International Journal of Geo-Information

MDPI

Article

Identifying Urban Functional Zones and Multi-Dimensional Flood Risk Connectivity: Integrating POI Semantic Representation and Spatial Network Analysis

Meiyan Gao * and Zongmin Wang

School of Water Conservancy and Transportation, Zhengzhou University, Zhengzhou 450001, China; zmwang@zzu.edu.cn
* Correspondence: gmeiyan123@163.com

Abstract

Refined urban flood risk assessment requires a paradigm beyond traditional single-scale analyses. Existing approaches often overlook fine-scale functional differentiation and cascading risk connectivity across urban systems. This study proposes an integrated framework combining semantic–spatial representation of Point-of-Interest (POI) data, hierarchical feature integration, and a hydrological–functional–road tripartite potential flood risk connectivity network. A semantic–spatial dual-branch autoencoder extracts POI features for functional zone identification, while hazard, exposure, and coping capacity indicators are integrated to quantify block-scale flood risk. Hydrological connectivity, functional coupling, and road cascading failure models are further combined to characterize multi-path flood risk connectivity patterns. Results show that (1) functional zones were derived from dominant POI categories. The semantic–spatial clustering achieved a silhouette coefficient of 0.564 with 0.2% noise blocks. The spatial-only configuration achieved 0.554. This suggests that semantic information provided complementary but limited refinement. (2) A total of 302 high-risk blocks (20.5%) were identified, while the median-based classification showed that public service zones had the highest proportion of the high-risk group (61.3%). (3) The coupled flood risk connectivity network covers 99.2% of blocks with significant spatial dependence (Moran's I = 0.703). Connectivity intensity varied among functional zones, but most pairwise differences were not significant after Bonferroni correction. Network centrality showed significant associations with connectivity intensity and range. (4) The mean standardized functional mixture degree was 0.629. Geographically weighted regression (GWR) revealed spatial variation in the local association between functional mixture and physical flood susceptibility, but the Monte Carlo test did not provide evidence of statistically significant global spatial non-stationarity (p = 0.960). This framework provides a function-oriented perspective for understanding urban flood resilience by integrating functional structure, multidimensional flood risk, and multi-path risk connectivity, and supports the characterization of potential flood-sensitive areas through function-based flood risk assessment.

Keywords: urban functional zones; flood risk connectivity; POI semantic representation; spatial network analysis; functional mixture; risk connectivity network

Check for updates

Academic Editors: Wolfgang Kainz and Godwin Yeboah

Received: 5 August 2026
Revised: 10 September 2026
Accepted: 15 September 2026
Published: 21 September 2026

Copyright: © 2026 by the authors. Published by MDPI on behalf of the International Society for Photogrammetry and Remote Sensing. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Urban functional zones serve as spatial carriers of diverse socioeconomic activities and urban services, providing important units for understanding spatial organization,

resource allocation, and urban governance. Their accurate identification has significant theoretical and practical implications. With the continuous expansion of cities and increasing complexity of urban configurations [1], conventional approaches can no longer meet the demand for fine-grained cognition and management of urban spaces. As spatial representations of urban functions, these zones support activities including transportation, residence, recreation, education, healthcare, and administration. Their identification also provides valuable information for urban planning and disaster risk assessment.

In recent years, multi-source geographic big data have offered a new foundation for identifying urban functional structures. Among them, Point-of-Interest (POI) data are particularly important. Existing studies predominantly rely on POI density statistics [2], spatial clustering [3], or semantic embedding methods [4] to delineate urban functional zones. However, in complex urban settings, POI data frequently exhibit both semantic sparsity and spatial heterogeneity [5]. This makes it difficult for conventional approaches to capture the underlying patterns of urban functional organization [3,6]. For functional mixture quantification, information entropy-based approaches have become mainstream, yet conventional entropy models may have limitations when certain POI categories are absent or unevenly distributed [7]. Consequently, achieving robust functional zone delineation and functional mixture characterization that accommodates both semantic information and spatial structure remains a key challenge.

