NG CIVIL ENGINEERING

NG Civil Engineering
ISSN (online): 3107-569X
Volume 2, Issue 2 (April-June), 2026, 31-51

NEXT GEN ACADEMIC SERVICES
CATALYZING KNOWLEDGE FOR TOMORROW

Review Article

Open Access

Remote Sensing and GIS-Based Flood Hazard and Susceptibility Mapping: A Review of Methods, Datasets and Applications

Sangita Hazra*

Department of Geospatial Science, Techno India University, West Bengal, India

*Corresponding author email ID: sangitahazra2001@gmail.com

HIGHLIGHTS

- Remote sensing and GIS improve flood susceptibility and hazard mapping through multi-source geospatial data.
- Model reliability depends on quality inventories, robust validation, and uncertainty assessment
- Future workflows should combine explainable AI, physical knowledge, and local evidence for decision support.

ARTICLE INFO

Article History:
Received: 20 March 2026
Revised: 22 April 2026
Accepted: 29 may 2026
Published: 30 June 2026

Keywords:
Flood hazard
Hazard Mapping
Geographic Information System (GIS)
Synthetic aperture radar (SAR)
Disaster risk
Machine learning

ABSTRACT

Flood hazard and susceptibility mapping is essential for disaster risk reduction because it translates hydrological, geomorphological, meteorological, land-cover, and exposure data into spatial information for preparedness, land-use planning, infrastructure development, agricultural risk management, and emergency response. Remote sensing and GIS have strengthened this field through Earth observation, terrain analysis, rainfall datasets, land-use monitoring, flood inventories, and spatial modeling tools. This review examines major data sources, conditioning factors, modeling approaches, validation methods, operational challenges, and future research priorities in flood hazard and susceptibility mapping. It distinguishes hazard, susceptibility, vulnerability, exposure, and risk, while evaluating optical and SAR imagery, DEMs, rainfall, soil, land-cover, and ancillary datasets. Common approaches include weighted overlay, analytical hierarchy process, frequency ratio, logistic regression, support vector machines, random forest, XGBoost, deep learning, and explainable AI. Despite technological progress, model accuracy remains limited by poor flood inventories, sampling bias, class imbalance, scale effects, DEM uncertainty, limited transferability, and weak validation using independent flood events. Future efforts should prioritize transparent, reproducible, multi-source, physically informed, and decision-oriented flood-mapping workflows that integrate Earth observation, local knowledge, and explainable geospatial AI.

1. INTRODUCTION

Floods are among the most spatially extensive and socio-economically disruptive natural hazards, but the technical challenge for planners is not merely to describe flood events after they occur. The central challenge is to identify where flooding is likely, why certain locations are repeatedly affected, and how the

https://doi.org/10.66132/ngce20260204

resulting maps can inform decisions before the next event. Remote sensing and GIS have become indispensable because flood processes are spatial, temporally dynamic, and influenced by interacting controls such as rainfall intensity, terrain, drainage proximity, soil permeability, land-cover conversion, river morphology, surface roughness, and built infrastructure. A review paper in this area must therefore examine both the data pipeline and the modeling logic, rather than treating flood mapping as a simple cartographic exercise.

The term flood mapping is often used loosely in the literature, but it covers at least three different analytical tasks. Flood inundation mapping delineates observed water extent during or shortly after an event. Flood hazard mapping estimates the probability, depth, velocity, or intensity of flooding under specified scenarios. Flood susceptibility mapping identifies locations that have an inherent tendency to flood based on conditioning factors, even when full hydrodynamic information is unavailable. Remote sensing contributes strongly to all three tasks: optical and SAR imagery support flood extent detection, digital elevation models provide topographic controls, satellite rainfall products represent hydrometeorological forcing, and land-cover products show changes in impervious surface, vegetation, and agricultural land. GIS provides the spatial integration environment in which these inputs are harmonized, reclassified, weighted, modeled, and validated.

Recent literature shows a clear shift from manual interpretation and subjective overlay toward more empirical workflows. Classical GIS-based multi-criteria decision analysis, especially analytical hierarchy process (AHP), continues to be widely used because it is transparent, low-cost, and applicable in data-scarce basins. However, studies increasingly use frequency ratio, weights of evidence, logistic regression, random forest, support vector machine, gradient boosting, convolutional neural networks, and explainable-AI methods to capture nonlinear relationships between flood inventories and conditioning factors. The move is not merely algorithmic. It reflects a broader transition toward data-driven geospatial risk assessment, cloud computing platforms such as Google Earth Engine, SAR-derived inventories, open rainfall products, high-resolution DEMs, and independent validation against observed flood extents.

Synthetic aperture radar has become particularly important because optical flood monitoring is frequently constrained by cloud cover, rainfall, smoke, darkness, and inconsistent acquisition timing during flood events. Sentinel-1 SAR, with weather-independent imaging, has expanded the ability to build flood inventories and detect inundation across large areas, including regions where field data are sparse. Nevertheless, SAR is not a magic solution. Flooded vegetation, double-bounce scattering in urban areas, wind-roughened water, layover, shadow, speckle, and confusion with permanent water bodies all complicate interpretation. The best recent work therefore treats SAR as one component of a multi-source evidence system rather than as a standalone answer.

This review is motivated by three practical problems observed in the flood susceptibility literature. First, many studies produce visually attractive maps but provide weak justification for factor selection, weighting, training-data sampling, and validation. Second, model comparison is often reduced to area under the curve (AUC), while spatial transferability, event independence, calibration, and decision usefulness are neglected. Third, many papers use the terms hazard, susceptibility, vulnerability, and risk interchangeably, which weakens methodological clarity and policy relevance. A serious review must confront these weaknesses directly because flood maps are not academic illustrations; they may influence zoning, insurance, infrastructure, evacuation planning, and public investment.

The objective of this paper is to review the application of remote sensing and GIS in flood hazard and susceptibility mapping, focusing on methods, datasets, conditioning factors, validation strategies, and emerging research directions. The review gives special attention to recent studies using SAR, cloud platforms, machine learning, and explainable geospatial artificial intelligence, while retaining the classical GIS and AHP literature because many regions still lack dense hydraulic, hydrological, and field-observation datasets. The paper is structured to move from concepts and data sources to modeling approaches, validation, applications, limitations, and future directions.

2. REVIEW APPROACH AND SCOPE

This paper is designed as a narrative-structured review with systematic elements. The review focuses on peer-reviewed studies and recent review articles dealing with flood hazard, susceptibility, flood-inundation inventory generation, GIS-based conditioning-factor modeling, remote sensing datasets, multi-criteria decision analysis, statistical learning, machine learning, and validation. Preference was given to literature from approximately the last decade, with stronger weight on 2024-2026 contributions where they directly advance SAR, Google Earth Engine, machine learning, or explainable-AI practice. Foundational works were retained only where they established widely used methods or conceptual distinctions that remain important for current research.

The scope is limited to spatial flood hazard and susceptibility mapping. It does not attempt a full review of hydraulic modeling, real-time hydrological forecasting, dam-break simulation, or socio-economic vulnerability assessment, although these areas are discussed where they intersect with remote sensing and GIS. The review also avoids presenting remote sensing as a substitute for hydrological reasoning. A flood susceptibility map built only from pattern recognition may perform well within one inventory but fail outside the sampled event space if rainfall mechanisms, land-use conditions, drainage interventions, or geomorphic settings change. Therefore, the review evaluates both empirical performance and physical plausibility.

The studies considered here can be grouped into six broad categories: flood-inundation extraction from remote sensing imagery; GIS-based multi-criteria hazard mapping; statistical susceptibility models; machine-learning and deep-learning susceptibility models; cloud-computing and big-data workflows; and validation-focused studies. These categories overlap in practice. For example, a study may use Sentinel-1 SAR to create an inventory, GIS to derive conditioning factors, XGBoost to generate susceptibility classes, SHAP to interpret variable influence, and flood-sensor or community-reported locations for independent validation. Such integrated workflows represent the current frontier of applied geospatial flood science.

The review pays particular attention to the difference between model input evidence and model output claims. A map derived from elevation, slope, drainage distance, land cover, rainfall, and soil texture may reasonably be called a susceptibility map if it ranks the tendency of locations to flood. It should not be presented as a risk map unless it incorporates exposed population, assets, infrastructure, vulnerability, or damage functions. Similarly, a single-event SAR flood extent should not be generalized as long-term hazard without multi-event analysis or scenario modeling. These distinctions are methodological safeguards against overclaiming.

3. CONCEPTUAL FOUNDATIONS: HAZARD, SUSCEPTIBILITY, VULNERABILITY AND RISK

Conceptual precision is one of the most common weaknesses in flood-mapping papers. Flood hazard generally refers to the potentially damaging physical event or condition, often expressed through flood extent, water depth, velocity, frequency, duration, or return period. Flood susceptibility refers to the spatial propensity of an area to experience flooding based on physical and environmental controls. Vulnerability refers to the sensitivity or lack of capacity of people, assets, crops, infrastructure, or ecosystems to withstand flooding. Exposure refers to the presence of people, property, roads, crops, utilities, and other elements in places that may be flooded. Risk is the expected consequence that emerges from the interaction of hazard, exposure, and vulnerability.

Remote sensing and GIS can support all of these concepts, but the input requirements differ. Susceptibility mapping can often be conducted using terrain, drainage, rainfall, land-cover, soil, and historical flood-inventory data. Hazard mapping is stronger when hydrological or hydraulic simulations, event frequency, flood depth, and return-period estimates are included. Vulnerability mapping requires socio-economic, infrastructural, institutional, and livelihood data. Risk mapping requires overlaying hazard or susceptibility with exposure and vulnerability. The mistake to avoid is using a physical susceptibility output and labeling it as comprehensive risk.

