Mapping the Scientific Evolution of Flood Risk Modelling with GIS: A Bibliometric and Thematic Analysis (1987–2024)

Jagtap, A. A.,[1,2*] Mane, P. B.[1] and Shedge, D. K.[1]
[1]Department of Electronics and Telecommunication, AISSMS Institute of Information Technology, Savitribai Phule Pune University, Pune, India
E-mail: anjalijagtap2306@gmail.com,* pbmane6829@gmail.com, dnyandeo.shedge@aissmsioit.org
[2]Department of Electronics and Telecommunication, Hope Foundation's International Institute of Information Technology, Savitribai Phule Pune University, Pune, India, E-mail: anjalij@isquareit.edu.in
Corresponding Author

DOI: https://doi.org/10.52939/ijg.v22i2.4791

Abstract
Flood risk modeling using Geographic Information Systems (GIS) has gained increasing global significance amid intensifying climate change and extreme weather events. This study presents a comprehensive bibliometric and thematic analysis of GIS-based flood risk modeling research published between 1987 and 2024. A total of 3,946 records were retrieved from the Scopus database and analyzed using the Bibliometrix R package to examine publication trends, authorship patterns, institutional collaboration, and thematic evolution. Results indicate a 15.08% annual growth rate in scientific output, with a pronounced surge after 2010 driven by advances in remote sensing technologies, global policy frameworks such as the Sendai Framework, and the rise of climate adaptation initiatives. China and India dominate in publication volume, whereas Italy and Malaysia show higher citation impacts, reflecting methodological depth and innovation. Thematic evolution analysis reveals a shift from traditional floodplain mapping to emerging themes such as AI-driven flood forecasting, real-time hydrological monitoring, and resilience assessment. Despite this progress, significant gaps persist in socio-economic integration, uncertainty quantification, and model interpretability. Based on these identified gaps, this study highlights the potential of integrating Machine Learning (ML), explainable AI techniques (e.g., SHAP), and Conformal Prediction (CP) as emerging technologies to enhance the predictive power, transparency, and confidence estimation of GIS-based flood risk models.

Keywords: AHP, Bibliometric Analysis, Climate Resilience, Conformal Prediction, Explainable AI, Flood Risk, GIS, Machine Learning, SDGs, SHAP

1. Introduction

Floods are among the most frequent and destructive natural hazards globally, responsible for extensive socio-economic damage, environmental degradation, and disruption to infrastructure and livelihoods. The increasing frequency and intensity of flood events, amplified by climate change, erratic rainfall, and rapid urbanization, demand robust flood risk modeling frameworks for effective disaster risk reduction and sustainable land-use planning, as emphasized by [1]. Over the past two decades, Geographic Information Systems (GIS) have become indispensable in flood modeling, offering spatial analysis, data integration, and hydrodynamic simulation capabilities. When integrated with hydrological models such as HEC-RAS and SWAT, GIS enables high-resolution floodplain delineation and real-time inundation prediction, as demonstrated by [2] and [3]. The incorporation of remote sensing datasets including LiDAR, radar, and multispectral imagery further enhances spatial accuracy and monitoring capabilities [4] and [5].

In recent years, Machine Learning (ML) has emerged as a transformative component in flood modeling. Algorithms such as Random Forest (RF), Support Vector Machines (SVM), and Neural Networks (NN) have demonstrated high predictive performance in flood susceptibility mapping by integrating meteorological, topographic, and land-use variables [6][7] and [8]. Hybrid models that combine ML with GIS and remote sensing such as RF–AHP, fuzzy–AHP, and Bayesian–AHP are increasingly used across diverse geographies including India [1][2] and [6], Nepal [9], Vietnam [10], and Nigeria [11].

Alongside these developments, Multi-Criteria Decision-Making (MCDM) tools particularly the Analytical Hierarchy Process (AHP) have provided structured and transparent approaches for flood hazard ranking and parameter weighting [12][13][14][15] and [16]. Advanced techniques such as CoCoSo and VIKOR have been introduced as alternative MCDM approaches for multi-hazard assessment [12] and [17], while entropy-based models contribute objectivity to weight assignment [18] and [19]. Recent studies have begun addressing model interpretability and uncertainty quantification long-standing limitations of conventional GIS–ML approaches. The integration of explainable AI techniques such as SHAP (Shapley Additive Explanations) allows identification of variable importance and enhances model transparency [20] [21] and [22]. Similarly, Bayesian and probabilistic approaches have been introduced to quantify confidence and uncertainty in flood susceptibility mapping [23].

Urban and coastal flood vulnerability mapping continues to be a prominent research frontier Studies by [18] [24] [25] and [26] emphasize the influence of impervious surfaces, inadequate drainage, and infrastructure density on flood susceptibility in rapidly urbanizing environments. Coastal and deltaic regions have seen increasing application of ensemble and hybrid geo-AI models for multi-hazard assessment [27] and [28]. Geographically, most GIS-based flood modeling studies are concentrated in South and Southeast Asia particularly India [1] [3], Pakistan [7], [17], and Bangladesh [18] reflecting the high exposure to monsoonal variability and land-use pressures. Comparative and methodological contributions from China [4], Egypt [8], Iran [13] and Iraq [29] further highlight regional diversity in modeling approaches.

Despite the rapid growth of literature, the field remains fragmented. Many studies prioritize model performance without addressing transparency, reproducibility, or policy relevance. Deterministic frameworks (e.g., AHP) often lack adaptive learning capacity, while data-driven models (e.g., ML) frequently operate as black boxes without explainability or confidence estimation. Consequently, there is limited understanding of how these methodological strands have evolved and where major research gaps persist. To address this knowledge gap, the present study performs a comprehensive bibliometric and thematic analysis of 3,946 Scopus-indexed publications (1987–2024) related to GIS-based flood risk modeling. Using the Bibliometrix R package and Multiple Correspondence Analysis (MCA), the study systematically maps the intellectual structure, thematic evolution, and methodological trends in this domain. The key objectives are to (i) analyze publication dynamics, collaboration networks, and influential sources; (ii) identify dominant and emerging research themes; and (iii) uncover methodological gaps particularly in explainable AI (SHAP), uncertainty quantification (Conformal Prediction), and hybrid GIS–ML–AHP integration that can guide future research directions.

Through this synthesis, the study provides a data-driven understanding of the global evolution of GIS-based flood modeling and highlights how transparent, interpretable, and uncertainty-aware technologies can advance next-generation flood risk assessment and climate resilience planning.

2. Bibliometric Data and Analytical Methods

Several bibliometric and scientometric reviews have analyzed the evolution of flood-related research from different disciplinary perspectives. Bibliometrics focuses on quantitative publication patterns (e.g., productivity, collaboration, citation), while scientometric emphasizes the intellectual and thematic structures of scientific domains. Both approaches are applied complementarily in this study to assess the global research landscape of GIS-based flood risk modeling.

