19th WORLD CONFERENCE OF THE ASSOCIATED RESEARCH CENTRES FOR THE URBAN UNDERGROUND SPACE
Belgrade, Serbia, November 4-7, 2025.

ACUUS BELGRADE 2025

Paper No: 7.3.149
DOI: 10.46793/ACUUS2025.7.3.149

FLOOD RISK ASSESSMENT USING A LITERATURE-BASED APPROACH

Wei-Wei Zhao[1], Shui-Long Shen[2], Annan Zhou[3], Yan-Ning Wang[4]

Abstract: Urban flood risk is intensifying under the dual pressures of rapid urbanization and climate change, highlighting the need for robust, transparent, and scalable assessment frameworks. This study presents a fully data-driven framework for urban flood risk assessment that integrates region-specific risk factors and minimizes subjective bias. A bibliometric-informed weighting method is employed to objectively assign importance to each factor based on its prominence in peer-reviewed literature. The framework was applied to a case study in Guilin, China, where twenty localized indicators were synthesized through spatial analysis to produce a composite flood risk map. Model validation against the major flood event of June 2024 demonstrated strong spatial agreement between predicted high-risk areas and actual flood-affected zones. These results confirm the framework's predictive validity and practical applicability. By enhancing transparency, reproducibility, and adaptability, this approach offers valuable support for urban flood resilience planning and disaster mitigation.

Keywords: Urban flood risk, Data-driven framework, Region-specific indicators, Literature-based weighting, Flood risk mapping

1. INTRODUCTION

Urban flood risk is increasingly critical due to rapid urbanization and climate change, which intensify both the frequency and severity of flooding (Zhou et al., 2019). Beyond infrastructure damage, floods disrupt transportation, public services, health systems, and local economies (Allaire., 2018; Hammond et al., 2015), necessitating robust and scalable assessment frameworks for resilience planning. Traditionally, flood risk is conceptualized as a function of hazard, exposure, and vulnerability (Maranzoni et al., 2023), typically assessed using proxy indicators and models. Physically based hydrodynamic models like HEC-RAS (Joyce et al., 2009) simulate flood dynamics but depend on high-resolution data, often unavailable in urban settings.

Considering these limitations, there is a pressing need to explore alternative, objective, and scalable frameworks for flood risk assessment. A conventional flood risk assessment framework typically conceptualizes risk as the intersection of hazard, exposure, and vulnerability (Maranzoni et al., 2023). These components are typically assessed using proxy variables supported by numerical simulations and geospatial analysis techniques. Physically based hydrodynamic models, such as HEC-RAS (Joyce et al., 2009), are commonly used to simulate flood propagation and estimate hazard intensity. While effective, these models rely on high-resolution data that are often scarce or costly in urban areas.

An alternative approach employs multi-criteria decision analysis to construct composite risk indices (Lin et al., 2019). Popular weighting methods include AHP (Lyu et al., 2020; Wu et al., 2022), entropy (Liu et al., 2019; Wu et al., 2022) and Principal Component Analysis (PCA) (Liu et al., 2021). However, these techniques often suffer from subjectivity, particularly in expert-driven weighting schemes, which can affect reproducibility and policy relevance (Munier et al., 2021). To mitigate this, literature analysis offers a data-informed means of determining indicator weights by examining keyword frequencies, co-occurrence, and research trends (Lieber et al., 2022).

This study proposes a novel urban flood risk assessment framework that combines literature insights with data-driven modeling to reduce bias and enhance objectivity. A case study in Guilin, China, demonstrates its practical applicability.

[1] Ph.D. Candidate, Zhao, Wei-Wei, M.Sc. Computer Science, Shantou University, Shantou, China; 14wwzhao@stu.edu.cn.
[2] Professor, Shen, Shui-Long, Ph.D. Civil Engineering, Dean College of Engineering, Shantou University, Shantou, China, shensl@stu.edu.cn.
[3] Professor, Zhou, Annan, Ph.D. Civil Engineering, School of Engineering, RMIT University, Melbourne, Australia, annan.zhou@rmit.edu.au.
[4] Professor, Wang, Yan-Ning, Ph.D. Civil Engineering, College of Engineering, Shantou University, Shantou, China, wangyn@stu.edu.cn.

2. METHODOLOGY

2.1. Risk Assessment Framework

The flood risk assessment framework adopted in this study consists of five main stages: risk factor identification, data preprocessing, Flood Risk Index (FRI) construction, comparison analysis, and validation.

