Slope; Aspect; Plan-Curvature; Prof-Curvature; DisRiv; DisRoa; LULC; Lithology; Soil depth; Rainfall; SPI; TWI	12.5 m × 12.5 m	1991–2019 (256 flood and 256 non-flood locations)
90	Bui et al. (2019)-A new	Hybrid, Neural Network, Kernel based, Tree based	PSO-MARS, BPNN, SVM, Classification Tree	Bac Ha and Bao Yen (BHBY) in the Lao Cai, Vietnam (1510.4)	Elevation; Slope; Plan-Curvature; TWI; Aspect; TWI; SPI; Riv-Den; NDVI; Soil type; Lithology; Rainfall	10 m × 10 m	2017 (654 flood locations)
91	Hoseini et al. (2020)	Ensemble, Statistical (Regression)	GLMBoost, RF, BayesGLM	Gorganroud River Basin in the north of Iran (11290)	Elevation; Slope; Aspect; TPI; TPI; FloAcc; TWI; Drain-Den; DisRiv; Precipitation; NDVI; Soil depth; Soil type; LULC; Lithology	30 m × 30 m	11 March 2019 to 10 April 2019 (368 flood and 368 non-flood locations)

No.	Author (year)	Used ML model		Catchment (size in km^2)	Input parameters	Spatial res.	Data span (data number)
		Configuration	Type				
92	Bui et al. (2019)-A novel	Hybrid, Neural Network, Kernel based, Tree based	PSO-ELM, MLP, SVM, C4.5 DT	Bac Ha and Bao Yen (BHBY) in the Lao Cai, Vietnam (1510.4)	Elevation; Slope; Plan-Curvature; TWI; Aspect; TWI; SPI; Riv-Den; NDVI; Soil type; Lithology; Rainfall	10 m × 10 m	2017 (654 flood locations)
93	Tehrany et al. (2019)	Tree based, Kernel based	DT, SVM	The watershed and Brisbane River, Australia	Elevation; Slope; Aspect; Plan-Curvature; SPI; TWI; TRI; STI	5 m × 5 m	January 2011 flood (400 locations)
94	Dodangeh et al. (2020)	Statistical (Regression), Ensemble, Hybrid, Neural Network, Kernel Based	GAM, BTR, MARS, RS-GAM, BT-GAM, RS-BTR, BT-BTR, RS-MARS, BT-MARS, MLP, SVM	Ardabil Province in northwest Iran (17953)	Elevation; Slope; Aspect; Plan-Curvature; DisRiv; Rainfall; NDVI; LULC; Lithology	30 m × 30 m	1988 and 2016 (147 flood and 147 non-flood locations)
95	Arora et al. (2021)	Hybrid	ANFIS, ANFIS-GA, ANFIS-DE, ANFIS-PSO	Area of upper and lower Ganga River Basin (GRB) (10138)	Rainfall; Soil type; Riv-Den; DisRiv; DisRoa; TWI; Elevation; Slope Aspect; Slope; Plan-Curvature; LULC; Geomorphology	30 m × 30 m	2008 (May, September and October) and 11–22 Feb 2000 (1000 flood and 1000 non-flood locations)
96	Andaryani et al. (2021)	Neural Network, Hybrid	MLP, FART, and SOM(all with four activation functions: Sigmoidal, Linear, Commitment, Typicality)	The Ajichay river, Iran (7567)	Elevation; Slope; Aspect; Plan-Curvature; SPI; TWI; Lithology; LULC; Rainfall; DisRiv	30 m × 30 m	1999, 2005–2012, and 2014–2017 (245 flood and 272 non-flood locations)
97	Razavi-Termeh et al. (2023)	Ensemble, Hybrid	RF, RF-IWO, RF-SMA, RF-SBO	Estahban town in the east of Fars province, Iran (2652)	Rainfall; LULC; Elevation; SPI; Slope Aspect; NDVI; Slope; Lithology; Plan-Curvature; DisRiv; Prof-Curvature; TWI	30 m × 30 m	2022-01-01 to 2022-07-06 (509 flood and 509 non-flood locations)
98	Lyu et al. (2023)	Ensemble (Tree based)	RF, GBDT, XGBoost, Categorical Boosting	The Guangdong Greater Bay in Hong Kong-Macau (55900)	Elevation; Slope; Aspect; TWI; LULC; Riv-Den	30 m × 30 m	2010–2022 (594 flood and 594 non-flood locations)
99	Liao et al. (2023)	Neural Network (deep learning), Ensemble (Tree based), Instance based	CNN, XGBoost, MORF, KNN	The Chebei River Basin (CRB), China (74)	Rainfall; Water depth	8 m × 8 m	Test case: Flood event in April 23, 2022 (852 locations)
100	Yu et al. (2023)	Ensemble (Tree based), Tree based, Instance based, Kernel based, Neural Network	RF, GBDT, KNN, SVM, ANN	Wuhan City, China (8569.15)	Elevation; Slope; Aspect; Plan-Curvature; NDVI; TWI; NDBI; SPI; LULC; Soil type; DisRiv	30 m × 30 m	Floods in 2016 to 2022 (77 flood locations)