Analysis of over 3,000 publications from the Web of Science, revealing a thematic transition from hazard-centric flood modeling toward integrated Flood Risk Management (FRM) and governance was carried out by [1]. The most influential contributions originated from the UK and the Netherlands, with a select group of authors and journals disproportionately shaping the discourse. A bibliometric overview of Regional Flood Frequency Analysis (RFFA) using both Scopus and Web of Science datasets was conducted by [2]. Their findings underscored global methodological evolution and advocated for machine learning–based RFFA and improved uncertainty quantification.

Similarly, the 2022 bibliometric study on flash flood susceptibility [30] (305 Web of Science records) revealed three developmental phases early (2007–2015), exploratory (2016–2019), and growth (post-2020) and emphasized the integration of ML and GIS approaches. A region-specific review published in Sustainability [31] assessed GIS and remote sensing-based flood research across South Asia (2004–2024), identifying “remote sensing,” “urban planning,” and “climate resilience” as dominant clusters. A global scientometric evaluation of flood mitigation strategies [11], highlighted the rise of hybrid AHP and ensemble ML models but noting persistent issues

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

of interpretability. Likewise, the Water journal's 2023 review on flood frequency analysis in mountain regions [4] found increasing AI applications for complex terrains but observed minimal focus on model transparency or uncertainty analysis.

While these reviews advance understanding of flood risk research structures, none systematically trace the convergence of expert-driven decision models (AHP), data-driven learning (ML), and explainable AI (e.g., SHAP) or uncertainty-aware methods (e.g., Conformal Prediction). Addressing this methodological gap, the present study performs a comprehensive bibliometric and thematic analysis of GIS-based flood modeling literature to evaluate research growth, collaboration patterns, and thematic transitions.

2.1 Bibliometric Methodology

The bibliometric study was conducted using a four-stage structured workflow. Figure 1 shows conceptual workflow of the bibliometric study, depicting the logical sequence and interconnections among database selection, bibliometric analysis, research trend analysis, and thematic interpretation. The first stage involved database selection and the development of a search strategy to retrieve relevant publications. The second stage focused on data preprocessing, including screening, cleaning, and standardization of bibliographic records to ensure data quality and consistency. The third stage comprised bibliometric analysis, encompassing productivity analysis, citation analysis, network analysis, and thematic mapping to examine publication patterns, influential contributions, and evolving research trends. The final stage involved thematic interpretation and reflection to synthesize the results, identify knowledge gaps, and outline potential directions for future research.

2.1.1 Database and search strategy

All bibliographic records were retrieved from the Scopus database on November 19, 2024. Scopus was chosen due to its broader coverage of engineering, environmental science, and geospatial research compared with Web of Science, along with its metadata compatibility with Bibliometrix R-package and Biblioshiny for advanced analyses. Table 1 shows Scopus search query details used.

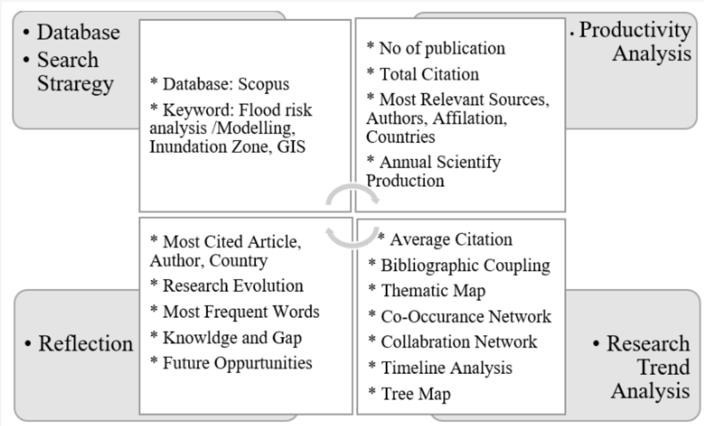

Figure 1: Conceptual workflow of the bibliometric study

Table 1: Scopus search configuration

Component	Query
2000 Keywords	("flood risk" OR "flood hazard" OR "inundation zone") AND ("GIS" OR "Geographic Information System") AND ("mapping")
Document type	Journal articles, conference papers, reviews
Time span	1987 to 2024
Language	English
Boolean query	TITLE-ABS-KEY ("flood* risk" OR "flood* hazard" OR "flood* susceptibility" OR "flood* vulnerability" OR "flood* mapping") AND TITLE-ABS-KEY ("GIS" OR "geographic information system*" OR "remote sensing" OR "geospatial") AND PUBYEAR > 1986 AND PUBYEAR < 2025 AND (LIMIT-TO (LANGUAGE, "English")) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re") OR LIMIT-TO (DOCTYPE, "cp"))

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) |

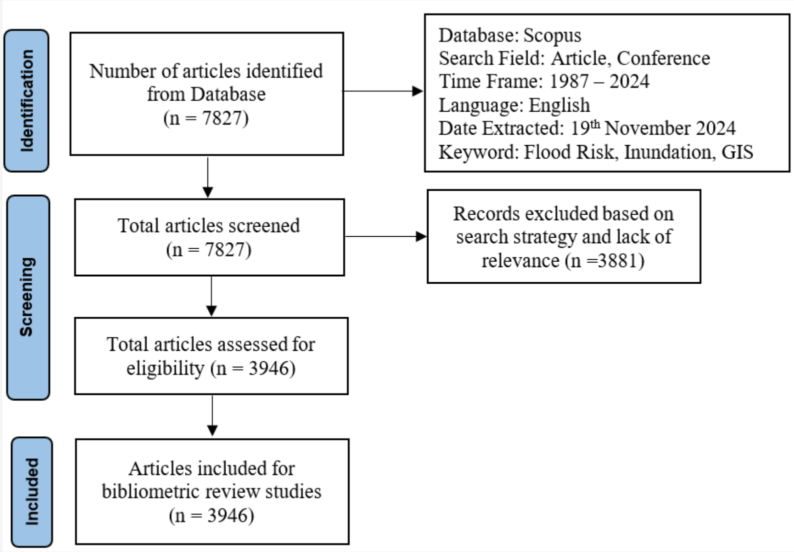

Figure 2: PRISMA chart for bibliometric review

2.1.2 Data filtering and inclusion criteria

The initial query returned 7,827 records. The dataset was refined through automated and manual screening using the following criteria:

- Inclusion:
  - Peer-reviewed studies focusing on flood hazard, risk, or susceptibility mapping using GIS, remote sensing, or spatial ML techniques.
  - English-language publications between 1987–2024.
- Exclusion:
  - Non-flood studies (e.g., groundwater modeling, drought), editorials, theses, non-GIS flood work and non-indexed grey literature.
  - Duplicates removed using duplicated Matching("DOI") in Bibliometrix.