In the first stage, six major categories of flood-related factors were identified through an integrated approach combining literature review and expert knowledge: topography, hydrology, natural buffering, anthropogenic activities, socioeconomic resilience, and infrastructure. This factor selection process ensures a comprehensive and context-specific representation of flood risk drivers. The selected factors were then standardized and spatially aligned during data preprocessing to maintain consistency across datasets.

The FRI construction stage incorporates two key innovations: the comprehensive factor selection mentioned above, and the use of bibliometric-based weighting methods to assign relative importance to each factor. Specifically, three weighting approaches were employed: entropy method, PCA, and literature frequency. For index classification, two methods were applied: the normal distribution method, which segments the FRI based on standard deviations from the mean, and the Natural Breaks method, which optimizes classification for skewed or non-normal distributions.

To assess the consistency and robustness of the FRI outputs generated by these different approaches, comparison analyses were conducted using metrics such as consistency level, Kappa coefficient, and Gi* statistic. Finally, the model's validity was evaluated by comparing the spatial distribution of flood risk against historical flood records, providing empirical evidence of the FRI's predictive accuracy.

2.2. Risk Factor Identification with Urban Characteristics

Accurate flood risk assessment requires selecting factors that reflect local geographic and socioeconomic conditions. In Guilin, the dense river network, rapid urbanization, and economic reliance on tourism and transport necessitate a context-specific approach. Key factors such as tourism density, river proximity, transport hubs, and flood control structures capture the city's unique risk profile, ensuring that both hazard and mitigation elements are appropriately represented in the evaluation framework.

Table 1. *Flood risk assessment factors*

Factor	Category	Direction
Digital evaluation model (DEM)	Topography	Negative
Slope	Topography	Negative
River density	Hydrology	Positive
River proximity	Hydrology	Positive
Precipitation	Hydrology	Positive
Natural landscape features	Natural Buffering	Negative
Land cover - risk	Natural Buffering	Positive
Land cover - mitigation	Natural Buffering	Negative
Population density	Anthropogenic	Positive
Tourism density	Anthropogenic	Positive
Land use - risk	Anthropogenic	Positive
Land use - mitigation	Anthropogenic	Negative
Per capita income	Socioeconomic Resilience	Negative
Public fiscal expenditure	Socioeconomic Resilience	Negative
Transport hubs	Infrastructure	Positive
Railways	Infrastructure	Positive
Roads	Infrastructure	Positive
Building density	Infrastructure	Positive
Flood control structures	Infrastructure	Negative
Reservoir influence	Infrastructure	Positive

Table 1 summarizes twenty flood risk factors grouped into six categories: Topography, Hydrology, Natural Buffering, Anthropogenic Activity, Socioeconomic Resilience, and Infrastructure. The factors were chosen for their roles in hazard, exposure, or mitigation, with the six categories collectively representing the different aspects that contribute to flood risk. The table also includes a “Direction” column indicates whether a factor increases (positive) or decreases (negative) flood risk. Topography includes the Digital Elevation Model (DEM) and slope, both of which influence how water moves across the landscape; higher elevation and steeper slopes generally reduce flood accumulation, so these factors are marked as negative contributors to flood risk. Hydrology captures the direct drivers of flooding, including river density, river proximity, and precipitation. Areas with dense river networks, closer proximity to rivers, or higher rainfall are more prone to flooding, and these factors are therefore considered positive. Natural Buffering reflects the protective capacity of the environment. Features such as natural landscapes and land cover designed for mitigation help absorb or slow floodwaters and are marked negative, whereas land cover that increases vulnerability (e.g., impervious surfaces or risk-prone land use) is marked positive. Anthropogenic Activity represents human-driven factors that influence exposure. Higher population density, tourism activity, and land use that increases flood exposure all contribute positively to flood risk. Conversely, land use designed for mitigation, such as parks or green spaces that reduce runoff, is treated as negative. Socioeconomic Resilience captures a community's capacity to withstand floods. Higher per capita income and greater public fiscal expenditure provide resources for adaptation and recovery, and thus are considered negative contributors to risk. Finally, Infrastructure encompasses built structures that can either increase or decrease risk. Transport hubs, railways, roads, building density, and reservoirs can amplify exposure and are marked positive, while flood control structures are designed to reduce risk and are therefore negative.