The conceptual distinction has practical consequences. A low-lying floodplain with no current settlement may be highly susceptible but low in current human risk. A dense urban settlement built on a moderate-hazard surface may carry higher risk because exposure and vulnerability are high. An agricultural field may be physically flood-prone but socially recoverable if farmers have crop insurance, drainage support, resilient varieties, warning information, and alternative livelihoods. Therefore, remote sensing and GIS outputs should be framed according to their actual content. The most defensible papers state clearly whether they map susceptibility, hazard, exposure, vulnerability, or risk.

In the literature, GIS-based AHP and weighted overlay studies often combine physical factors to create a hazard or susceptibility index. Frequency ratio and weights-of-evidence models estimate the association between past flood locations and factor classes. Machine-learning models learn more complex relationships between observed flooded and non-flooded samples and predictor layers. Hydrodynamic models estimate water movement through physical equations, but they require more detailed boundary conditions, topography, roughness, river geometry, and calibration data. Each approach has value, but the output should be named and interpreted according to the data and assumptions used.

For review purposes, the strongest position is to treat susceptibility mapping as a practical middle ground. It is more spatially explicit than generic flood inventories and less data-demanding than full hydraulic simulation. It can support preliminary planning, hotspot screening, agricultural zoning, emergency preparedness, and prioritization of field surveys. However, it must be validated and communicated as a relative susceptibility product, not as a definitive prediction of flood depth, velocity, duration, or damage

.

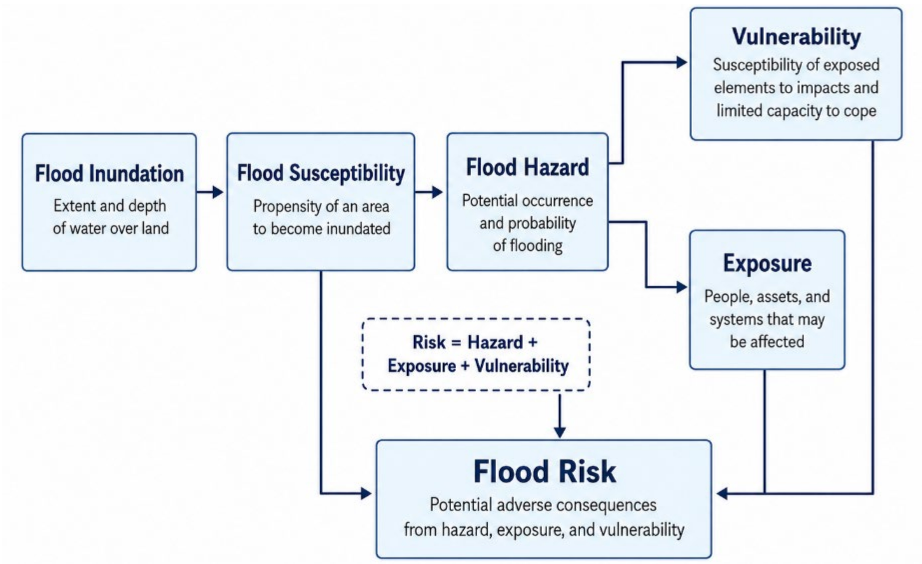

Figure 1. Conceptual distinction among flood inundation, susceptibility, hazard, exposure, vulnerability and risk in geospatial flood assessment.

Table 1. Conceptual distinction among key flood-mapping terms

Term	Primary meaning	Typical geospatial inputs	Appropriate output label	Common misuse to avoid
Flood inundation	Observed water extent during or after a specific flood event	Optical imagery, SAR imagery, water indices, change detection, field points	Event flood extent map	Treating one event as a long-term hazard map
Flood susceptibility	Relative tendency of a location to experience flooding	DEM derivatives, drainage, rainfall, LULC, soil, flood inventory	Susceptibility zonation or probability map	Calling a physical susceptibility map a risk map
Flood hazard	Physical intensity or probability of flooding	Return-period rainfall, hydrology, hydraulic depth, velocity, extent, duration	Hazard map or scenario map	Ignoring flood depth, frequency, or intensity
Exposure	People, infrastructure, crops, assets, and services located in flood-prone zones	Population grids, buildings, roads, crop maps, utilities	Exposure layer or exposed asset inventory	Confusing exposure with vulnerability
Vulnerability	Degree of potential harm or limited capacity to cope	Socio-economic indicators, building type, livelihood sensitivity, preparedness	Vulnerability index or profile	Using only physical terrain factors
Risk	Expected adverse consequence from hazard, exposure, and vulnerability	Integrated hazard/susceptibility, exposure, vulnerability, consequence data	Flood risk map	Using hazard alone as risk

4. REMOTE SENSING DATA SOURCES FOR FLOOD MAPPING

Remote sensing contributes to flood mapping in two major ways: direct observation of water extent and indirect derivation of flood-conditioning variables. Direct observation includes optical or SAR-based detection of inundated areas. Indirect derivation includes DEM-based terrain metrics, land-use and land-cover classification, vegetation condition, impervious surface estimates, soil moisture proxies, surface roughness, rainfall products, and drainage network extraction. The strength of a flood susceptibility model depends not only on the algorithm but on whether these input layers represent the actual flood-generating processes at the selected spatial scale.

Optical satellite imagery, including Landsat, Sentinel-2, MODIS, and commercial high-resolution imagery, is useful for detecting open water, mapping land-cover change, identifying floodplain agriculture, monitoring river migration, and deriving vegetation or water indices. Indices such as normalized difference water index, modified normalized difference water index, normalized difference vegetation index, and normalized difference built-up index are common in flood studies. Optical data are intuitive and spectrally rich, but they are limited during active flood events because clouds and rainfall frequently obscure the surface. This limitation is especially serious in monsoon regions and cyclone-affected coastal basins where flood peaks occur under persistent cloud cover.

SAR imagery overcomes several optical limitations because microwave signals can image the surface independent of daylight and most weather conditions. Sentinel-1 has become the dominant open SAR source for flood studies because of free access, repeat acquisition, and broad spatial coverage. SAR-based flood extraction is often built on backscatter thresholding, change detection, coherence analysis, machine learning, or object-based classification. A recent review of SAR flood detection emphasizes that open SAR data have stimulated extensive research but that fair assessment requires quantitative metrics and high-quality reference datasets; it also warns that vegetated and urban areas remain problematic because scattering mechanisms can mask water signals (Amitrano et al., 2024).

SAR is especially useful for generating flood inventories for susceptibility modeling. Instead of relying only on historical reports or digitized flood polygons, researchers can derive multi-event flood observations and use them as dependent variables in statistical or machine-learning models. Hitouri et al. (2024), for example, used SAR-derived inventory information and 12 conditioning variables to train random forest, CART, SVM, and XGBoost models in a small Moroccan watershed, finding that random forest performed best among the tested models. Such studies illustrate how SAR shifts susceptibility mapping from expert-only weighting toward empirical learning from observed flood evidence.

Digital elevation models remain among the most important datasets for flood susceptibility mapping because water movement is controlled by topography. DEMs support derivation of elevation, slope, curvature, drainage direction, flow accumulation, stream power index, topographic wetness index, topographic position, relative relief, and distance to drainage. However, DEM quality can strongly affect results. Coarse DEMs may miss levees, embankments, drainage channels, road barriers, and micro-topography. Vertical errors can be especially problematic in flat coastal plains and urban areas, where small elevation differences control flood pathways. Researchers should therefore report DEM source, resolution, vertical accuracy, preprocessing, sink filling, and hydrological conditioning.

Satellite rainfall products and gridded climate datasets, including CHIRPS, GPM IMERG, ERA5 derivatives, and regional precipitation products, are frequently used where gauge networks are sparse. Rainfall can be represented as annual or seasonal totals, extreme-event rainfall, maximum daily rainfall, antecedent rainfall, or long-term climatological averages. For susceptibility mapping, rainfall should be aligned with the flood mechanism. A model of pluvial urban flooding should not rely only on annual rainfall if short-duration high-intensity rainfall is the trigger. Similarly, riverine flood susceptibility may require upstream catchment rainfall, not only local rainfall at the flooded pixel.

Land-use and land-cover data are central because land-cover change modifies runoff generation, infiltration, surface roughness, drainage obstruction, and exposure. Urban expansion generally increases imperviousness and shortens runoff response time, while deforestation and agricultural intensification can modify infiltration, sedimentation, and channel capacity. In agricultural flood studies, crop type, cropping season, floodplain encroachment, paddy distribution, and drainage infrastructure may be as important as generic LULC classes. The use of remote sensing for LULC mapping therefore should be tied to a clear hydrological interpretation instead of being included only because it is easily available.

Soil and geological datasets are often used to represent infiltration, permeability, lithological resistance, and groundwater-surface water interaction. Soil texture, hydrologic soil group, drainage class, clay content, and permeability influence how rainfall becomes runoff. Geological structure can influence drainage density, channel pattern, and subsurface flow. These layers are useful, but they are frequently coarser than the target mapping resolution. A common beginner mistake is to combine a 30 m DEM with soil or geology data mapped at regional scale and then interpret the output as if every pixel has high-resolution soil information. Proper review and methods sections should acknowledge such scale mismatch.

Cloud platforms, especially Google Earth Engine, have changed flood mapping by allowing users to process long time series of satellite imagery, rainfall, DEM, and land-cover data without downloading massive datasets. Swain et al. (2020) demonstrated the usefulness of combining AHP, GIS, remote sensing, and GEE for flood susceptibility mapping in Bihar, India. More recent studies extend this logic by creating multi-year SAR flood inventories, deriving repeated ponding hotspots, and integrating them with machine learning. The advantage is reproducibility and scalability; the risk is black-box processing when authors do not report thresholds, masks, preprocessing choices, and validation logic in sufficient detail.A typical remote sensing and GIS data-processing chain for flood hazard and susceptibility mapping is presented in Figure 2

.