After refinement, 3,946 publications were finalized for analysis. The exclusion process (Figure 2) followed the PRISMA 2020 standard:

2.2 Bibliometric Analysis and Parameters

Data analysis was conducted using R Studio (v4.3) with the Bibliometrix package, while visualization and minor refinements were performed using Microsoft Excel where necessary. To ensure comparability across authors, journals, and countries with uneven publication volumes, association strength normalization was applied, as it effectively reduces scale bias in bibliometric networks. The identification of thematic and collaboration structures was performed using the Louvain modularity clustering algorithm, selected for its ability to efficiently detect cohesive communities in large and heterogeneous networks. Network visualization employed the Fruchterman–Reingold force-directed layout, which spatially separates clusters based on relational intensity and improves interpretability. Keyword co-occurrence analysis was based on author keywords, with a minimum occurrence threshold of five to retain statistically meaningful and frequently recurring research themes while minimizing noise. Citation analysis was conducted using fractional counting to account for multi-authored publications and to avoid inflation of citation influence by highly collaborative works. To capture the temporal evolution of research themes, the analysis was stratified into three periods (1987–2010, 2011–2017, and 2018–2024), reflecting foundational, transitional, and rapid growth phases in flood-related geospatial and machine learning research.

2.3 Analytical Dimensions

- Productivity Analysis: Annual publication growth, citation trajectory, and authorship patterns.
- Source and Author Metrics: Most prolific journals, institutions, and countries.
- Co-authorship Network: Louvain clustering of country and author collaborations.
- Thematic Mapping: Co-word and “Keyword Plus” analysis using Multiple Correspondence Analysis (MCA).
- MCA applied to authors’ keywords using K-means clustering (k = 5).
- Three temporal slices analyzed to visualize topic evolution.
- Intellectual Structure: Co-citation and bibliographic coupling to trace conceptual linkages.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

- Thematic Evolution: Keyword growth and burst detection through Biblioshiny visual timelines.

The selected indicators publication growth rate, total citations, collaboration index, and thematic centrality align with the study's aim to uncover research productivity, intellectual structure, and thematic shifts. Bibliometrix was prioritized for its open-source reproducibility and integration with PRISMA-compliant workflows.

2.4 Reflection and Interpretation

Through this structured workflow, the study ensures transparency, replicability, and methodological rigor in analyzing 3,946 GIS-based flood modeling publications. The integration of MCA and clustering analyses enables identification of thematic transitions such as the movement from deterministic GIS modeling to explainable and uncertainty-aware frameworks thereby providing an empirical foundation for subsequent discussion and future research recommendations.

3. Bibliometric Analysis

The bibliometric evidence demonstrates a rapidly evolving, globally interconnected field with consistent growth and expanding interdisciplinarity. The shift from traditional GIS–hydrological modeling toward AI-driven and explainable flood-risk assessment underscores a paradigm emphasizing transparency, interpretability, and uncertainty quantification. These insights reveal both the maturity of the domain and the opportunity for future integration of AHP–ML–SHAP–CP approaches to advance robust, interpretable flood-modeling frameworks. The bibliometric results are supported by quantitative indicators derived from the Bibliometrix framework, including an average annual growth rate (AGR) of 15.08%, a citation density of 13.28 citations per document, and a collaboration index (CI) of 1.46, indicating a moderately collaborative research field. Thematic strength and maturity were evaluated using Callon's centrality and density metrics, enabling differentiation between motor, basic, niche, and emerging themes.

3.1 Annual Publication Growth

The annual scientific output on GIS-based flood risk modeling exhibits a nonlinear, exponential growth pattern between 1987 and 2024 (Figure 3). While the average annual growth rate across the entire period is 15.08%, this value represents a long-term mean rather than a constant rate. Publication growth remained modest during the foundational phase (1987–2005), accelerated during the transitional phase (2006–2010), and increased sharply after 2010, coinciding with advances in remote sensing, machine learning, and the adoption of global disaster-risk frameworks such as the Sendai Framework. This exponential trajectory indicates a rapidly expanding and maturing research domain rather than steady linear growth. The exponential increase in publication output reflects not only rising flood risk under climate change but also a fundamental methodological transformation in flood modeling research. Early growth was dominated by deterministic GIS–hydrological approaches, whereas the post-2010 surge aligns with the widespread integration of machine learning, high-resolution remote sensing, and multi-criteria decision analysis. Similar exponential trends have been reported in broader flood-risk and climate-adaptation bibliometric studies, indicating that GIS-based flood modeling has evolved from a niche technical domain into a core component of sustainability and disaster-resilience research.

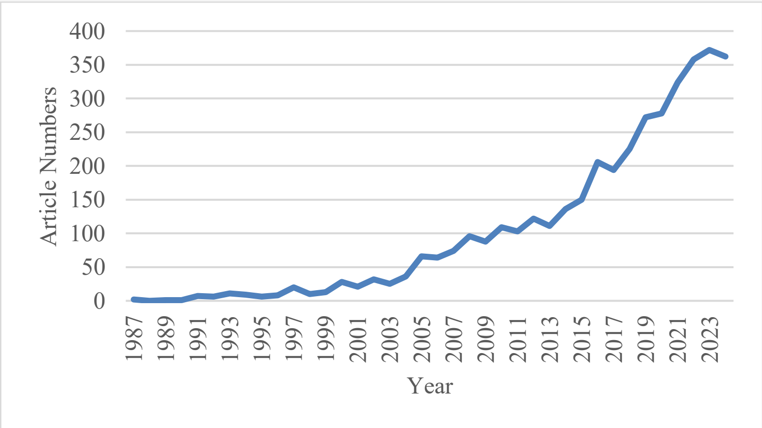

Figure 3: Annual growth of publications (1987–2024)

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

The pronounced increase in publications after 2010 further coincides with the adoption of major global policy frameworks, including the Sendai Framework for Disaster Risk Reduction (2015) and the Paris Agreement (2015), as well as the growing availability of open-access geospatial datasets (e.g., MODIS, CHIRPS, and Copernicus). These policy and technological drivers facilitated large-scale comparative analyses and international collaboration, reinforcing flood-risk modeling as a critical pillar of climate-adaptation and sustainable development research.

3.2 Citation Performance and Research Visibility

The average citation density for the 3,946 publications is 13.28 citations per document, indicating moderate-to-high impact (Figure 4). Foundational studies before 2005 received steady citations as conceptual baselines for flood mapping and hydrological modeling. The 2010–2020 decade shows peak citation intensity corresponding with adoption of machine learning (ML) and hybrid GIS–AHP techniques. Recent papers (post-2021) have not yet reached full citation potential due to natural time-lag effects, but those involving AI, ensemble approaches, and uncertainty analysis are likely to become future citation leaders.

3.3 Source and Author Productivity

Table 2 lists the top 10 publication sources. Natural Hazards (160 articles) and Water (Switzerland) (130) lead the field, showing the strong interdisciplinary nature of flood-related research spanning hydrology, sustainability, and geospatial science. Journal of Hydrology and International Journal of Disaster Risk Reduction further demonstrate integration between scientific modeling and policy frameworks. Leading contributors include B. Pradhan, J. Zhang, Y. Chen, and Y. Wang, who collectively advanced GIS–ML integration for flood susceptibility mapping. The collaboration index (CI), average authors per multi-authored paper, is 1.46, reflecting a moderately collaborative but growing research community. Increased co-authorship post-2010 mirrors the rise of interdisciplinary data-driven projects.