Appendix 4

Flowchart

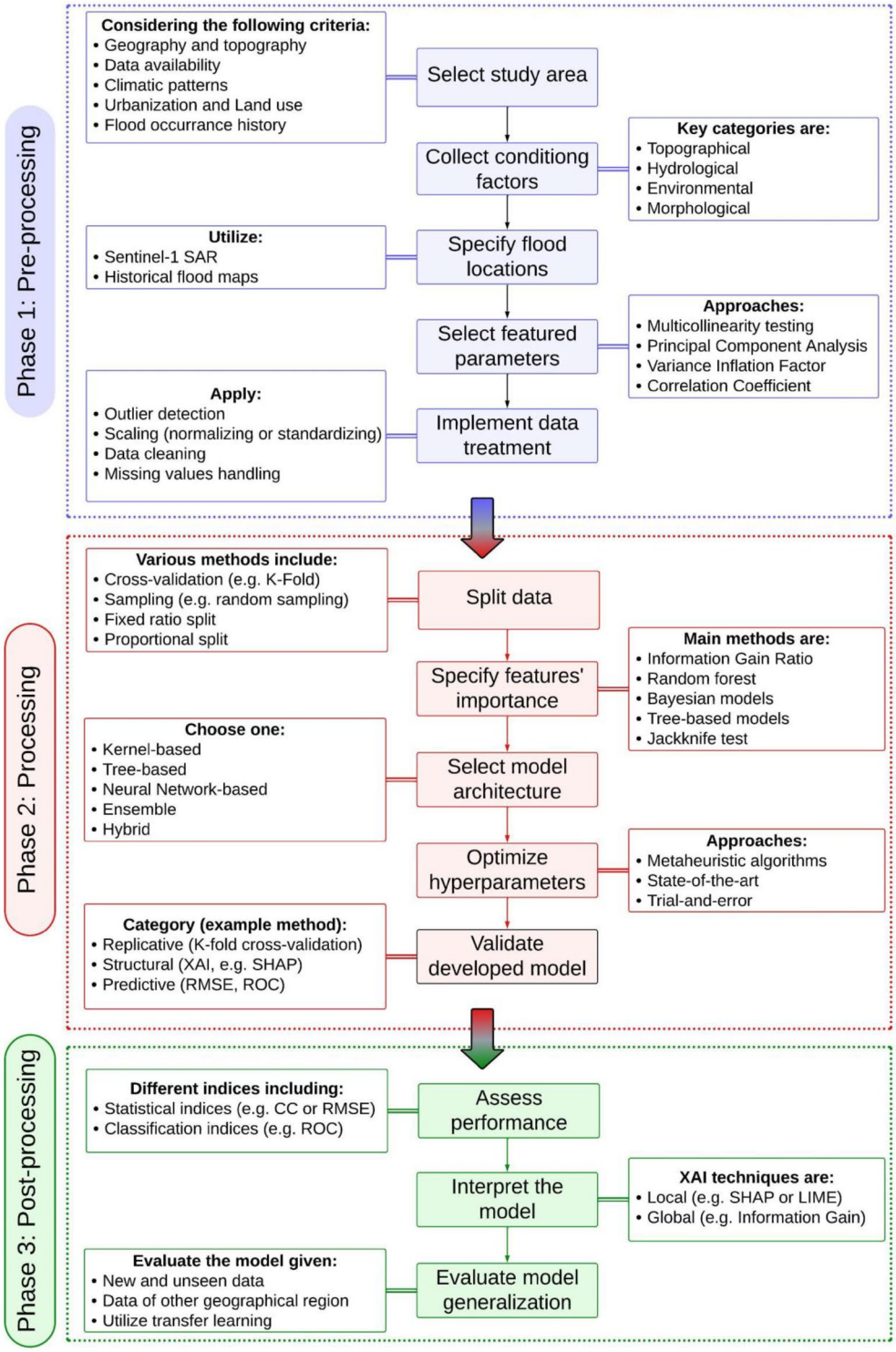

FIGURE A1 | Phases and steps involved in developing a machine learning-based model for flood susceptibility mapping, including pre-processing, processing, and post-processing stages.