Building on the analysis of urban functional structure, flood risk assessment typically considers hazard, exposure, and coping capacity. Vulnerability aspects are partly represented through exposure and coping capacity indicators [8,9]. Existing studies seldom systematically incorporate the urban functional organization characteristics depicted by POIs into the flood risk analysis framework. The intrinsic coupling between urban functional structure and flood risk therefore remains insufficiently understood [10]. Moreover, POI semantic representations and flood risk indicators differ substantially in feature characteristics and dimensionality. Simple concatenation or weighted fusion can therefore easily cause information imbalance and weaken the integration of functional and risk-related information [11]. An effective feature fusion strategy that enables interaction between functional and risk information therefore remains an important research need. In addition to evaluating composite flood risk, separating physical hazard conditions from exposure-related consequences is essential when examining the relationship between urban functional structure and flooding.

Although POI data are increasingly incorporated into flood risk assessment, their analytical roles vary considerably across studies. One line of research has treated POIs primarily as indicators of exposed assets and activity locations, often by combining them with population data to refine the spatial representation of flood impacts [12]. Another line of work has integrated POI-derived information into vulnerability or exposure assessment by linking POI records with mobility patterns to estimate flood-affected populations [13]. A more function-oriented approach has emerged in studies that use POI data to delineate functional zones and then examine flood vulnerability variations across these zones [10]. More recently, POI and area-of-interest data have been jointly employed to identify building functions, which are then combined with flood susceptibility to evaluate building-level risk [14]. Despite these advances, the explicit representation of the semantic and spatial heterogeneity of POIs for functional zone identification remains limited. In addition, the propagation of flood risk among functional units through multiple connectivity pathways has received less attention.

Furthermore, urban flood impacts exhibit strong cascading propagation characteristics, driven jointly by hydrological processes, functional dependencies, and infrastructure networks [15]. Previous studies have mainly investigated hydrological connectivity, functional

interactions, and road-network cascading effects as separate pathways [16]. Hydrological connectivity studies have focused on physical flow processes, functional coupling studies have emphasized human activity patterns derived from POI data, and road-network studies have examined infrastructure cascading. However, few studies have integrated these three types of pathways within a unified network framework. To better characterize multi-pathway flood risk propagation, this study integrates hydrological connectivity, functional coupling, and road-network cascading effects within a unified analytical framework. Additionally, the complexity of urban spatial structure leads to marked spatial heterogeneity in risk distribution. Traditional methods based on information entropy or local spatial statistics still have limitations in depicting functional mixture patterns and their spatial variations [17]. Local modeling approaches such as Geographically Weighted Regression (GWR) offer a pathway to reveal this spatial heterogeneity [18]; yet their application in the urban flood domain remains relatively limited.

Addressing these gaps, this study constructs a comprehensive analytical framework that integrates POI semantic representation, a multi-dimensional flood risk indicator system, and a hydrological–functional–road tripartite potential flood risk connectivity network. Taking Zhengzhou City as a case study, the framework characterizes the intrinsic relationships between urban functional structure and flood risk from three perspectives: functional representation, function–risk coupling, and risk connectivity. The main contributions of this paper are: (1) developing a semantic–spatial joint representation strategy that integrates spatial configuration and POI composition to reveal latent functional structures and support functional-zone identification; (2) integrating functional attributes into flood risk assessment to examine whether risk patterns differ across functional zones after controlling for network position; (3) establishing a hydrological–functional–road tripartite potential flood risk connectivity network to characterize multi-path connectivity mechanisms; and (4) evaluating the spatial association between functional mixture and physical flood susceptibility using GWR, while testing whether the local relationships exhibit statistically significant spatial non-stationarity.