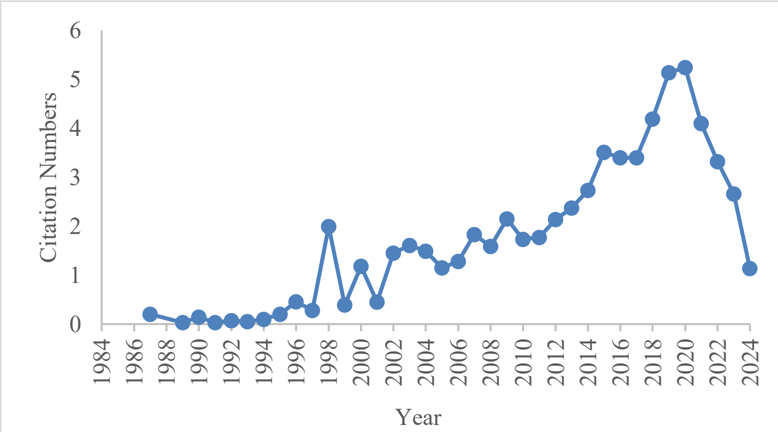

Figure 4: Average citations per publication (1987–2024)

Table 2: Top 10 sources in GIS-based flood-risk modeling

Sources	Article Numbers
Natural Hazards	160
Water (Switzerland)	130
Iop Conference Series: Earth and Environmental Science	84
Journal Of Hydrology	60
International Journal of Disaster Risk Reduction	58
Sustainability (Switzerland)	58
International Archives of The Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives	47
Journal of Flood Risk Management	45
Water Resources Management	41
Arabian Journal of Geosciences	40

Table 3: Country-level scientific production and international collaboration (MCP = Multi-Country Publications)

Country	Articles (% of Total)	MCP Ratio (%)
China	433 (10.97)	17.8
India	341 (8.64)	10.6
USA	210 (5.32)	13.8
Malaysia	145 (3.67)	24.0
Italy	138 (3.50)	26.1
Iran	126 (3.19)	15.5
Indonesia	121 (3.06)	18.4
UK	118 (2.99)	29.3
Germany	107 (2.71)	21.4
Japan	95 (2.41)	12.3

3.4 Country Productivity and Global Collaboration

Country-level results (Table 3) indicate that China (10.97 %), India (8.64 %), and the USA (5.32 %) contribute almost 25 % of global output. Values are normalized to ensure consistency with the 3,946-record dataset. All frequencies and percentages have been normalized and cross-verified to ensure consistency with the final dataset of 3,946 Scopus-indexed records. Rather than reflecting publication volume alone, the country-level distribution of flood-risk modeling research reveals structural drivers of scientific dominance. Countries such as China and India lead not only due to high flood exposure but also because of sustained public investment in geospatial infrastructure, national remote sensing programs, and large research networks focused on disaster risk reduction. In contrast, countries with lower total outputs but higher multiple-country publication (MCP) ratios such as the United Kingdom and Italy function as international knowledge hubs, contributing advanced methodological expertise and facilitating cross-regional research integration. These patterns indicate that global flood-risk research is shaped by both hazard-driven demand and asymmetric access to data, computational resources, and funding mechanisms.

While China and India dominate in total output due to large flood-prone regions and state-funded initiatives, Italy, Malaysia, and the UK display higher MCP ratios, highlighting stronger international partnerships. The global collaboration map (Figure 5) highlights extensive cross-regional knowledge exchange among Asia, Europe, and North America, where technical expertise from advanced research systems intersects with region-specific flood-risk contexts. Countries are shaded using a blue color gradient representing relative publication output, with darker shades indicating higher research productivity. The connecting lines indicate the presence of international co-authorship between countries, reflecting cross-border research collaboration rather than collaboration intensity or direction. The map reveals that countries with high publication output such as China, India, and the United States are embedded within dense international collaboration networks, while several regions with lower output show fewer cross-national linkages. Overall, the figure highlights uneven global participation in flood-risk research and underscores the role of internationally connected research hubs in facilitating global scientific exchange.

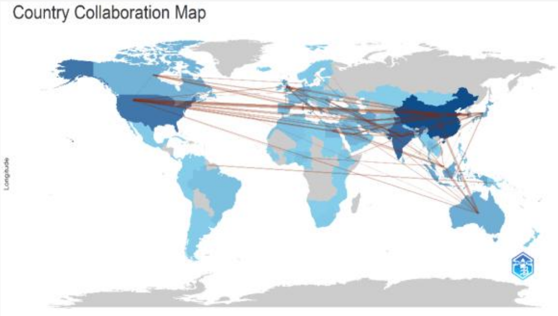

Figure 5: Global collaboration network

3.5 Three-Field and Thematic Network Analysis

The three-field plot (Figure 6) illustrates the relationships among Authors (AU), Author Keywords (DE), and Sources (SO) by linking influential authors, dominant research themes, and leading publication outlets. The strong coupling between author [16] such as keywords including *AHP*, *flood susceptibility*, and *GIS*, and high-impact journals such as *Natural Hazards* and *Sustainability* indicates the emergence of a stable methodological core centered on multi-criteria decision analysis and geospatial modeling. This convergence suggests that certain authors and journals act as epistemic anchors, shaping methodological standards and guiding research trajectories. At the same time, the presence of multiple journals connected to overlapping keyword clusters reflects methodological diffusion and interdisciplinary uptake across environmental, geospatial, and sustainability research domains. Beyond the three-field relationships, the broader bibliometric structure reveals that leading outlets such as Natural Hazards, Water (Switzerland), and the Journal of Hydrology serve as central platforms for interdisciplinary knowledge exchange. Influential scholars including [13][21][27] and [29] have advanced hybrid GIS–ML frameworks, while institutions such as Hohai University, Huazhong University of Science and Technology, and the University of Burdwan act as regional research hubs.