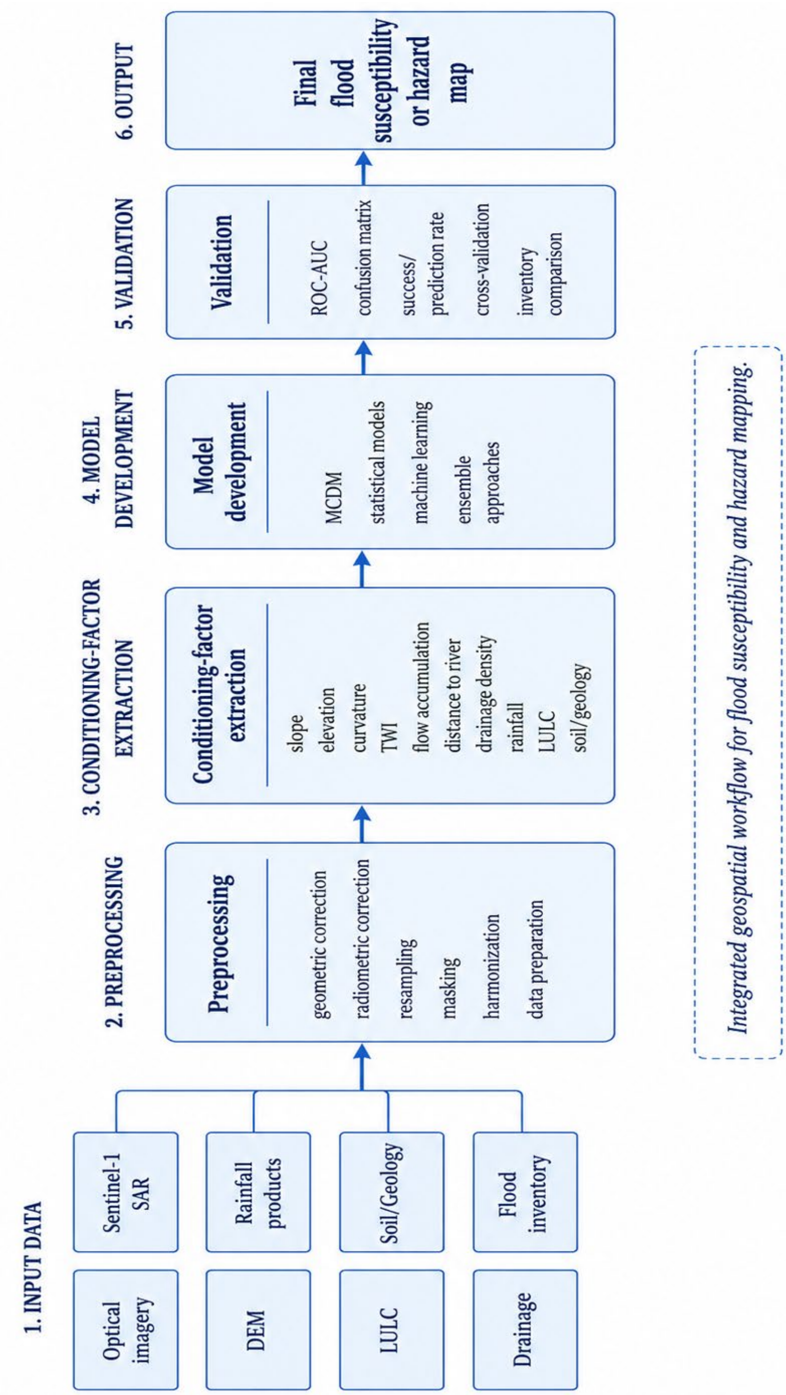

Figure 2. Remote sensing and GIS data pipeline for flood hazard and susceptibility mapping.

Table 2. Major remote sensing and GIS datasets used in flood hazard and susceptibility mapping

Dataset or product type	Typical spatial role	Strengths	Limitations	Representative use in literature
Optical satellite imagery	Open-water extraction, LULC mapping, vegetation and water indices	Spectrally rich, intuitive, long archives	Cloud cover, flood-timing mismatch, shadow confusion	Water-index and LULC-based conditioning layers in GIS workflows
Sentinel-1 SAR	Event flood extent, multi-year flood inventory, all-weather monitoring	Cloud-penetrating, day-night imaging, suitable for emergencies	Speckle, urban double bounce, vegetated flood ambiguity	SAR flood detection and inventories (Amitrano et al., 2024; Hitouri et al., 2024)
Digital elevation models	Elevation, slope, curvature, TWI, flow accumulation, drainage derivation	Core control on runoff and inundation tendency	Vertical error, resolution limits, hydrological conditioning needs	DEM derivatives in AHP, FR, ML models (Swain et al., 2020; Feizbahr et al., 2025)
Satellite rainfall products	Rainfall forcing, extreme rainfall zones, antecedent moisture proxy	Useful where gauges are sparse; spatially continuous	Bias, temporal aggregation, limited local extremes	Rainfall as conditioning factor in susceptibility models
Land-use/land-cover maps	Imperviousness, agriculture, vegetation, wetlands, built-up areas	Captures human modification of runoff and exposure	Classification error, temporal mismatch	LULC conditioning in urban and rural flood studies
Soil and geology maps	Infiltration, permeability, runoff potential	Supports process interpretation	Often coarse and generalized	Soil type, lithology, hydrologic groups in GIS/ML models
Ancillary spatial data	Roads, drainage, river distance, flood sensors, field points	Supports validation and operational relevance	May be incomplete or institutionally fragmented	Independent validation in sensor- and community-supported studies (Feizbahr et al., 2025)

5. FLOOD-CONDITIONING FACTORS IN GIS-BASED MAPPING

Flood-conditioning factors are the explanatory variables used to estimate why some locations are more flood-prone than others. In remote sensing and GIS-based studies, these factors usually include topographic, hydrological, meteorological, land-cover, soil, geological, and anthropogenic variables. The most defensible factor set is not necessarily the longest list. It is the set that is physically meaningful, spatially available at appropriate resolution, not excessively redundant, and validated against observed flood evidence. Adding many poorly justified variables can improve apparent model performance while reducing interpretability and transferability.

Topographic factors are almost universal. Elevation is a primary control because low-lying areas are more likely to receive and retain floodwaters. Slope influences the speed of runoff and the potential for water accumulation; flatter terrain is often associated with ponding and slow drainage, while steep slopes may be associated with flash runoff but limited local storage. Curvature affects convergence and divergence of flow paths. Relative relief and topographic position can distinguish valley bottoms from ridges. Topographic wetness index and flow accumulation are commonly used to represent zones where water tends to concentrate. These factors are computationally straightforward, but their values depend heavily on DEM quality and preprocessing.

Hydrological proximity factors, especially distance to rivers, streams, drainage channels, and wetlands, are widely used because flood frequency generally increases near flow corridors. However, distance-based indicators must be interpreted carefully. In engineered urban areas, flood-prone zones may occur far from natural channels due to blocked drains, undersized culverts, impervious surfaces, and local depressions. In alluvial plains, flood pathways may follow abandoned channels or distributary networks not captured in standard drainage layers. Therefore, drainage-distance variables should ideally be complemented with flow accumulation, drainage density, floodplain morphology, and historical inundation evidence.

Rainfall is both a trigger and a conditioning factor depending on how it is represented. Long-term rainfall climatology may explain broad regional differences, while event rainfall, maximum daily rainfall, antecedent rainfall, and intensity-duration variables are more relevant for specific flood mechanisms. In susceptibility mapping, rainfall layers are often static averages, but flooding is inherently episodic. This creates a mismatch: the model may identify areas with high long-term rainfall rather than areas vulnerable to specific extreme rainfall patterns. Future studies should better distinguish rainfall climatology from event-scale forcing and should test whether rainfall variables improve model performance beyond topography and drainage alone.

Land-use and land-cover conditioning factors are increasingly important because flood susceptibility is not purely natural. Urbanization increases impervious surface, reduces infiltration, alters drainage, and concentrates exposure. Agricultural land can either absorb or convey runoff depending on crop type, soil condition, field bunding, drainage design, and season. Wetlands can store water, but encroached wetlands may become flood hotspots. Forest loss can increase runoff and sedimentation, although effects vary by scale and rainfall regime. A high-quality review should therefore treat LULC not as a single categorical layer but as a dynamic representation of human modification of hydrological response.

Soil and lithology determine infiltration, storage, and runoff generation. Clay-rich soils, low-permeability formations, shallow groundwater, and compacted surfaces can increase flood susceptibility by limiting infiltration. Sandy soils and permeable lithologies may reduce surface ponding but can be associated with groundwater flooding in some settings. Because soil and geology datasets are often generalized, their role should be interpreted at the scale at which they were mapped. When a machine-learning model assigns high importance to soil type, the analyst should check whether the variable reflects real hydrological mechanisms or merely spatially coincides with past flood locations.

Anthropogenic variables are underdeveloped in many susceptibility studies. Distance from roads, drainage infrastructure, embankments, canals, settlements, culverts, and land reclamation can meaningfully affect flood pathways. Road embankments may obstruct natural drainage; canals may convey or store floodwater; levees may reduce frequent flooding but increase residual risk when overtopped. Urban studies increasingly include impervious surface, building density, road density, drainage density, and depression areas. Feizbahr et al. (2025) included depression areas, soil hydrologic group, major and minor stream distance, distance to roads, and other geospatial factors, then validated results using multiple independent sources including sensors and local task-force information.

Multicollinearity and factor redundancy are serious issues. Elevation, slope, flow accumulation, TWI, drainage density, and distance to streams can be correlated. In AHP, correlated variables may cause double-counting if weights are assigned independently. In machine learning, correlated variables may inflate feature importance or make interpretation unstable. A professional workflow should include correlation screening, variance inflation checks, feature importance analysis, ablation testing, or domain-based factor reduction. The goal is not to maximize the number of layers but to build a parsimonious model that represents flood processes and remains interpretable.