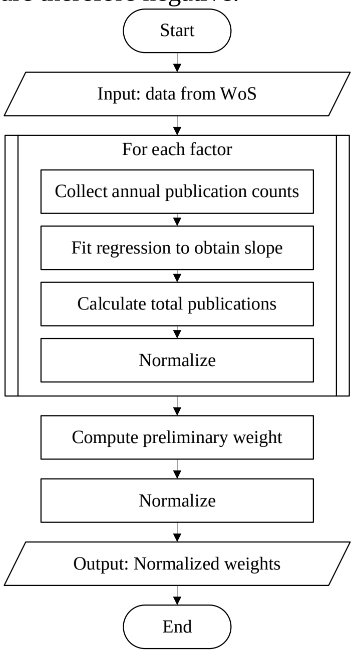

Figure 1. *Flowchart of Literature Frequency Analysis*

2.3. Literature-Based Weighting Method

To reduce subjectivity in assigning weights to flood risk factors, this study proposes a literature-based weighting approach that draws on academic literature as an objective source of evidence. Literature analysis quantifies the extent and attention each factor has received in flood-related research, thereby reflecting its relative importance in the scientific community. The core idea of the literature frequency analysis is to use the volume of relevant publications as a proxy for the significance of each risk factor. To operationalize this, we carried out keyword-based searches in the Web of Science database, carefully defining search terms for each risk factor category (e.g., "slope AND flood," "infrastructure AND flood"). The number of retrieved publications serves as an indicator of how extensively each factor has been studied. Factors with higher publication counts are thus assigned greater weights, while those with fewer references were down-weighted. This approach offers a data-driven and reproducible method for weighting, grounded in the collective focus of flood research. By relying on publication volume, it avoids potential biases inherent in expert opinion-based methods while reflecting research community consensus.

Figure 1 presents the Literature Frequency Analysis, which derives factor weights from both publication volume and observed research trends. The flowchart breaks down the procedure into a series of systematic steps. The process begins with collecting raw bibliometric data from the Web of Science database. For each flood risk factor, annual publication counts are compiled over the study period, providing a temporal profile of research attention. A regression analysis is then calculated to estimate the trend, represented by the slope, which captures whether interest in a particular factor is increasing, decreasing, or stable over time. Next, the total number of publications for each factor is provided to measure overall academic focus. These values are subsequently normalized to bring all factors onto a common scale, facilitating direct comparison. From this, preliminary weights are computed by combining the normalized total publication counts and the trend information. A final normalization step converts these preliminary weights into standardized bibliometric weights, which are assigned to each factor for integration into the Flood Risk Index.

3. RESULTS

This study conducts a case study in Guilin, China. Located in the Guangxi Zhuang Autonomous Region of southern China, Guilin lies within the upper reaches of the Pearl River basin and is characterized by a complex karst landscape with steep limestone hills and an extensive river network. This unique geographic and hydrological setting makes Guilin particularly vulnerable to flooding, as rapid runoff from the steep terrain and concentrated river flows increase flood risk in the urban area.

Flood Risk Index (FRI) maps were generated using the literature frequency analysis Frequency weighting method, which assigns factor weights based on publication volume in relevant literature. To enhance comparability across the three weighting methods, the continuous FRI values were reclassified using the Normal Distribution Rule. According to this rule, the FRI values are stratified into four ordinal levels based on their positive deviation from the mean (μ): (i) Level I (Low Risk): FRI $\leqslant \mu + 1\sigma$ (~68%); (ii) Level II (Moderate-Low Risk): $\mu + 1\sigma <$ FRI $\leqslant \mu + 2\sigma$ (~13.5%); (iii) Level III (Moderate-High Risk): $\mu + 2\sigma <$ FRI $\leqslant \mu + 3\sigma$ (~2.35%); (iv) Level IV (High Risk): FRI $> \mu + 3\sigma$ (~0.15%).