2. Materials and Methods

2.1. Study Area

Zhengzhou City is located in north-central China, at the transitional zone between the middle and lower reaches of the Yellow River. The study area covers the central urban area consisting of five administrative districts, approximately between 113°27′ E and 113°52′ E, and 34°36′ N and 34°58′ N (Figure 1). The terrain descends from southwest to northeast, characterized by hilly landscapes in the southwest and the Yellow River alluvial plain in the northeast. Zhengzhou was severely affected by the "7·20" extreme rainfall event in 2021, making it a representative case for investigating urban flood risk mechanisms. Extensive impervious surfaces accelerate surface runoff generation and increase the likelihood of urban flooding during intense rainfall events. The rainy season extends from June to September, during which approximately 70% of annual precipitation occurs, with a mean annual precipitation of about 640 mm. This study focuses on the central urban area, covering five administrative districts, including Huiji, Jinshui, Zhongyuan, Erqi, and Guancheng. This densely developed urban core features concentrated urban functions, high population density, and complex road networks, providing a representative setting for examining how urban functional structures influence flood risk connectivity.

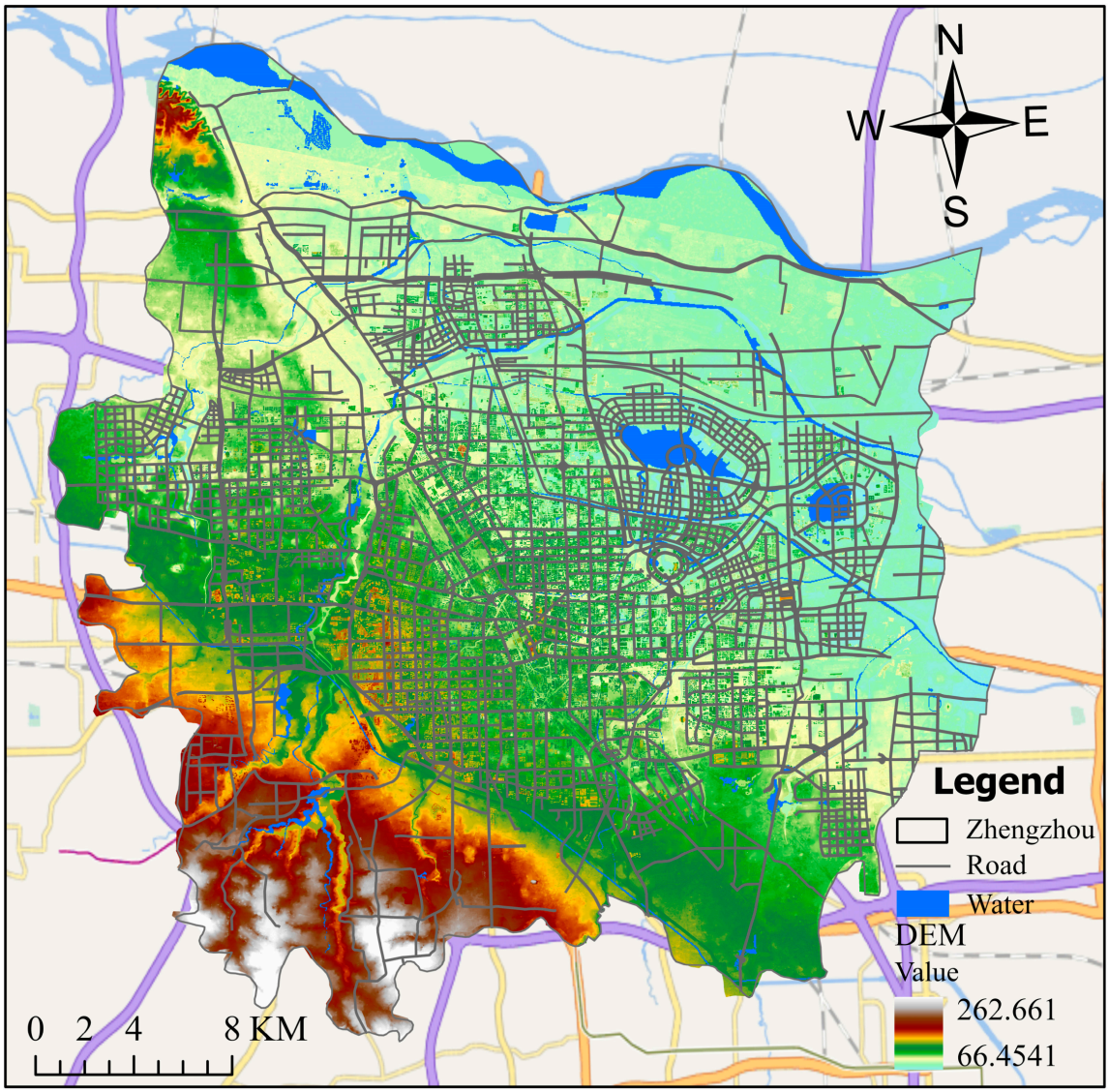


Figure 1. Overview of the study area in Zhengzhou City.

2.2. Data

The datasets used in this study include Point-of-Interest (POI) data, road network data, topographic and land use data, and waterlogging point data. POI data were obtained from AMap (8963 records, https://lbs.amap.com/ (accessed on 14 September 2026)) and OpenStreetMap (250,368 records, https://www.openstreetmap.org/ (accessed on 14 September 2026)). Additional emergency shelter locations were collected from official sources (https://m.henan.gov.cn/2022/07-20/2490634.html (accessed on 14 September 2026)) and converted into POI format. After deduplication, name standardization, and coordinate correction, 106,010 high-quality POI records were retained and reclassified into five functional categories: residential, industrial, commercial, ecological and leisure, and public service. Historical land use data (1985, 1995, 