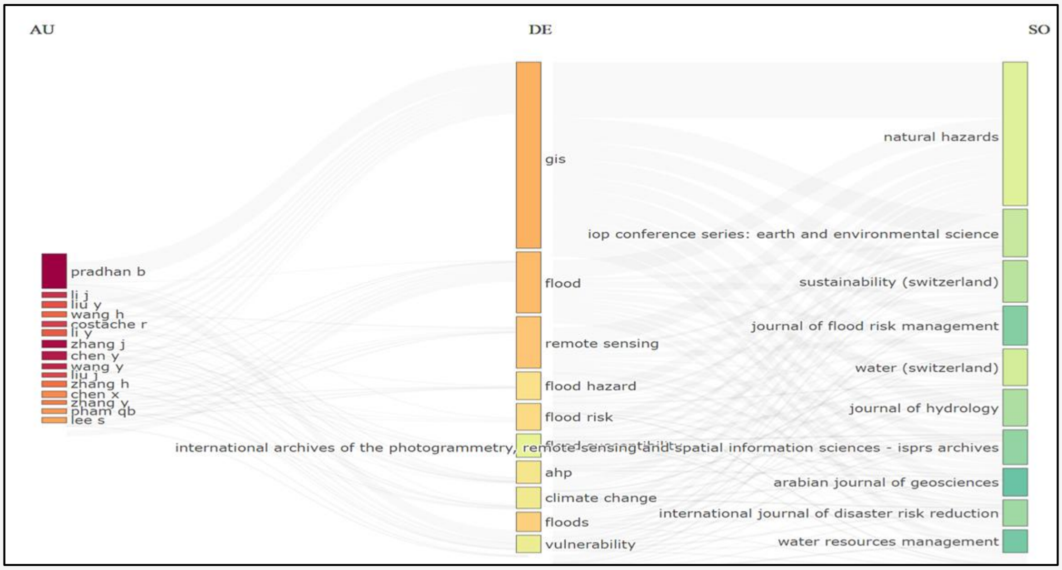

Figure 6: Three field analysis of authors keywords, and sources

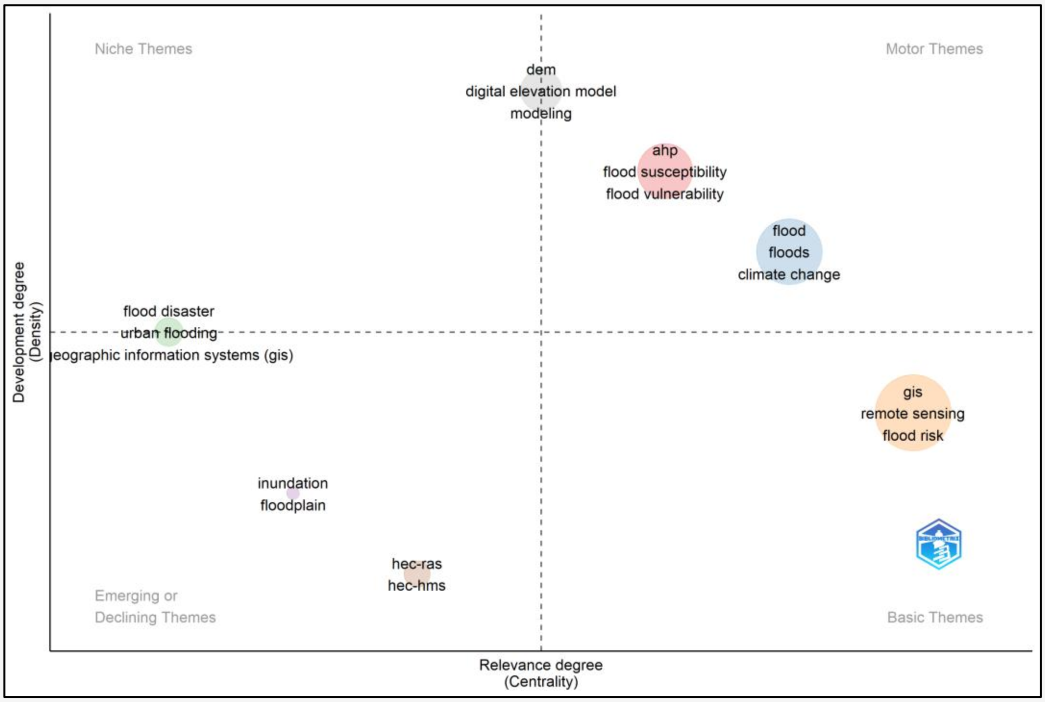

Figure 7: Thematic evolution map

Collectively, these entities strengthen cross-disciplinary integration among hydrology, remote sensing, climate adaptation, and urban planning research. The thematic evolution map (Figure 7), generated via Multiple Correspondence Analysis (MCA), delineates the intellectual trajectory of GIS-based flood risk modeling across four distinct phases:

- 1987–2005 – Hydrological modeling and floodplain mapping: Early studies emphasized deterministic models such as HEC-RAS and HEC-HMS, focusing primarily on hydraulic simulation and terrain-based floodplain delineation.
- 2006–2010 – AHP and multi-criteria GIS assessment: This period marked the formal integration of decision-support frameworks (AHP, fuzzy logic) with GIS for hazard zoning and risk prioritization.
- 2011–2018 – Machine learning and urban vulnerability analysis: Rapid digitalization and

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

data availability enabled the adoption of ML algorithms (RF, SVM, ANN) for spatial prediction and urban flood risk studies.

- 2019–2024 – Explainable AI (SHAP), uncertainty quantification, and resilience assessment: Recent studies increasingly focus on interpretability and probabilistic confidence estimation, representing a paradigm shift toward transparent, data-driven modeling.

The thematic map (Figure 7) positions research themes according to Callon's centrality (relevance) and density (development) quantitatively derived from keyword co-occurrence networks. The upper-right quadrant ('motor themes') includes AHP, flood susceptibility, and flood vulnerability, representing mature and influential topics with high centrality and density values driving the field. The upper-right quadrant ('motor themes') includes AHP, flood susceptibility, and flood vulnerability, representing mature and influential topics driving the field. GIS, remote sensing, and flood risk appear as 'basic themes' methodological foundations underpinning most studies. Digital elevation model and modeling are 'niche themes', reflecting technically specialized subfields. 'Emerging or declining themes', such as inundation, floodplain, and HEC-RAS/HEC-HMS, indicate either transitioning or fading methods. This distribution reveals a methodological evolution from deterministic hydrological modeling to data-driven, interpretable, and uncertainty-aware approaches. Bubble size reflects the frequency of keyword co-occurrence; position shows theme centrality (relevance) and density (development level) based on Callon's algorithm.

As shown in Figure 3, this rise in publication output corresponds with the expansion of global collaboration patterns depicted in Figure 5. The co-authorship structure in Figure 5 complements the source-author relationships identified in Figure 6. The thematic clusters in Figure 7 build upon these quantitative trends, revealing how methodological focus has shifted over time.

3.6 Limitations and Data Bias

Despite Scopus's broad coverage, certain limitations persist. Restricting records to English-language publications excludes regional studies published in national journals. Excluding non-indexed grey literature may underrepresent localized flood-management practices. Bibliometric thresholds (e.g., minimum keyword frequency) may also underweight emerging research themes. Nonetheless, these biases do not alter macro-level patterns observed in publication growth and thematic transition.

4. Critical analysis and Discussion

The bibliometric analysis confirms that GIS-based flood risk modeling has evolved from deterministic hydrological applications to adaptive, data-driven frameworks. The field's sustained annual growth rate (15.08 %) and the post-2010 surge (Figure 3) coincide with the institutionalization of the Sendai Framework (2015) and wider access to open geospatial datasets such as MODIS and Copernicus. These drivers expanded research from localized floodplain mapping to global comparative analyses. Thematic evolution (Figure 7) shows four major developmental stages:

- 1987–2005 – Hydrological modeling and floodplain delineation (foundational stage).
- 2006–2010 – GIS–MCDA integration, introducing AHP and fuzzy-logic decision systems.
- 2011–2018 – Machine-learning adoption (RF, SVM, ANN) for flood susceptibility and urban vulnerability analysis.
- 2019–2024 – Explainable AI and uncertainty quantification, emphasizing transparency through SHAP and probabilistic modeling.