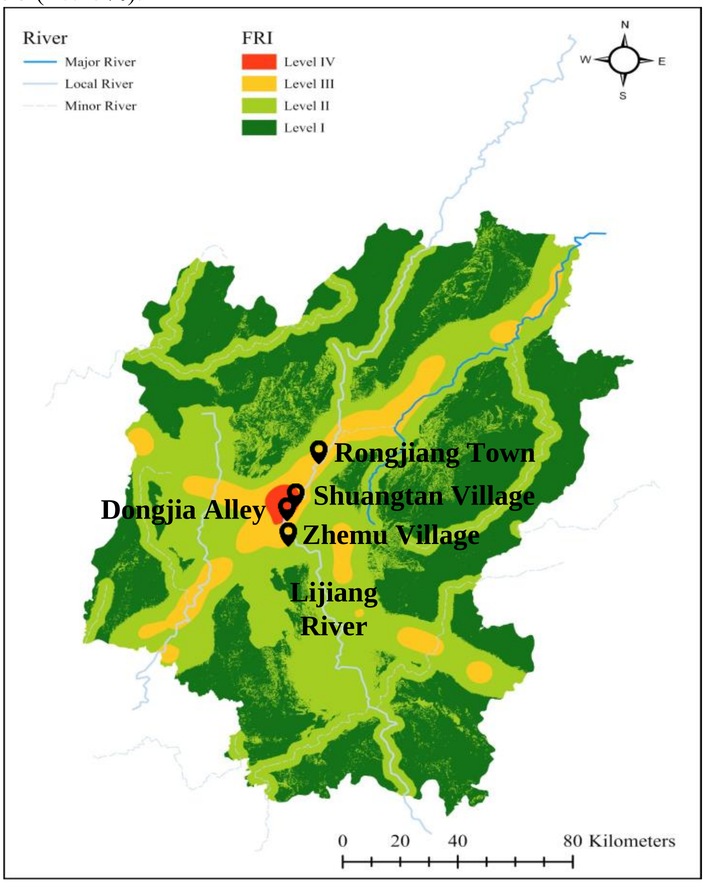

Figure 2. *FRI Map Classified by the Normal Distribution Rule*

4. DISCUSSIONS

The Flood Risk Index (FRI) map shown in Figure 2 was generated solely using the literature frequency weighting method. Its practical validity was assessed by comparing the spatial distribution of flood risk with a significant flood event in Guilin in June 2024, which exceeded a 30-year recurrence interval. Figure 3 presents the FRI map alongside satellite images from before (June 15) and during (June 19) the flood event. Four representative sites—Rongjiang Town, Shuangtan Village, Dongjia Alley, and Zhemu Village—highlighted in Figure 3, were selected for detailed spatial comparison. At each site, satellite imagery reveals substantial water expansion consistent with inundation, which aligns well with the high-risk areas identified in the FRI map. This strong spatial agreement confirms the robustness of the literature-based FRI model and underscores its applicability for urban flood risk assessment in Guilin.

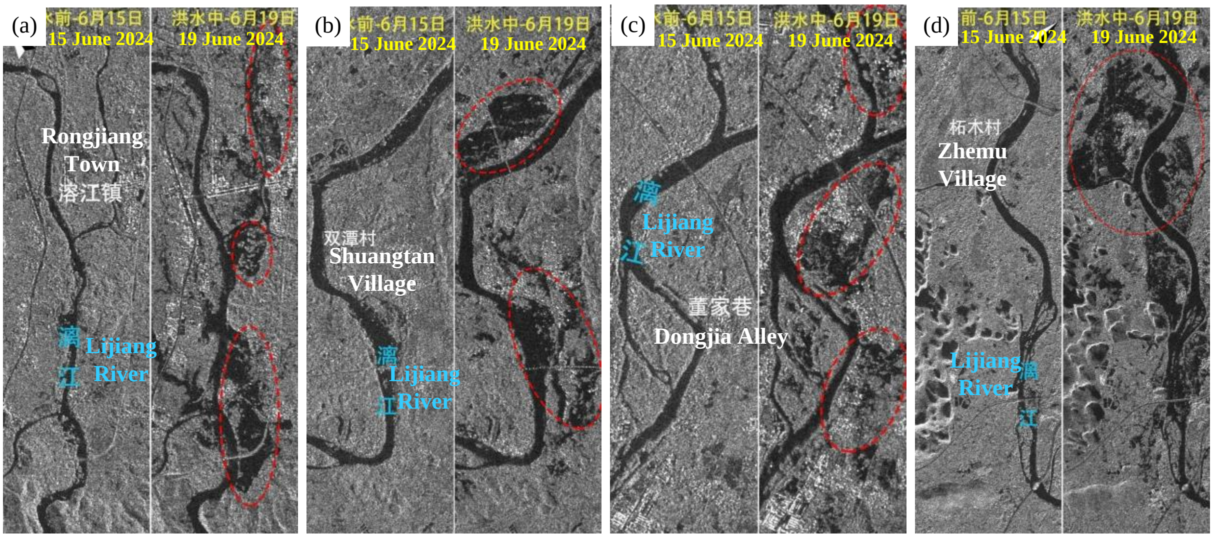

Figure 3. *Validation Using Satellite Imagery, recreated based on News (2025)*

Future developments should prioritize the integration of dynamic and real-time data sources to improve model responsiveness. Although this study used June 2024 rainfall data and validated with June 2025 flood events, the lack of real-time inputs constrains adaptability to fast-changing flood conditions. Incorporating satellite rainfall estimates, real-time hydrological data, and IoT-based monitoring would enable more accurate assessments, particularly in rapidly urbanizing or climate-sensitive regions. Additionally, the model can support flood governance by simulating human interventions such as reservoir releases or pumping station operations. These improvements would allow the framework to evolve from a static risk map into a dynamic, data-driven tool for adaptive urban flood management.