2023) were derived from the China Land Cover Dataset (CLCD) at 30 m resolution [19], providing a basis for analyzing urban expansion and its relationship with flood risk. All spatial datasets involved in spatial analysis were transformed into a projected coordinate system suitable for metric operations (CGCS2000/3-degree Gauss-Kruger zone 38, EPSG:4528). Spatial operations, including distance, area, buffering, and Euclidean distance calculations, were conducted in meters.

Road network data were acquired from OpenStreetMap and processed by removing auxiliary roads, sidewalks, and non-motorized lanes while retaining urban express-

ways, elevated roads, and primary to quaternary roads. Based on road buffer analysis, 1474 traffic analysis zones were generated as the basic spatial units for subsequent functional zone identification and flood risk analysis (Figure 2). Topographic data were derived from a ZY-3 satellite-based digital elevation model with a spatial resolution of 5 m. Land use data included a 1 m resolution land cover dataset [20] and building footprint distributions, which were used to characterize urban physical conditions and exposure-related factors. In addition, the official comprehensive plan of Zhengzhou (2010–2020) was obtained from the Zhengzhou Municipal Natural Resources and Planning Bureau (https://zrzyhghj.zhengzhou.gov.cn/tdlygh/8354220.jhtml (accessed on 14 September 2026)). This completed plan covers the central urban area at a block-level resolution and was used as an independent reference for the manual validation of functional-zone labels, rather than for the automated classification.



Figure 2. Basic functional blocks in the study area.

Waterlogging point data were compiled from multiple sources, including field surveys, official reports, government websites, and publicly available online information during the "7·20" extreme rainfall event. After screening and verification, a total of 905 validation points were compiled, including 732 field-observed flood depth points and 173 multi-source flood-affected locations obtained from official reports and web-based sources. These multi-source observations were integrated as independent validation data for evaluating the spatial consistency of the identified high-risk blocks.