AHP, and flood susceptibility as motor themes that drive innovation, while GIS and remote sensing remain basic themes underpinning the discipline. Digital elevation model and modeling emerge as niche themes, whereas explainable AI and conformal prediction occupy the emerging quadrant, signalling the next methodological frontier.

4.1 Geographic Distribution and Knowledge Exchange

Geographically, China (10.97 %) and India (8.64 %) dominate global output, reflecting their exposure to monsoonal variability, large flood-prone basins, and strong government funding for disaster research. By contrast, Italy, Malaysia, and the UK exhibit higher multi-country publication ratios, indicating strong international partnerships and methodological novelty. The collaboration index (CI = 1.46) points to moderately increasing interdisciplinarity. The global collaboration network (Figure 5) demonstrates pronounced North–South knowledge flows, where technological expertise from developed regions complements data-rich but flood-vulnerable nations in Asia and Africa. Node color in Figure 5 represents countries; line thickness indicates collaboration intensity.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

4.2 Research Gaps and Emerging Frontiers

Despite rapid progress, three persistent gaps constrain current research:

- Explainability and transparency – Few studies employ SHAP or equivalent frameworks, leaving ML models largely opaque.
- Uncertainty quantification – Probabilistic and conformal prediction methods remain underused in flood-susceptibility mapping.
- Socio-economic integration – Limited incorporation of human vulnerability and adaptive capacity reduces policy relevance.

Bridging these deficits requires hybrid, interpretable frameworks uniting GIS, MCDA, and explainable AI with socio-economic indicators. The bibliometric evidence thus provides the empirical basis for proposing future integration of AHP–ML–SHAP–CP approaches not as validated models but as next-generation research directions emerging from observed thematic gaps.

4.3 Policy Implications and Future Directions

The synthesis underscores several priorities:

- Couple real-time hydrological data with GIS for adaptive flood forecasting.
- Cross-regional and intercontinental research collaboration to enhance shared access to advanced flood-modeling expertise and region-specific knowledge.
- Bridge socio-technical divides by integrating social vulnerability indices within GIS workflows.
- Align flood-risk research with SDG 13 (Climate Action) and the Sendai Framework, ensuring translation from research to resilience policy.

Overall, GIS-based flood modeling has transitioned from deterministic, model-driven analysis to transparent, uncertainty-aware systems underpinned by AI and MCDA integration. This paradigm shift, documented through bibliometric evidence, reveals a maturing but still evolving field that increasingly prioritizes interpretability, collaboration, and policy applicability. The bibliometric evidence reveals clear methodological evolution from deterministic GIS–MCDA frameworks to explainable and uncertainty-aware AI-driven models. These observed transitions form the empirical foundation for proposing an integrated conceptual framework. Accordingly, Section 5 synthesizes these insights into a structured model (AHP–ML–SHAP–CP) that reflects emerging research directions rather than experimental implementation

5. Emerging Methodological Themes and Research Gaps in Flood-Risk Modelling

Building upon the bibliometric insights derived in Section 4, this section conceptualizes an integrated analytical framework that addresses the interpretability and uncertainty gaps identified in GIS-based flood risk modeling literature.

5.1 Analytical Hierarchy Process (AHP) and Its Role in Flood Mapping

The Analytical Hierarchy Process (AHP), introduced by Saaty in the 1980s, remains one of the most widely applied Multi-Criteria Decision-Making (MCDM) tools in GIS-based flood risk mapping. It enables the systematic weighting of factors such as slope, land use/land cover (LULC), soil type, rainfall intensity, and drainage density through pairwise comparisons. AHP's transparency, simplicity, and adaptability make it ideal for data-scarce regions. However, its deterministic nature introduces subjectivity and potential bias, especially when multiple experts or diverse criteria are involved.

Bibliometric findings from Section 4 confirm that AHP-based approaches dominated the 2006–2010 thematic phase, serving as foundational decision-support tools for flood hazard zoning. Yet, later phases (2011–2024) revealed a shift toward machine learning and explainable AI, driven by the need for automation, interpretability, and uncertainty quantification areas where traditional AHP is limited. To mitigate these limitations, the literature has explored Fuzzy AHP, Entropy weighting, and Bayesian AHP have emerged to handle uncertainty in expert judgments. Simultaneously, Machine Learning (ML) models Random Forests (RF), Support Vector Machines (SVM), and Gradient Boosting have gained traction for their predictive accuracy. Nonetheless, their interpretability limitations (the "black box" problem) necessitate further enhancement.

5.2 SHAP and Explainable AI in Flood Risk Analysis

Explainable AI (XAI) has recently become a focal point in flood susceptibility mapping, as identified in the 2019–2024 thematic cluster (Figure 7). Among XAI methods, SHAP (SHapley Additive exPlanations) offers robust interpretability by quantifying each feature's contribution to the model's output. For instance, in a flood susceptibility model, SHAP can show how rainfall or LULC individually affect the probability of flooding. By bridging statistical inference and decision support, SHAP allows researchers to improve transparency by explaining the contribution of individual variables within trained machine learning models, thereby supporting model interpretation and stakeholder communication.

5.3 Emerging Needs in Interpretability and Uncertainty Assessment in Flood-Risk Modelling

The bibliometric analysis highlights two persistent methodological gaps within flood susceptibility and flood-risk modelling research. First, despite the rapid adoption of machine learning techniques, model interpretability remains limited, with relatively few studies providing transparent explanations of variable influence on flood occurrence. Second, uncertainty quantification is largely absent, as most flood susceptibility maps present deterministic outputs without associated confidence measures. These gaps are not addressed through the proposal of a new modelling framework in this study. Instead, they are discussed here as emerging methodological needs identified through the temporal evolution of author keywords, thematic clusters, and highly cited literature. The increasing prominence of multi-criteria decision analysis, machine learning, explainable artificial intelligence, and uncertainty-aware modelling reflects a broader shift toward more transparent and reliable geospatial flood-risk assessments.

Multi-criteria decision analysis methods, particularly the Analytic Hierarchy Process (AHP), continue to be used for structuring expert knowledge and weighting geospatial predictors. Concurrently, machine learning approaches have demonstrated strong predictive capabilities for modelling complex, non-linear relationships in flood-prone environments. More recently, explainable artificial intelligence techniques, such as SHAP-based feature attribution, have gained attention for their potential to enhance transparency by clarifying the relative influence of conditioning factors.

Uncertainty quantification remains an underexplored dimension in flood susceptibility mapping. Although conformal prediction has not yet been widely adopted in operational flood-risk studies, its growing presence in adjacent geospatial and environmental modelling literature suggests its potential relevance for future research. In this context, conformal prediction is discussed not as a validated solution, but as an emerging research direction aimed at addressing the long-standing absence of confidence bounds in susceptibility outputs. Overall, the co-occurrence of these methodological themes in recent literature underscores a transition from purely accuracy-driven modelling toward approaches that emphasize interpretability, transparency, and uncertainty awareness. These trends provide a roadmap for future flood-risk research rather than a prescriptive or integrated modelling framework.