Table 3. Common flood-conditioning factors and their methodological implications

Factor group	Examples	Hydrological interpretation	Main caution	Suitable methods
Topography	Elevation, slope, curvature, TWI, flow accumulation	Controls water movement, concentration, and retention	DEM resolution and vertical error may dominate results	AHP, FR, WoE, LR, RF, XGBoost, CNN
Drainage proximity	Distance to river, distance to stream, drainage density	Represents riverine exposure and flow concentration	Urban drainage and abandoned channels may be missed	AHP, FR, WoE, logistic regression, ML
Rainfall	Mean rainfall, extreme rainfall, event rainfall, antecedent rainfall	Represents flood trigger and spatial climatic forcing	Temporal mismatch between rainfall layer and flood inventory	Statistical, ML, event-based GIS analysis
Land cover	Built-up area, agriculture, forest, wetland, vegetation indices	Represents infiltration, roughness, runoff and human alteration	Classification error and temporal mismatch	GIS overlay, ML, change analysis
Soil and geology	Soil texture, hydrologic group, lithology, permeability	Controls infiltration, runoff and subsurface storage	Often mapped at coarser scale than DEM or imagery	GIS-MCDA, FR, LR, ML
Anthropogenic features	Roads, embankments, canals, drainage, settlements	Represents obstruction, drainage modification and exposure	Data incompleteness and local engineering complexity	Urban GIS, object-based analysis, ML
Flood inventory	Past flood polygons, SAR flood extent, field points	Provides target variable or validation evidence	Sampling bias and event dependence	FR, WoE, LR, supervised ML, validation

6. GIS-BASED MULTI-CRITERIA AND STATISTICAL METHODS

GIS-based multi-criteria methods remain widely used because they are transparent, relatively simple, and applicable where historical flood inventory data are incomplete. The typical workflow involves selecting conditioning factors, reclassifying each factor into susceptibility classes, assigning weights, overlaying the weighted layers, and classifying the final index into very low, low, moderate, high, and very high susceptibility zones. The method is attractive for planners because the steps can be explained to non-specialists. Its weakness is that factor weights can be subjective and may reflect expert opinion more than empirical flood behavior.

The analytical hierarchy process is the most common multi-criteria decision method in flood susceptibility mapping. AHP uses pairwise comparisons to derive relative weights for factors and checks consistency through a consistency ratio. Swain et al. (2020) used AHP with GIS, remote sensing, and Google Earth Engine to map flood susceptibility in Bihar, India, selecting broad criteria including hydrologic, morphometric, permeability, land-cover dynamics, and anthropogenic interference. Their study demonstrated how a large number of parameters can be integrated in a cloud environment, but it also illustrates the core AHP issue: the final map depends on the weighting logic and the quality of expert judgment.

Weighted overlay and AHP are useful when the objective is preliminary screening, especially in data-scarce regions. They are less suitable when rich flood inventories are available because they do not learn directly from observed flood and non-flood samples. Another limitation is that AHP often assumes linear, additive effects. In reality, flood susceptibility may emerge from interactions. A low slope may be harmless if drainage is excellent, but highly flood-prone if combined with poor drainage, high rainfall, impervious cover, and proximity to river backwater. Additive scoring can miss these threshold and interaction effects.

Table 4. Comparison of major GIS-based flood susceptibility modeling approaches

Approach	Data requirement	Main strengths	Main limitations	Best use case
Weighted overlay	Conditioning factors and rule-based weights	Simple, transparent, easy to communicate	Highly subjective; additive assumptions	Preliminary planning and data-poor basins
AHP / fuzzy AHP	Conditioning factors and expert pairwise comparison	Structured weighting; consistency check; stakeholder-friendly	Expert bias; correlation and double-counting risk	Multi-criteria planning where inventory data are limited
Frequency ratio	Flood inventory and classified conditioning factors	Empirical, interpretable, GIS-friendly	Bivariate; sensitive to class intervals	Moderate-data susceptibility screening
Weights of evidence	Flood inventory and factor classes	Probabilistic logic; interpretable weights	Conditional independence assumption	Exploratory spatial association mapping
Logistic regression	Flood and non-flood samples with predictors	Multivariate probability estimation; interpretable coefficients	Functional-form assumptions; sampling sensitivity	Statistical baseline for model comparison
Random forest	Training inventory and predictor layers	Handles nonlinearities and interactions; robust performance	Less transparent; spatial sampling bias possible	General supervised susceptibility mapping
XGBoost / boosting	Training inventory and predictor layers	High predictive power; handles complex relationships	Hyperparameter sensitivity; overfitting risk	High-performance susceptibility and hotspot modeling
CNN / deep learning	Large labeled imagery or raster stacks	Learns spatial patterns and texture	Data hungry; weak interpretability if unmanaged	Flood segmentation and high-resolution pattern detection

Frequency ratio is a bivariate statistical approach that estimates the relationship between each factor class and flood occurrence. If a particular elevation class contains a higher proportion of flooded locations than expected by area, it receives a higher frequency ratio. Rahmati et al. (2016) applied frequency ratio and weights-of-evidence models for flood susceptibility mapping in Golestan Province, Iran, providing one of the widely cited examples of statistical susceptibility mapping. These methods are more empirical than AHP because they use observed flood locations, but they still treat variables largely independently and may be sensitive to class boundaries.

Weights of evidence is another bivariate approach derived from Bayesian probability logic. It estimates positive and negative weights for factor classes based on the presence or absence of flood occurrence. Like frequency ratio, it is interpretable and relatively easy to implement in GIS. Its limitations include assumptions of conditional independence among predictors and sensitivity to inventory quality. If flood inventory points are biased toward accessible locations, reported damage locations, or specific events, the weights may not generalize.

Logistic regression is a multivariate statistical model that estimates the probability of flood occurrence from predictor variables. It is more flexible than bivariate methods because it can account for multiple variables simultaneously. However, logistic regression assumes a particular functional relationship unless transformations and interactions are specified. It also requires careful sampling of flooded and non-flooded points. Poor selection of non-flood samples can produce inflated accuracy, especially if non-flood points are sampled from clearly upland areas rather than from plausible floodplain contrasts.

Comparative studies of AHP, fuzzy AHP, frequency ratio, and related methods show that model choice should depend on data availability and decision context. AHP and fuzzy AHP may be useful where local expert knowledge is strong but inventory data are weak. Frequency ratio and weights of evidence are useful when historical flood locations are available but sample sizes are moderate. Logistic regression provides a bridge toward machine learning by allowing multivariate probability estimation. A recent comparative flood hazard assessment emphasized that GIS and multi-criteria analysis are widespread because they can process distributed spatial inputs and produce easy-to-read maps for authorities and local users (Taoukidou et al., 2025).

The strongest practice is not to treat traditional methods as obsolete. In many regions, an AHP or frequency-ratio map may be more defensible than a poorly trained deep-learning model. What matters is methodological honesty. AHP maps should report pairwise matrices, consistency ratio, factor classes, and expert basis for weights. Frequency-ratio and weights-of-evidence studies should report inventory source, class intervals, sampling strategy, and validation. Logistic regression studies should report coefficients, predictor screening, multicollinearity, calibration, and independent test results. Without these details, the map is not reproducible.

7. MACHINE LEARNING, DEEP LEARNING AND EXPLAINABLE GEOAI

Machine learning has become central to recent flood susceptibility mapping because it can model nonlinear relationships and interactions among conditioning factors. Random forest, support vector machine, gradient boosting, XGBoost, artificial neural networks, convolutional neural networks, and hybrid ensemble approaches are now common. Their appeal is straightforward: flood occurrence is rarely controlled by one variable, and machine-learning algorithms can learn complex combinations of elevation, slope, drainage, rainfall, LULC, soil, road proximity, depression areas, and remote-sensing indices. However, high accuracy claims should be treated critically unless the study uses strong validation design.

Random forest is widely used because it is robust, relatively easy to tune, and capable of handling nonlinear predictor interactions. It generates variable-importance measures, although these can be biased by correlated predictors and variable scale. Hitouri et al. (2024) reported that random forest outperformed CART, SVM, and XGBoost in their SAR-supported small-watershed study in Morocco, achieving the best AUC among the tested models. This does not mean random forest is universally superior; it means that in a particular dataset, sampling design, predictor set, and landscape context, random forest captured the flood signal better than competing models.

Support vector machine has been used frequently in susceptibility mapping because it can model nonlinear class boundaries through kernel functions. It is effective with moderate sample sizes but can be sensitive to kernel choice, parameter tuning, class imbalance, and feature scaling. Artificial neural networks and multilayer perceptrons can model complex relationships but may lack interpretability and require careful training. Gradient boosting and XGBoost often deliver strong predictive performance because they sequentially correct errors and capture complex feature interactions. Feizbahr et al. (2025) found that XGBoost outperformed random forest, CNN, and frequency-ratio models in a Sentinel-1-integrated flood susceptibility study in Southeast Texas, with AUC reported as 0.92 and additional validation against sensors and local high-risk information.

Deep learning is more common in flood segmentation from imagery than in traditional conditioning-factor susceptibility mapping, although the boundary is blurring. Convolutional neural networks can use spatial context, texture, and neighborhood patterns from raster stacks. They are especially relevant where flood extent needs to be extracted from optical or SAR imagery. Recent SAR flood-detection literature emphasizes the importance of benchmark datasets, reproducibility, and careful validation because deep-learning models may perform well in open-water settings but struggle in vegetated and urban areas (Amitrano et al., 2024).

Semi-supervised learning and weakly supervised learning are increasingly relevant because high-quality flood labels are scarce. Zhao et al. (2019) used a semi-supervised machine-learning model for urban flood susceptibility assessment, demonstrating how partial labels can be leveraged when complete inventories are unavailable. This is important for urban settings where flood reports may be incomplete, insurance data may be inaccessible, and official records may underrepresent informal settlements or small recurrent pluvial floods. Yet semi-supervised approaches require careful assumptions about unlabeled data; unlabeled does not mean non-flooded.