2.3. Methods

This study develops a comprehensive analytical framework integrating POI semantic–spatial representation, multidimensional flood risk characterization, and a multi-path flood risk connectivity network (Figure 3). The framework comprises four interconnected components. The first component performs urban functional zone delineation based on semantic–spatial POI representations. A semantic–spatial dual-branch autoencoder extracts heterogeneous POI features, and a two-stage clustering method achieves refined functional zone identification. The second component performs POI and flood risk feature fusion, where a multidimensional flood risk indicator system is constructed with CRITIC-based weighting and a cascaded feature organization strategy to formulate a function–risk classification. The third component conducts multi-path risk connectivity analysis by integrating hydrological connectivity, a functional coupling gravity model, and a road cascading failure model to characterize the tripartite risk connectivity mechanism. The fourth component analyzes the coupling relationship between functional mixture and flood risk through an improved entropy-based model and Geographically Weighted Regression (GWR) to reveal the spatial non-stationarity of the mixture–risk relationship.

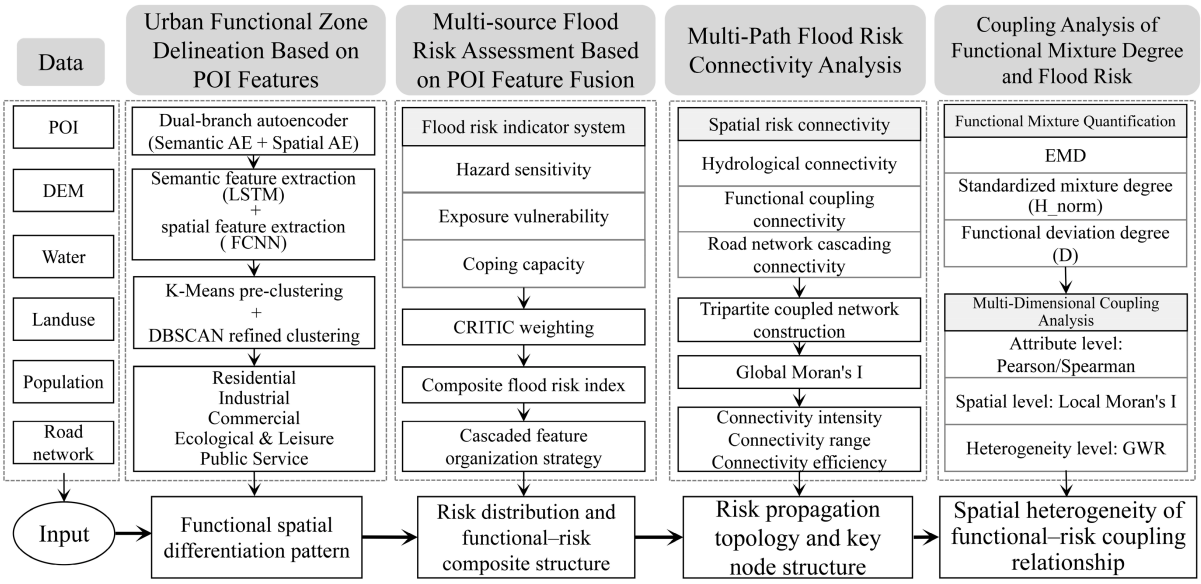


Figure 3. Technical roadmap of the study.

2.3.1. Urban Functional Zone Delineation Based on POI Features

POI preprocessing was conducted using POI records aggregated within the 1474 spatial units, which served as the basic analytical units, with the workflow illustrated in Figure 4. The original POI dataset contained 259,331 records. After deduplication, name standardization, and coordinate validation, 106,010 valid POI records were retained. Common POI categories such as restaurants and convenience stores were removed to reduce background noise, following the approach of [21]. This removal mainly reduced generic high-frequency POI signals, while functional zone identification relied on semantic–spatial representation rather than POI counts alone, allowing residential areas to be identified through their spatial characteristics. The DBSCAN algorithm was subsequently applied to remove spatial noise points [22]. The cleaned POIs were reclassified into five functional zone categories: Residential, Industrial, Commercial, Ecological and Leisure, and Public Service (Table 1).