5.4 Uncertainty Quantification as an Emerging Research Gap

Bibliometric analysis indicates that uncertainty quantification remains a comparatively underrepresented topic in flood susceptibility and flood-risk modelling research. Most studies continue to rely on conventional accuracy assessment metrics such as AUC, F1-score, and Kappa statistics to evaluate model performance, with limited attention to the communication of predictive uncertainty. Recent methodological discussions in broader geospatial and environmental modelling literature have highlighted post-hoc uncertainty quantification techniques, including conformal prediction, as potential tools for characterizing prediction reliability after model validation. However, the application of such techniques in flood-risk modelling remains limited and fragmented. Consequently, uncertainty-aware mapping should be viewed as a research gap rather than an established methodological direction within the flood literature.
In the context of this scientometric review, conformal prediction is referenced solely to illustrate the growing recognition of uncertainty assessment in geospatial modelling, not as a modelling framework or substitute for conventional validation practices.

5.5 Implications for Future Flood-Risk Research

The bibliometric trends identified in this study suggest that future flood-risk research should continue to emphasize methodological transparency, rigorous validation, and improved communication of uncertainty. While expert-based decision-support approaches and data-driven machine learning techniques have each demonstrated value, their application has largely progressed along separate methodological pathways. Future studies should prioritize the careful selection of modelling approaches based on data availability, study objectives, and validation requirements, while maintaining standard accuracy assessment as the primary means of performance evaluation. Advances in explainable artificial intelligence and uncertainty quantification may complement these efforts by enhancing interpretability and reliability, provided they are applied appropriately and independently. Importantly, this scientometric review does not propose or evaluate any integrated modelling framework. Instead, it highlights methodological gaps and emerging themes that warrant further empirical investigation through dedicated modelling studies.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

6. Conclusion

This study provides a comprehensive scientometric assessment of flood risk and flood susceptibility modelling research based on bibliometric evidence extracted from the Scopus database. By analyzing publication growth, citation structures, keyword co-occurrence networks, thematic evolution, and international collaboration patterns, the study offers an empirical synthesis of how flood-risk modelling has evolved over the past two decades. The bibliometric results reveal three key empirical trends. First, flood-risk research has experienced rapid and non-linear growth since approximately 2010, coinciding with advances in remote sensing, geospatial data availability, and machine learning techniques. Second, the thematic evolution demonstrates a clear methodological transition from early dominance of GIS-based multi-criteria decision analysis toward data-driven machine learning approaches, followed more recently by increasing attention to explainable artificial intelligence and uncertainty-aware modelling. Third, the global collaboration analysis highlights a geographically uneven but increasingly interconnected research landscape, where high-output countries in Asia intersect with methodologically influential research hubs in Europe and North America.

From a theoretical perspective, this study contributes by clarifying how flood-risk modelling research has shifted from deterministic, factor-weighting paradigms toward approaches that prioritize predictive performance, interpretability, and reliability. Rather than proposing new models, the analysis synthesizes how methodological priorities have changed over time and identifies persistent gaps, particularly in model transparency and uncertainty communication. This perspective advances the understanding of flood-risk modelling as an evolving geospatial science shaped by both technological innovation and policy-driven research agendas.

Practically, the findings provide guidance for researchers and practitioners by identifying dominant methods, influential publication outlets, and emerging research themes. The results underscore the continued importance of rigorous accuracy assessment in flood susceptibility mapping, while also highlighting growing interest in complementary approaches that enhance interpretability and reliability. For policymakers and disaster-risk managers, the observed increase in international collaboration suggests expanding opportunities for cross-regional knowledge transfer and methodological standardization. Based on the bibliometric evidence, future flood-risk research should focus on three actionable directions: (i) improving transparency in machine learning–based flood models through systematic use of explainable AI techniques, (ii) strengthening uncertainty awareness by explicitly reporting reliability measures alongside conventional accuracy metrics, and (iii) fostering international and interdisciplinary collaboration to address region-specific flood challenges using globally informed methodologies. These recommendations emerge directly from observed publication trends and thematic gaps, rather than from hypothetical or untested modelling frameworks.

Overall, this scientometric review consolidates the state of flood-risk modelling research, identifies critical methodological transitions, and provides an evidence-based roadmap for future investigations within the geoinformatics community.

7. Limitations and Future Research Direction

Despite providing a comprehensive overview of research trends in GIS- and AHP-based flood risk modeling, this study is subject to several limitations inherent to bibliometric analyses. First, the analysis relies on a single bibliographic database, which may underrepresent relevant studies published in regional journals, non-English outlets, or grey literature, potentially introducing database and language bias. Second, citation-based indicators are influenced by publication age, disciplinary norms, and self-citation practices, which may not always reflect the true scientific or societal impact of a study. Third, keyword-based analyses depend on author-defined terminology; variations in keyword selection, synonym usage, and evolving nomenclature may affect thematic clustering results. Additionally, collaboration metrics may overemphasize highly networked institutions while underrepresenting informal or emerging research partnerships. Finally, bibliometric methods capture structural and quantitative patterns in the literature but do not directly evaluate methodological rigor or empirical validity of individual studies. These limitations should be considered when interpreting the results, and future research could integrate systematic content analysis or expert-driven review to complement bibliometric insights.

Future studies can address the limitations of bibliometric analyses by integrating multi-database sources (e.g., Scopus, Web of Science, and regional indexing platforms) to reduce database and language bias and provide a more inclusive representation of global flood-risk research. Complementary systematic content analysis and meta-analytical approaches may be employed to evaluate methodological rigor, model performance, and uncertainty handling beyond citation-based influence.

The incorporation of semantic text mining and natural language processing techniques could improve thematic detection by accounting for evolving terminology and synonym variability in keyword usage. In addition, future research may explore dynamic collaboration networks and longitudinal knowledge diffusion models to better capture informal, emerging, and research linkages often underrepresented in co-authorship analyses. Finally, combining bibliometric insights with case-based validation and expert elicitation would strengthen the translation of scientometric patterns into actionable guidance for flood-risk modeling, climate adaptation, and disaster-resilience policy.

References

[1] Chauhan, V., Gupta, L. and Dixit, J., (2025). Machine Learning and GIS-Based Multi-Hazard Risk Modeling for Uttarakhand. *Environmental and Sustainability Indicators,* Vol. 20. https://doi.org/10.1016/j.indic.2025.100236.

[2] Gaikwad, V. P., Pande, S. R. and Joshi, R., (2025). Assessing the Flood Susceptibility Mapping in the Krishna River Basin. *SSRN Electronic Journal.* [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5153613. [Accessed: Jan. 5, 2025].

[3] Liu, X., Lyu, H. M. and Shen, S. L., (2025). Assessment of Geo-Disaster Risk Levels Induced by Rainfall. *Geomatics, Natural Hazards and Risk,* Vol. 16. https://doi.org/10.1080/17499518.2025.2503254.