Explainable artificial intelligence is one of the most important recent developments. Flood susceptibility maps are used by decision-makers, so a black-box probability surface is not enough. SHAP values, permutation importance, partial dependence plots, accumulated local effects, and local explanation methods can reveal whether models are using plausible flood controls or exploiting artifacts. Feizbahr et al. (2025) used SHAP to identify elevation, slope, and topographic wetness index as influential features across models. This type of interpretation helps bridge predictive modeling and hydrological reasoning, although SHAP results must still be interpreted under correlation and sampling constraints.

A crucial issue in machine-learning flood mapping is the creation of non-flood samples. Many studies randomly sample non-flood points outside flood polygons, but this can make the classification task artificially easy if non-flood points are far from floodplains or on high terrain. A more rigorous design samples non-flood points from hydrologically plausible areas, uses spatial cross-validation, tests on independent events, and avoids training-test leakage. If flood and non-flood samples are spatially autocorrelated, random train-test splits can overestimate performance because nearby pixels share the same terrain and land-cover conditions.

Another issue is class imbalance. Flooded pixels or points are usually much fewer than non-flooded locations. Accuracy alone can be misleading because a model can classify most pixels as non-flooded and still appear accurate. Metrics such as precision, recall, F1-score, AUC, area under the precision-recall curve, balanced accuracy, Kappa, and calibration curves should be considered. For planning purposes, false negatives may be more serious than false positives because missing a flood-prone settlement can lead to under preparedness. Therefore, model threshold selection should be linked to decision cost, not only statistical optimum.

The machine-learning literature also faces a transferability problem. A model trained in one basin may not perform well in another basin with different rainfall mechanisms, drainage networks, land-use patterns, soil conditions, engineering structures, or flood types. Even within one basin, a model trained on past events may fail when climate extremes, urban expansion, or drainage interventions alter flood behavior. The future of machine learning in flood susceptibility mapping therefore depends on spatial cross-validation, multi-event testing, domain adaptation, uncertainty quantification, and integration with hydrological knowledge.

The review conclusion is that machine learning is not automatically better than GIS-MCDA. It is better only when there is a reliable inventory, appropriate predictor data, careful sampling, robust validation, and transparent interpretation. AHP can be poor when weights are arbitrary; machine learning can be poor when labels are biased. Both can be useful when their assumptions are respected. The professional standard should be comparative modeling with baseline methods, independent validation, uncertainty reporting, and explicit discussion of decision relevance.

The major methodological families and their practical strengths and limitations are compared schematically in Figure 3.

	GIS-MCDA	Statistical Models	Machine Learning	Deep Learning
Examples	• AHP • Weighted Overlay	• Frequency Ratio • Weights of Evidence • Logistic Regression	• Random Forest • SVM • XGBoost	• CNN • Transformer Segmentation
Strengths	Simple, transparent, and useful with limited data	Interpretable and moderately data-efficient	High predictive accuracy and handles nonlinear interactions	Strong feature learning for complex spatial and image patterns
Limitations	Subjective weighting and weaker predictive power	Sensitive to assumptions and less effective for complex nonlinear patterns	Needs quality training data and is less interpretable	Data- and compute-intensive with lower interpretability

Figure 3. Comparative methodological framework of GIS-MCDA, statistical models, machine learning and deep learning in flood susceptibility mapping

8. VALIDATION AND ACCURACY ASSESSMENT

Validation is the weakest link in many flood hazard and susceptibility studies. A map cannot be trusted simply because it appears visually plausible. Validation should answer whether predicted high-susceptibility zones correspond to independent flood evidence, whether the model generalizes beyond the training sample, and whether the classification thresholds are appropriate for decision-making. The literature uses AUC, success-rate curves, prediction-rate curves, confusion matrices, precision, recall, F1-score, Kappa, and comparison with historical flood extents, but not all validation designs are equally strong.

AUC is widely used because it measures the ability of a model to discriminate flooded from non-flooded samples across thresholds. However, AUC can be inflated by spatial autocorrelation, easy negative samples, and random train-test splits. A high AUC does not guarantee that the susceptibility map is well calibrated, transferable, or useful for planning. If training and validation points are drawn from the same flood event and nearby locations, the test result may represent interpolation rather than true prediction. Therefore, AUC should be interpreted as one metric, not as final proof.

Independent flood inventories provide stronger validation. These may include flood polygons from different dates, Sentinel-1-derived flood extent not used in training, field GPS points, official disaster records, water-level sensors, local reports, UAV observations, or hydrodynamic simulation outputs. Feizbahr et al. (2025) strengthened validation by comparing model outputs with high-risk locations from flood sensors, base-level engineering maps, and local community task-force information. This multi-source validation is more useful than relying only on internal AUC because it tests whether the map aligns with operational evidence.

SAR-derived validation is valuable but must be used carefully. If Sentinel-1 data are used both to create the training inventory and validate the output, the validation is not independent. A stronger design uses different events, different sensors, or temporally separated SAR extents. Optical Sentinel-2 or Landsat imagery can validate flood extent when cloud-free acquisitions exist, but optical validation may underrepresent flooded areas during cloudy events. Alonso-Sarria et al. (2025) showed that validation with optical imagery can reduce apparent accuracy due to timing and false-negative issues, illustrating why validation source and acquisition timing matter.

Spatial cross-validation is increasingly necessary. In random cross-validation, neighboring pixels from the same landscape can appear in both training and testing sets, producing optimistic performance. Spatial blocking or basin-wise cross-validation better tests whether a model can generalize to unseen areas. Event-based validation is also important: a model trained on one or several flood events should be tested on a later event if the objective is prediction. Without spatial or temporal separation, reported accuracy may overstate real-world performance.

Validation should also consider map-class agreement. Many studies classify susceptibility into five classes and then report the percentage of observed flood points falling within high and very high classes. This is intuitive for planners, but class breaks can be arbitrary. Natural breaks, quantiles, equal intervals, and expert thresholds produce different class areas. A model that labels half the basin as high susceptibility will capture many flood points but may be too broad for action. Therefore, the area proportion of each class and the concentration of observed floods within each class should both be reported.

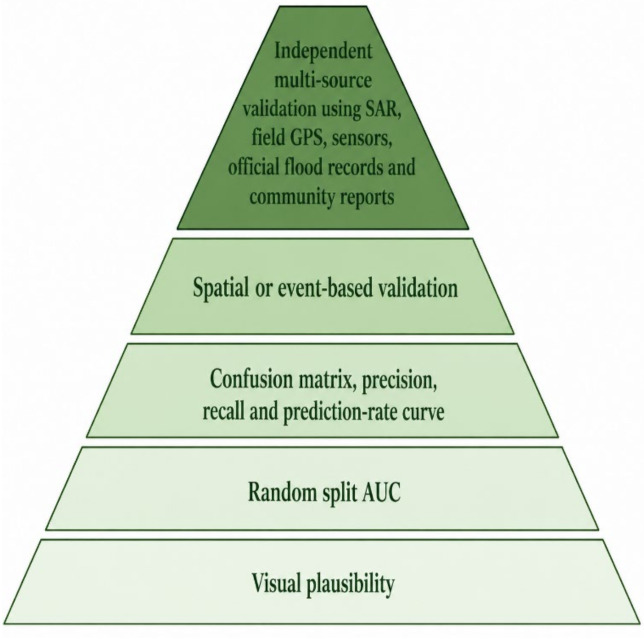

Figure 4. Validation hierarchy for flood susceptibility maps from internal metrics to independent multi-source evidence

Uncertainty reporting remains underdeveloped. Flood susceptibility maps often present crisp classes as if boundaries are exact. In reality, uncertainty arises from input data, DEM error, flood inventory incompleteness, sampling design, algorithm selection, parameter tuning, classification thresholds, and future climate variability. Ensemble modeling, bootstrap resampling, probability intervals, sensitivity analysis, and map-difference analysis can help communicate uncertainty. A professional map should show not only where susceptibility is high but also where model confidence is low.

The most useful validation is decision-oriented. For emergency management, a map should identify evacuation routes, shelters, and critical infrastructure likely to be affected. For agriculture, it should identify crop areas, planting calendars, drainage needs, and insurance zones. For urban planning, it should support zoning, stormwater investment, and protection of natural retention areas. A technically high AUC map that does not translate into decisions has limited value. Review papers should therefore evaluate validation not only by statistical metrics but also by operational relevance.

The validation hierarchy proposed in this review is shown in Figure 4, moving from weak internal checks toward stronger independent, multi-source evidence

Table 5. Validation approaches for flood hazard and susceptibility maps

Validation method	What it tests	Strength	Main weakness	Recommended practice
ROC-AUC	Discrimination between flood and non-flood samples	Comparable across models; widely reported	Can be inflated by spatial autocorrelation and easy negatives	Use with spatial or event-based test sets
Confusion matrix	Class-level correctness at selected threshold	Provides precision, recall, F1 and errors	Threshold dependent; class imbalance sensitive	Report balanced metrics and decision threshold
Prediction-rate curve	Predictive fit against validation floods	Useful for susceptibility-class evaluation	Depends on inventory quality	Use independent inventory where possible
Historical flood extent	Agreement with past observed floods	Operationally intuitive	May reflect only past event conditions	Use multi-event and temporally separated validation
SAR-derived flood extent	Cloud-independent event validation	Strong for monsoon and emergency contexts	Urban/vegetation ambiguity; independence issue if used for training	Use separate SAR events for validation
Field survey / GPS	Ground-level verification	High credibility for local studies	Limited spatial coverage and access bias	Combine with satellite and official records
Hydraulic model comparison	Physical scenario comparison	Can include depth and velocity	Requires model assumptions and calibration	Use as complementary evidence
Stakeholder/sensor validation	Decision relevance and local credibility	Connects maps with operational reality	May be incomplete or qualitative	Triangulate with quantitative metrics

9. APPLICATIONS IN DISASTER MANAGEMENT, AGRICULTURE, URBAN PLANNING AND CLIMATE ADAPTATION

Flood hazard and susceptibility maps have value only when connected to decisions. In disaster management, they support early warning, evacuation planning, shelter siting, emergency-route analysis, pre-positioning of relief materials, critical-infrastructure protection, and public awareness. Remote sensing contributes timely event information, while GIS integrates hazard layers with roads, settlements, hospitals, schools, power infrastructure, and communication networks. The operational challenge is ensuring that maps are understandable, updated, and trusted by agencies before disasters occur.