5. CONCLUSIONS

This study introduces a fully data-driven framework for urban flood risk assessment, demonstrated through a case study in Guilin, China. By integrating region-specific risk factors, objective weighting, and robust validation, the framework produces reliable and actionable flood risk maps to support informed decision-making.

The main contributions are summarized as follows:

(1) A comprehensive data-driven framework was developed to assess urban flood risk. The framework integrates region-specific flood risk factors tailored to Guilin’s geographic and socio-economic conditions. This localized approach improves model sensitivity and accuracy in capturing spatial flood risk patterns.

(2) A novel literature-based weighting method was introduced to quantify the relative importance of risk factors. This technique reduces subjectivity by aligning indicator weights with their empirical prominence in peer-reviewed literature.

(3) The framework was validated using the major flood event of June 2024 in Guilin. The results demonstrate strong predictive performance and practical value, confirming its effectiveness in supporting targeted flood mitigation and urban resilience planning.

6. ACKNOWLEDGMENTS

The research work was funded by Guangdong Provincial Basic and Applied Basic Research Fund Committee (2022A1515240073).

7. BIBLIOGRAPHY

[1] Allaire, M. J. W. S. (2018). Socio-economic impacts of flooding: A review of the empirical literature. *3*, 18-26.

[2] Hammond, M. J., Chen, A. S., Djordjević, S., Butler, D., & Mark, O. J. U. W. J. (2015). Urban flood impact assessment: A state-of-the-art review. *12*(1), 14-29.

[3] Lieber, M., Chin-Hong, P., Kelly, K., Dandu, M., & Weiser, S. D. J. G. P. H. (2022). A systematic review and meta-analysis assessing the impact of droughts, flooding, and climate variability on malnutrition. *17*(1), 68-82.

[4] Lin, L., Wu, Z., & Liang, Q. J. N. H. (2019). Urban flood susceptibility analysis using a GIS-based multi-criteria analysis framework. *97*, 455-475.

[5] Liu, Y., You, M., Zhu, J., Wang, F., & Ran, R. J. I. J. o. D. R. R. (2019). Integrated risk assessment for agricultural drought and flood disasters based on entropy information diffusion theory in the middle and lower reaches of the Yangtze River, China. *38*, 101194.

[6] Liu, Z., Jiang, Z., Xu, C., Cai, G., & Zhan, J. J. N. H. (2021). Assessment of provincial waterlogging risk based on entropy weight TOPSIS–PCA method. *108*, 1545-1567.

[7] Lyu, H.-M., Sun, W.-J., Shen, S.-L., Zhou, A.-N. J. J. o. C. E., & Management. (2020). Risk assessment using a new consulting process in fuzzy AHP. *146*(3), 04019112.

[8] Lyu, H.-M., Zhou, W.-H., Shen, S.-L., & Zhou, A.-N. (2020). Inundation risk assessment of metro system using AHP and TFN-AHP in Shenzhen. *Sustainable Cities and Society, 56*, 102103. doi:https://doi.org/10.1016/j.scs.2020.102103

[9] Maranzoni, A., D'Oria, M., & Rizzo, C. J. J. o. F. R. M. (2023). Quantitative flood hazard assessment methods: A review. *16*(1), e12855.

[10] Munier, N., & Hontoria, E. (2021). *Uses and Limitations of the AHP Method*: Springer.

[11] News, S. (2025). Gaofen-3 satellite assists in flood disaster relief in Guilin. Sohu.com. https://www.sohu.com/a/789788556_120502346 (in Chinese: accesed June 25, 2025)

[12] Wu, J., Chen, X., & Lu, J. J. I. J. o. D. R. R. (2022). Assessment of long and short-term flood risk using the multi-criteria analysis model with the AHP-Entropy method in Poyang Lake basin. *75*, 102968.

[13] Zhou, Q., Leng, G., Su, J., & Ren, Y. (2019). Comparison of urbanization and climate change impacts on urban flood volumes: Importance of urban planning and drainage adaptation. *658*, 24-33.