[4] Ghosh, S., Bhowmik, R. and Bora, A., (2025). *Forest Vulnerability and Flood Mapping in Assam*. Sciforum. [Online]. Available: https://sciforum.net/paper/view/23124. [Accessed: Jan. 5, 2025].

[5] Pandey, P. C., Kumar, R., Pandey, M., Giuliani, G., Sharma, R. K. and Srivastava, P. K., (2024). *Geo-Information for Disaster Monitoring and Management.* Springer, Cham. https://doi.org/10.1007/978-3-031-51053-3.

[6] Kendagannaswamy, M. S., Tripathi, A. and Kumar, S., (2025). Regional-Scale Flood Susceptibility Mapping in Kerala. *Discover Applied Sciences,* Vol. 5. https://doi.org/10.1007/s42452-025-07182-z.

[7] Tayyab, M., Khan, M. A. and Mehmood, M., (2024). Assessment of Flood Types Susceptibility in Peshawar Using GIS and Machine Learning. *Journal of Environmental Management,* Vol. 360. https://doi.org/10.1016/j.jenvman.2024.120811.

[8] Tehrany, A., Pradhan, B. and Jebur, M. N., (2023). Integrated Flash Flood Indexing Using Random Forest and MLP with AHP. *Hydrology Research,* Vol. 54; 130–148. https://doi.org/10.2166/nh.2022.123.

[9] Shrestha, S., Paudel, D. and Gautam, R., (2023). RF–AHP Flood Susceptibility Mapping in Nepal. *International Journal of Disaster Risk Reduction,* Vol. 89. https://doi.org/10.1016/j.ijdrr.2023.103591.

[10] Linh, P. V. and Nguyen, Q. V., (2023). SHAP-Integrated Deep Learning for Flood Susceptibility in Vietnam. *Water Resources Research,* Vol. 59. https://doi.org/10.1029/2023WR032120.

[11] Olatunji, B. and Akinwale, M., (2024). Urban Flood Hazard Modeling Using SVM and AHP in Lagos. *GeoJournal,* Vol. 89, 143–158. https://doi.org/10.1007/s10708-023-10724-8.

[12] Saha, S., (2025). *Introduction to Multicriteria Decision Making in Hazard Monitoring. In: Multi-Hazard Early Warning Systems Using Geospatial and AI Technologies*. Springer, Cham; 1–12. https://doi.org/10.1007/978-3-031-89246-2_1.

[13] Mirzaei, M., Kalantar, B. and Kabiri, K., (2022). Logistic Regression versus Random Forest versus AHP for Flood Zoning. *Hydrological Processes,* Vol. 36. https://doi.org/10.1002/hyp.14560.

[14] Chen, X., Zhang, H. and Li, R., (2022). Multi-Hazard Modeling Using GIS and AHP in China. *ISPRS International Journal of Geo-Information,* Vol. 11. https://doi.org/10.3390/ijgi11100542.

[15] Sarkar, S. and Mondal, S., (2020). Flood Risk Zoning Using Fuzzy–AHP. *Sustainability,* Vol. 12. https://doi.org/10.3390/su12093853.

[16] Pradhan, B. and Lee, S., (2010). GIS-Based Flood Hazard Mapping Using Frequency Ratio and AHP. *Environmental Modelling and Software,* Vol. 25, 714–726. https://doi.org/10.1016/j.envsoft.2009.10.016.

[17] Nair, A. and Ramachandran, P., (2022). MCDA–CoCoSo Based Flood Risk in Kerala. *Sustainability,* Vol. 14. https://doi.org/10.3390/su14073953.

[18] Islam, S., Ahmed, R. and Rahman, T., (2024). Urban Flood Mapping in Bangladesh Using AHP and Entropy-Weighted Machine Learning. *Environmental Monitoring and Assessment,* Vol. 196. https://doi.org/10.1007/s10661-024-5678-2.

[19] Das, S., (2020). Flood Vulnerability Assessment Using Entropy–AHP. *Natural Hazards Review,* Vol. 21. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000359.

[20] Dou, J., Pham, B. T. and Shahabi, H., (2024). Explainable Machine Learning in Flood Susceptibility Mapping Using SHAP and AHP. *Remote Sensing,* Vol. 16. https://doi.org/10.3390/rs16030385.

[21] Wang, G. and Yu, X., (2023). Transparent Machine Learning–AHP Fusion for Hazard Prediction. *Science of the Total Environment,* Vol. 899. https://doi.org/10.1016/j.scitotenv.2023.165890.

[22] Barzegar, R., Pradhan, B. and Shafizadeh-Moghadam, H., (2020). SHAP and XGBoost for Flood Explainability. *Environmental Modelling and Software,* Vol. 133. https://doi.org/10.1016/j.envsoft.2020.104849.

[23] Mofidi, A. and Abolfazli, M., (2021). Bayesian–AHP Hybrid Model for Flood Mapping. *Hydrology and Earth System Sciences,* Vol. 25, 4435–4452. https://doi.org/10.5194/hess-25-4435-2021.

[24] Gonsalis, S. S., Mishra, M. and Bhatnagar, S., (2024). Soft Computing Techniques for Flood Risk in Indian Coastal Zones. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS),* 1–4. https://doi.org/10.1109/IGARSS.2024.10956247.

[25] Sahana, M. and Sajjad, H., (2021). AHP–GIS Urban Flood Vulnerability Mapping. *Hydrological Sciences Journal,* Vol. 66, 710–726. https://doi.org/10.1080/02626667.2021.1898254.

[26] Mahalingam, S., Pradeep, V. and Sundaram, A., (2023). SAR–AHP Integrated Real-Time Flood Mapping. *Remote Sensing Applications: Society and Environment,* Vol. 29. https://doi.org/10.1016/j.rsase.2023.100890.

[27] Roy, A. and Das, S., (2023). Geo-AI and AHP Fusion for National Flood Zoning. *ISPRS Archives,* Vol. 48, 237–244. https://doi.org/10.5194/isprs-archives-XLVIII-237-2023.

[28] Jain, S. K. and Yadav, R., (2021). Flash Flood Modeling in the Himalayas with AHP. *Sustainability,* Vol. 13. https://doi.org/10.3390/su131810025.

[29] Al-Abadi, A. and Al-Taiee, T., (2021). Comparative Study of AHP and Machine Learning for Tigris Flood Zones, Environmental Earth Sciences, Vol. 80. https://doi.org/10.1007/s12665-021-09862-8.

[30] Liu, Y., Wang, Y., Zhang, S. and Li, Z., (2022). Susceptibility Assessment of Flash Floods: A Bibliometric Analysis and Review. *Remote Sensing,* Vol. 14(19). https://doi.org/10.3390/rs14194791

[31] Madushani, J. A. T., Wijesekera, N. T. S. and Gunasekara, S. S., (2024). Thematic and Bibliometric Review of Remote Sensing and Geographic Information System-Based Flood Disaster Studies in South Asia (2004–2024). *Sustainability,* Vol. 17(1). https://doi.org/10.3390/su17010217.