In agriculture, flood susceptibility maps can inform crop planning, drainage management, seed selection, contingency cropping, crop insurance, input distribution, and post-flood recovery. Floods affect agriculture through waterlogging, soil erosion, sediment deposition, nutrient loss, pest and disease outbreaks, delayed sowing, crop lodging, livestock loss, and damage to irrigation infrastructure. Remote sensing can monitor crop extent, flood duration, and post-flood vegetation recovery. GIS can overlay susceptibility zones with crop maps, soil types, irrigation canals, and village boundaries. This agricultural application deserves more attention because many flood studies focus on urban assets while rural livelihoods remain underrepresented.

Urban planning is a major application because flood susceptibility is strongly shaped by built environment. Impervious surfaces increase runoff; road embankments can block drainage; informal settlements may occupy low-lying or reclaimed land; and stormwater systems may be undersized for changing rainfall extremes. Urban flood susceptibility mapping should include high-resolution DEMs, drainage networks, imperviousness, building density, depressions, road connectivity, and historical waterlogging points. Islam et al. (2025) reviewed urban flood susceptibility mapping and emphasized the increasing role of remote sensing, machine learning, and modeling approaches. The urban context shows why susceptibility mapping must evolve beyond simple river-distance logic.

Infrastructure planning requires identifying roads, bridges, culverts, power stations, schools, hospitals, and water-supply systems exposed to flooding. GIS-based overlay can prioritize retrofitting, embankment strengthening, culvert enlargement, drainage redesign, and maintenance. However, infrastructure decisions require more than relative susceptibility. Engineers often need flood depth, velocity, return period, scour potential, debris load, and failure consequences. Therefore, susceptibility maps should be treated as screening tools that guide detailed hydrological and hydraulic studies, not as replacements for engineering design.

Climate adaptation planning is another important use. Climate change can alter rainfall intensity, storm tracks, snowmelt timing, sea-level conditions, and compound flooding. Static susceptibility maps based only on historical conditions may underestimate future hazards. Integrating land-use scenarios, rainfall projections, sea-level rise, drainage development, and urban growth can help produce future susceptibility or scenario maps. However, scenario mapping introduces additional uncertainty. A professional review should argue for transparent scenario assumptions rather than deterministic claims about future flood zones.

Community-based planning can improve both data quality and map acceptance. Local residents often know recurrent waterlogging points, blocked drains, historical flood depths, evacuation constraints, and socially vulnerable households. Combining community knowledge with remote sensing and GIS can reduce blind spots, especially for small, recurrent, pluvial floods that may not appear in official databases. Feizbahr et al. (2025) included local community task-force information as part of validation, illustrating how technical and local evidence can be combined. The future of flood mapping should not be purely satellite-driven; it should be participatory where decisions affect local communities.

Insurance and compensation schemes can use susceptibility maps for risk zoning, premium design, loss verification, and targeting of support. In agricultural insurance, satellite-derived flood duration and crop-condition indices can help verify damage. However, ethical risks arise if susceptibility maps are used to deny services or increase premiums without considering social vulnerability and adaptation capacity. Therefore, flood maps should be used with transparent methods, stakeholder review, and safeguards against penalizing vulnerable communities for living in historically marginalized locations.

Environmental management also benefits from flood maps. Floodplains, wetlands, riparian buffers, and natural retention areas can reduce flood peaks and store water. Remote sensing can monitor wetland loss, river encroachment, sand mining, floodplain agriculture, and urban expansion into drainage corridors. GIS can identify areas where nature-based solutions, detention basins, floodplain restoration, or green infrastructure would be most effective. This is an important future direction because structural flood control alone is often insufficient under rapid urbanization and climate extremes.

10. CRITICAL LIMITATIONS AND COMMON METHODOLOGICAL MISTAKES

The first common mistake is conceptual overclaiming. Many papers call the final output a flood risk map even when it contains only physical conditioning factors. Unless exposure and vulnerability are included, the output is not risk. Similarly, a susceptibility map should not be interpreted as flood depth or return period. This distinction is not semantic; it determines whether the map can be used for zoning, engineering design, insurance, or emergency planning. Reviewers should reject manuscripts that use risk terminology without risk components.

The second mistake is weak flood inventory construction. Supervised models are only as reliable as the inventory used to train them. Inventories derived from historical reports may be biased toward populated or accessible areas. Inventories derived from SAR may miss flooded vegetation, urban floodwater, or shallow water. Inventories from one event may represent a specific rainfall pattern rather than general susceptibility. Studies should report inventory date, source, preprocessing, permanent water masking, sampling strategy, and uncertainty. A flood inventory should be treated as a research product, not a routine input.

The third mistake is arbitrary factor selection. Authors often include elevation, slope, rainfall, LULC, soil, geology, drainage density, distance to river, NDVI, TWI, and road distance because earlier studies used them. This copy-and-paste factor selection ignores local hydrology. A coastal city, mountain catchment, alluvial floodplain, semi-arid basin, and deltaic agricultural region do not need identical predictors. Factor selection should be justified by flood mechanism, data quality, scale, and availability. More variables do not automatically make a better model.

The fourth mistake is spatial-scale mismatch. Flood processes can be controlled by micro-topography, drainage structures, and local obstructions, while many input datasets are coarse. A 30 m DEM may be acceptable for regional screening but inadequate for street-level urban flooding. Soil and geology maps may be coarser than the susceptibility output. Rainfall products may miss convective storm variability. If the input scale is coarse, the final map should not claim fine-scale precision. A professional paper should align claims with data resolution.

The fifth mistake is excessive reliance on random train-test splits. Random splitting at pixel or point level can put spatially neighboring samples into training and testing sets, inflating performance because the model is tested on conditions similar to what it has already seen. Spatial cross-validation, blocked validation, or independent event testing should become standard. This is especially important for machine-learning models that can exploit spatial autocorrelation.

The sixth mistake is using AUC as the only accuracy metric. AUC is helpful but insufficient. A flood map used for planning must also be evaluated in terms of recall, precision, false negatives, class distribution, calibration, and alignment with independent flood evidence. High AUC does not tell whether the high-susceptibility zone is too large for practical intervention. It also does not reveal whether the model is well calibrated. Accuracy reporting should be connected to decision thresholds and error costs.

The seventh mistake is lack of interpretability. Machine-learning models can generate high-performing maps that are difficult to explain. If a model assigns high susceptibility to areas for reasons that contradict hydrological logic, decision-makers may distrust the output or, worse, use it uncritically. Explainable-AI tools help, but they do not replace domain interpretation. SHAP, feature importance, and partial dependence should be used to check whether model behavior is physically plausible.

The eighth mistake is insufficient uncertainty communication. Many maps show crisp classes and authoritative colors without conveying confidence. Input uncertainty, model uncertainty, and classification uncertainty should be disclosed. This is particularly important when maps guide land-use restrictions or infrastructure investment. Decision-makers need to know where predictions are robust and where additional field verification is required.

The ninth mistake is ignoring temporal dynamics. Land use changes, river channels shift, drainage systems are modified, embankments are built or breached, and rainfall extremes evolve. A susceptibility map produced from old land-cover data and historical flood inventory may quickly become outdated. Recurrent updating using satellite time series, new flood events, and local observations is necessary. Static maps should be dated and revised, not treated as permanent truths.

The tenth mistake is weak reproducibility. Many studies do not provide code, thresholds, sampling details, model parameters, or data-processing steps. With cloud platforms and open data, reproducible workflows are increasingly possible. Reviewers and journals should require clear reporting of data sources, preprocessing, factor derivation, sampling, hyperparameters, validation splits, and map classification. Without reproducibility, the study is difficult to trust or build upon.

11. FUTURE DIRECTIONS

The first future direction is multi-source flood inventory generation. SAR, optical imagery, UAV observations, official flood records, social media reports, field GPS points, water-level sensors, and community knowledge should be triangulated. No single source captures all flooding. SAR may miss urban or vegetated water; optical imagery may be cloud-limited; official reports may miss small events; community reports may be qualitative. Combining sources can improve inventory completeness and reduce bias.

The second direction is dynamic susceptibility mapping. Many current maps are static representations of terrain and long-term averages. Flood susceptibility, however, changes with antecedent soil moisture, reservoir operations, land-cover change, drainage blockage, storm trajectory, and seasonal cropping. Dynamic maps that update with rainfall forecasts, soil moisture, river levels, and recent satellite observations would be more useful for early warning. This requires integration of remote sensing, GIS, hydrological forecasting, and operational decision systems.

The third direction is stronger integration between susceptibility modeling and hydrodynamic modeling. Susceptibility models are efficient for screening, while hydraulic models provide flood depth, velocity, duration, and scenario interpretation. Hybrid workflows could use susceptibility maps to identify priority zones for detailed hydraulic simulation or use hydraulic outputs to train machine-learning models across larger areas. This would reduce the false choice between empirical GIS mapping and process-based modeling.

The fourth direction is explainable and physics-informed GeoAI. Deep learning and ensemble models should not be adopted only because they are fashionable. They should incorporate hydrological prior knowledge, terrain constraints, drainage connectivity, mass-balance logic, and uncertainty estimates. Wang et al. (2025) represents the emerging direction of accelerated and interpretable flood susceptibility mapping using explainable deep learning with hydrological prior knowledge. Such approaches are promising because they attempt to combine predictive strength with process credibility.

The fifth direction is better validation against independent events and real decisions. Future papers should test models on temporally separated floods, spatially independent areas, and operational datasets such as flood sensors, insurance claims, emergency reports, or field surveys. Validation should also ask whether maps help decision-makers allocate resources, improve warnings, reduce losses, or protect vulnerable populations. Technical accuracy without decision relevance is insufficient.

The sixth direction is high-resolution urban and peri-urban flood mapping. Urban flooding is often controlled by drainage infrastructure, road grades, building footprints, culverts, and micro-topography. Coarse DEMs and generic LULC classes are inadequate. LiDAR DEMs, UAV imagery, high-resolution SAR, street-level reports, drainage network data, and urban hydrological models should be integrated. Urban susceptibility reviews show that remote sensing and modeling are advancing, but local drainage data remain a bottleneck.

The seventh direction is agricultural and livelihood-sensitive flood mapping. Many flood-prone regions are rural, agrarian, and economically vulnerable. Future studies should overlay susceptibility with crop calendars, crop type, irrigation infrastructure, livestock locations, rural roads, grain storage, and market access. This would make maps more useful for agricultural extension, contingency planning, crop insurance, and livelihood resilience. The agricultural dimension is still weaker than the urban-infrastructure dimension in much of the geospatial flood literature.

The eighth direction is uncertainty-aware communication. Instead of presenting only final class maps, researchers should provide probability maps, confidence layers, sensitivity results, and limitations. Map legends should avoid giving false precision. Decision-makers should be told which areas require field verification and which model assumptions matter most. This is especially important when maps are used for regulatory decisions or public communication.

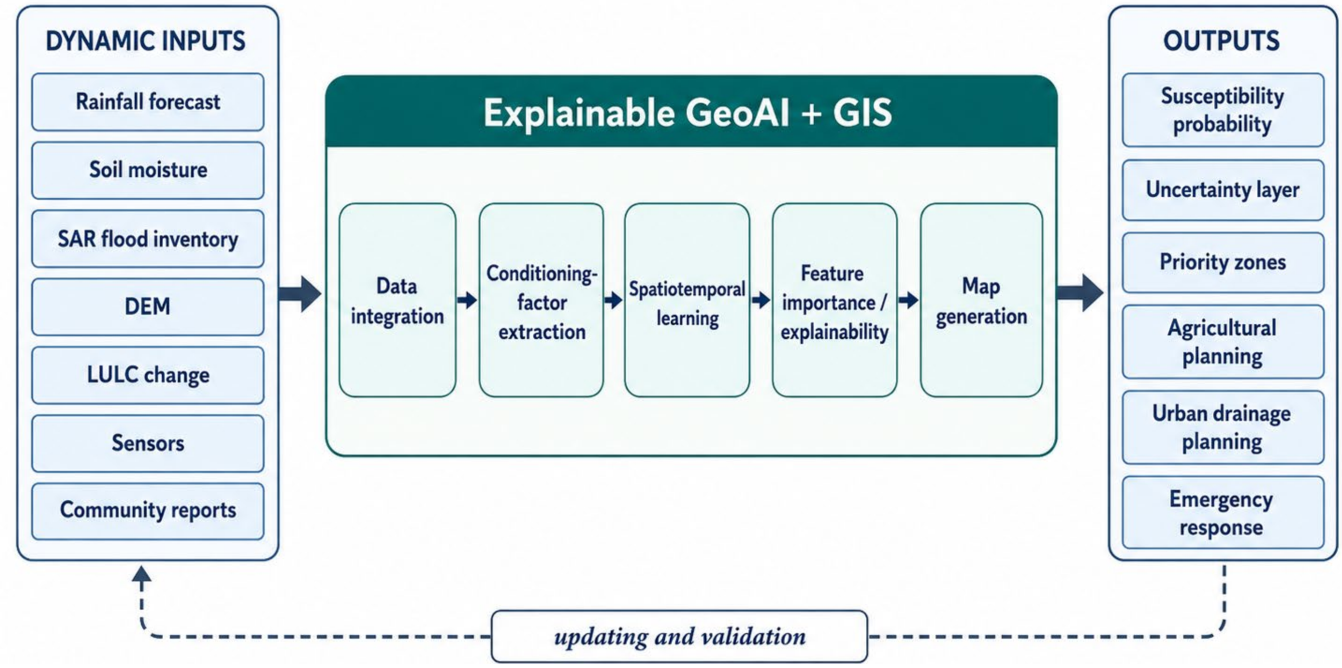

Figure 5. Future integrated framework for dynamic, explainable and decision-oriented flood susceptibility mapping

The ninth direction is reproducible, open, and interoperable workflow design. Studies using open satellite data, open DEMs, open rainfall products, and cloud platforms should share scripts or at least detailed workflow steps. Reproducibility allows comparison across basins, adaptation to new events, and operational updating. It also reduces the risk of one-off maps that cannot be maintained after publication.

The final direction is ethical use. Flood maps can influence property values, insurance, relocation, compensation, and public investment. They should be technically sound, transparent, and socially responsible. Communities should not be labeled high-risk without pathways for mitigation, support, and adaptation. Remote sensing and GIS provide powerful evidence, but evidence must be used with institutional accountability and local engagement.

A future-oriented framework for dynamic, explainable and decision-oriented flood susceptibility mapping is outlined in Figure 5.

Table 6. Research gaps and future priorities in remote sensing and GIS-based flood mapping

Current gap	Why it matters	Recommended future direction
Weak conceptual separation of hazard, susceptibility, vulnerability and risk	Leads to overclaiming and inappropriate use of maps	Use precise terminology and match output labels to input data
Incomplete or biased flood inventories	Training labels determine model reliability	Build multi-source, multi-event inventories using SAR, field, official and local data
Overreliance on AUC and random splits	Inflates model performance and hides transferability problems	Use spatial blocking, event-based validation and multiple metrics
Limited uncertainty reporting	Users may treat class boundaries as exact	Provide confidence layers, ensemble ranges and sensitivity analysis
Limited urban micro-topography and drainage data	Urban floods are controlled by local infrastructure	Integrate LiDAR, drainage networks, UAVs and local waterlogging records
Underdeveloped agricultural applications	Flood impacts on crops and livelihoods are often neglected	Overlay susceptibility with crop calendars, rural infrastructure and livelihood data
Black-box machine learning	Reduces trust and may hide nonphysical model behavior	Apply SHAP, ALE, partial dependence and hydrological plausibility checks
Static maps in changing landscapes	Land use and climate change alter flood patterns	Develop updateable and scenario-based susceptibility systems
Poor reproducibility	Limits comparison and operational adoption	Publish workflows, parameters and data-processing steps

12. PROPOSED BEST-PRACTICE WORKFLOW FOR FUTURE STUDIES

A rigorous flood hazard or susceptibility mapping study should begin with problem definition. The authors must state whether they are mapping observed inundation, susceptibility, hazard, exposure, vulnerability, or risk. They should specify the flood type: riverine, pluvial, flash flood, coastal, compound, urban drainage, agricultural waterlogging, or glacial-lake outburst. The flood type determines the conditioning factors, data sources, and validation strategy. Without this first step, the rest of the workflow becomes generic.

The second step is data audit. Researchers should list all datasets with source, date, spatial resolution, temporal resolution, projection, preprocessing, and known limitations. DEM, rainfall, LULC, soil, drainage, roads, flood inventory, and validation data should be documented. If datasets have different resolutions, the resampling method and target grid should be justified. If land-cover data are older than the flood inventory, the mismatch should be disclosed.

The third step is flood inventory construction. For SAR-based inventories, authors should describe preprocessing, speckle filtering, polarization, thresholding or classification method, permanent water masking, terrain correction, and validation. For historical inventories, they should report sources, dates, uncertainty, and digitization method. For field inventories, they should describe survey design and GPS accuracy. The inventory should be divided into training and validation in a way that reduces spatial and temporal leakage.

The fourth step is factor selection and derivation. Factors should be selected based on flood mechanism and data quality, not copied blindly from previous studies. Derivation methods for slope, TWI, flow accumulation, drainage density, distance measures, indices, and rainfall summaries should be reported. Correlation and redundancy should be checked. If AHP is used, pairwise matrices and consistency ratios should be provided. If machine learning is used, feature scaling, encoding, and selection should be explained.

The fifth step is comparative modeling. At minimum, one transparent baseline method should be compared with more complex models. For example, AHP or frequency ratio can be compared with random forest and XGBoost. This prevents unsupported claims that a single method is best. Hyperparameter tuning should be performed only on training or validation subsets, not on the final test set. Class imbalance should be addressed through appropriate sampling, weighting, or metrics.

The sixth step is validation. Internal accuracy metrics should be supplemented with independent flood evidence. AUC, precision, recall, F1-score, prediction-rate curves, class-area distribution, and independent event overlays should be reported. If possible, validation should be spatially blocked or event based. The paper should explicitly discuss false positives and false negatives because these errors have different practical consequences.

The seventh step is interpretation. Variable importance, SHAP, partial dependence, or sensitivity analysis should be used to check whether the model behaves in hydrologically plausible ways. If the model identifies road distance as the dominant factor, the author should explain whether roads obstruct drainage, indicate built-up areas, or simply correlate with observed reporting. Interpretation must connect statistics back to physical and social processes.

The eighth step is communication. Final maps should include clear legends, class definitions, scale, projection, data date, uncertainty notes, and appropriate terminology. A susceptibility map should not be sold as a risk map. Decision recommendations should match the map resolution and validation strength. For local planning, field verification should be recommended before major investment or relocation decisions.

The ninth step is reproducibility. Authors should share scripts, model parameters, preprocessing steps, and where possible, derived non-sensitive data. Cloud-based scripts can be archived or described in detail. Reproducibility is not a luxury; it is necessary for flood mapping because maps must be updated after new events, land-use changes, and improved datasets.

13. CONCLUSION

Remote sensing and GIS have fundamentally reshaped flood hazard and susceptibility mapping by making it possible to integrate terrain, rainfall, drainage, land cover, soil, geology, infrastructure, and historical flood evidence in spatially explicit workflows. The field has progressed from manual interpretation and simple overlays to cloud-based SAR inventories, multi-criteria analysis, statistical learning, machine learning, deep learning, and explainable GeoAI. This progression has improved the ability to identify flood-prone zones, but it has also created new risks of overclaiming, black-box modeling, and weak validation.

The review shows that no single method is universally superior. GIS-MCDA and AHP remain useful for transparent preliminary mapping in data-poor contexts. Frequency ratio, weights of evidence, and logistic regression provide interpretable empirical alternatives when flood inventories exist. Random forest, XGBoost, SVM, CNNs, and deep learning can improve predictive performance when high-quality training data and appropriate validation are available. SAR, especially Sentinel-1, is now central to flood inventory generation and event monitoring because of its all-weather capability, but it still faces challenges in urban and vegetated environments.

The most important conclusion is methodological rather than technological: a strong flood susceptibility study requires clear terminology, physically meaningful conditioning factors, reliable inventories, appropriate spatial scale, independent validation, uncertainty reporting, and decision-oriented communication. Future research should move toward dynamic, multi-source, explainable, reproducible, and ethically used geospatial flood-mapping systems. Such systems should support not only technical hazard assessment but also agricultural resilience, urban planning, infrastructure protection, community preparedness, and climate adaptation.

Conflicts of Interest: The authors declare no conflicts of interest.

Funding: This research received no external funding.
Acknowledgments: NA

Author Contributions: The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and preparation of the manuscript.

AI Disclosure: The authors declare that artificial intelligence (AI) tools were used solely to assist in language refinement, organization, and editing of the manuscript. AI was not used to generate original research data, conduct analyses, or draw scientific conclusions. All technical content, interpretations, and final decisions were developed and verified by the authors, who take full responsibility for the accuracy, integrity, and originality of the work.

REFERENCES

Addis, A. (2023). GIS-based flood susceptibility mapping using frequency ratio and information value models in Upper Abay River Basin, Ethiopia. Natural Hazards Research, 3(2), 247-256. https://doi.org/10.1016/j.nhres.2023.02.003

Al-Omari, A. A., Shatnawi, N. N., Shbeeb, N. I., Istrati, D., Lagaros, N. D., & Abdalla, K. M. (2024). Utilizing remote sensing and GIS techniques for flood hazard mapping and risk assessment. Civil Engineering Journal, 10(5), 1423-1436. https://doi.org/10.28991/CEJ-2024-010-05-05

Alonso-Sarria, F., Valdivieso-Ros, C., & Molina-Perez, G. (2025). Detecting flooded areas using Sentinel-1 SAR imagery. Remote Sensing, 17(8), 1368. https://doi.org/10.3390/rs17081368

Amitrano, D., Di Martino, G., Iervolino, P., Riccio, D., Ruello, G., & Papa, M. N. (2024). Flood detection with SAR: A review of techniques and datasets. Remote Sensing, 16(4), 656. https://doi.org/10.3390/rs16040656

Coltin, B., McMichael, S., Smith, T., & Fong, T. (2016). Automatic boosted flood mapping from satellite data. International Journal of Remote Sensing, 37(5), 993-1015. https://doi.org/10.1080/01431161.2016.1145366

Costache, R., Pham, Q. B., Sharifi, E., Linh, N. T. T., Abba, S. I., Vojtek, M., Vojtekova, J., Nhi, P. T. T., & Khoi, D. N. (2020). Flash-flood susceptibility assessment using multi-criteria decision making and machine learning supported by remote sensing and GIS techniques. Remote Sensing, 12(1), 106. https://doi.org/10.3390/rs12010106

Feizbahr, M., Brake, N., Arbabkhah, H., Hariri Asli, H., & Woods, K. (2025). Flood susceptibility mapping using machine learning and geospatial-Sentinel-1 SAR integration for enhanced early warning systems. Remote Sensing, 17(20), 3471. https://doi.org/10.3390/rs17203471

Hamidi, E., Peter, B. G., Munoz, D. F., Moftakhari, H., & Moradkhani, H. (2023). Fast flood extent monitoring with SAR change detection using Google Earth Engine. IEEE Transactions on Geoscience and Remote Sensing, 61, 1-19. https://doi.org/10.1109/TGRS.2023.3240097

Hitouri, S., Mohajane, M., Lahsaini, M., Ali, S. A., Setargie, T. A., Tripathi, G., D'Antonio, P., Singh, S. K., & Varasano, A. (2024). Flood susceptibility mapping using SAR data and machine learning algorithms in a small watershed in northwestern Morocco. Remote Sensing, 16(5), 858. https://doi.org/10.3390/rs16050858

Islam, T., Zeleke, E. B., Afroz, M., Islam, A. R. M. T., & Rahman, M. M. (2025). A systematic review of urban flood susceptibility mapping: Remote sensing, machine learning, and other modeling approaches. Remote Sensing, 17(3), 524. https://doi.org/10.3390/rs17030524

Kader, Z., Islam, M. R., Aziz, M. T., Hossain, M. M., Islam, M. R., Miah, M., & Wan Jaafar, W. Z. (2024). GIS and AHP-based flood susceptibility mapping: A case study of Bangladesh. Sustainable Water Resources Management, 10, 170. https://doi.org/10.1007/s40899-024-01126-0

Klemas, V. (2015). Remote sensing of floods and flood-prone areas: An overview. Journal of Coastal Research, 31(4), 1005-1013. https://doi.org/10.2112/JCOASTRES-D-14-00160.1

Mangkhaseum, S., Bhattarai, Y., Duwal, S., & Hanazawa, A. (2024). Flood susceptibility mapping leveraging open-source remote-sensing data and machine learning approaches in Nam Ngum River Basin (NNRB), Lao PDR. Geomatics, Natural Hazards and Risk, 15(1), 2357650.

Megahed, H. A., Abdo, A. M., AbdelRahman, M. A., Scopa, A., & Hegazy, M. N. (2023). Frequency ratio model as tools for flood susceptibility mapping in urbanized areas: A case study from Egypt. Applied Sciences, 13(16), 9445.

Misra, A., Gorelick, N., Reiche, J., & colleagues. (2025). Mapping global floods with 10 years of satellite radar data. Nature Communications, 16, 5604. https://doi.org/10.1038/s41467-025-60973-1

Mohammed, Z. T., Hussein, L. Y., & Abood, M. H. (2024). Potential flood hazard mapping based on GIS and analytical hierarchy process.

Journal of Water Management Modeling. https://doi.org/10.14796/JWMM.C528

Munasinghe, D., Cohen, S., Huang, Y.-F., Tsang, Y.-P., Zhang, J., & Fang, Z. (2018). Intercomparison of satellite remote sensing-based flood inundation mapping techniques. Journal of the American Water Resources Association, 54(4), 834-846. https://doi.org/10.1111/1752-1688.12626

Pimenta, L., Duarte, L., Teodoro, A. C., Beltrão, N., Gomes, D., & Oliveira, R. (2025). GIS-Based Flood Susceptibility Mapping Using AHP in the Urban Amazon: A Case Study of Ananindeua, Brazil. Land, 14(8), 1543.

Pradhan, B., Lee, R., Dikshit, A., & Kim, H. (2023). Spatial flood susceptibility mapping using an explainable artificial intelligence model. Geoscience Frontiers, 14(6), 101625. https://doi.org/10.1016/j.gsf.2023.101625

Rahman, Z. U., Zhang, M., Chen, F., Ullah, S., Ahmad, M., Feroz, A., ... & Azzam, A. (2025). Flood susceptibility mapping using supervised machine learning models: insights into predictors' significance and models' performance. Geomatics, Natural Hazards and Risk, 16(1), 2516728.

Rahmati, O., Pourghasemi, H. R., & Zeinivand, H. (2016). Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golestan Province, Iran. Geocarto International, 31(1), 42-70. https://doi.org/10.1080/10106049.2015.1041559

Schumann, G. J.-P. (2024). Breakthroughs in satellite remote sensing of floods. Frontiers in Remote Sensing, 4, 1280654. https://doi.org/10.3389/frsen.2023.1280654

Sellami, E. M., Ben Khalifa, A., & others. (2024). Google Earth Engine and machine learning for flash flood susceptibility mapping. Geosciences, 14(6), 152. https://doi.org/10.3390/geosciences14060152

Seydi, S. T., Kanani-Sadat, Y., Hasanlou, M., Sahraei, R., Chanussot, J., & Amani, M. (2023). Comparison of machine learning algorithms for flood susceptibility mapping. Remote Sensing, 15(1), 192. https://doi.org/10.3390/rs15010192

Swain, K. C., Singha, C., & Nayak, L. (2020). Flood susceptibility mapping through the GIS-AHP technique using the cloud. ISPRS International Journal of Geo-Information, 9(12), 720. https://doi.org/10.3390/ijgi9120720

Taoukidou, N., Koutalakis, P., Zaimes, G. N., & others. (2025). Flood hazard assessment through AHP, fuzzy AHP, and frequency ratio methods: A comparative analysis. Water, 17(14), 2155. https://doi.org/10.3390/w17142155

Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2014). Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. Journal of Hydrology, 512, 332-343. https://doi.org/10.1016/j.jhydrol.2014.03.008

Wang, J., Sanderson, J., Iqbal, S., & Woo, W. L. (2025). Accelerated and interpretable flood susceptibility mapping through explainable deep learning with hydrological prior knowledge. Remote Sensing, 17(9), 1540. https://doi.org/10.3390/rs17091540

Wangchuk, S., Bolch, T., & Robson, B. A. (2022). Monitoring glacial lake outburst flood susceptibility using Sentinel-1 SAR data, Google Earth Engine, and persistent scatterer interferometry. Remote Sensing of Environment, 271, 112910. https://doi.org/10.1016/j.rse.2022.112910

Zhao, G., Pang, B., Xu, Z., Peng, D., & Xu, L. (2019). Assessment of urban flood susceptibility using semi-supervised machine learning model. Science of the Total Environment, 659, 940-949. https://doi.org/10.1016/j.scitotenv.2018.12.217