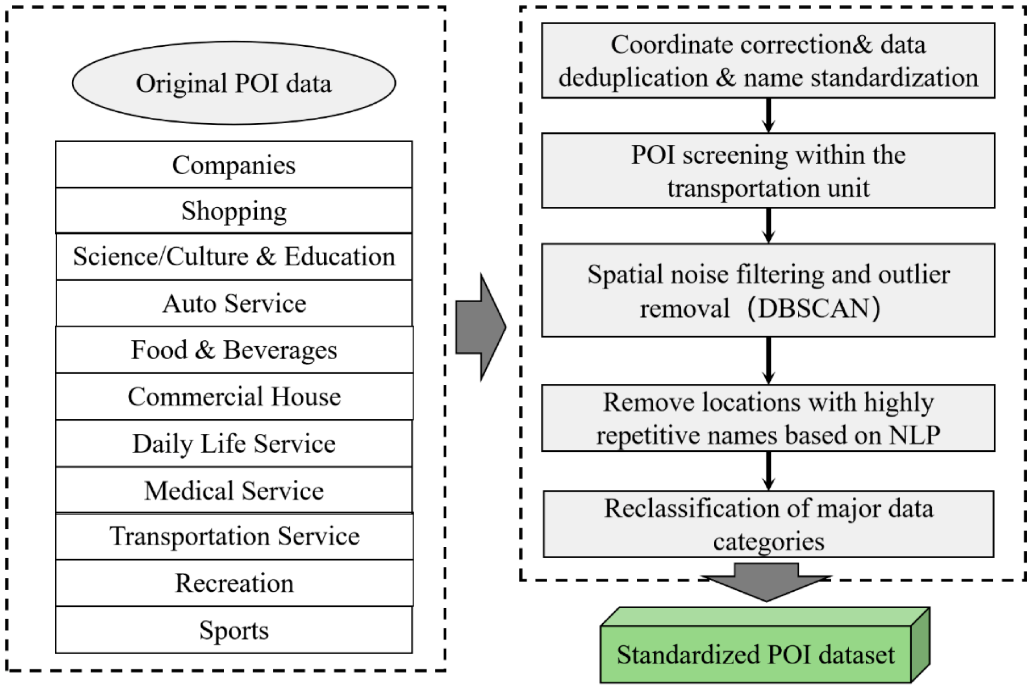


Figure 4. Flowchart of POI data preprocessing.

Table 1. POI reclassification system and flood-related implications of functional zones.

Functional Zone Type	Core Flood-Related Characteristics	Original Reclassified Subcategories	POI Count	Proportion	Key Flood Association Points
Residential	High population concentration, abundant water-prone spaces like underground garages	Residential	7136	6.73%	Population evacuation, underground space flood prevention, community emergency capacity
Industrial	Concentration of industrial and office buildings, some containing storage facilities	Companies and industrial parks	23,571	22.23%	Industrial facility flood prevention, hazardous material control, post-disaster industrial recovery
Commercial	High building density, integrated underground commercial and transport spaces, high pedestrian flow	Commercial and entertainment	31,210	29.45%	Underground space drainage, crowd emergency evacuation, commercial facility recovery
Ecological and Leisure	High proportion of green space and water bodies, stormwater regulation capacity, low population density	Parks, squares, scenic spots, sports and leisure	29,113	27.46%	Stormwater regulation, ecological buffering, emergency shelter layout
Public Service	Agglomeration of public services, medical rescue and emergency shelter functions, transport nodes for emergency evacuation	Schools, healthcare, science and cultural venues, train and bus stations	14,980	14.13%	Medical emergency rescue, campus safety, emergency shelter, public transportation

Upon completing POI data cleaning and reclassification, a dual-branch autoencoder (Figure 5) was constructed to extract both semantic and spatial features of POIs [23]. For semantic feature extraction, the five normalized POI category frequencies were arranged

in a fixed order, including residential, industrial, commercial, ecological and leisure, and public service, before being passed through a 64-dimensional fully connected embedding layer, followed by a single-layer LSTM with two hidden units. The LSTM was used as a nonlinear encoder to obtain a compact representation of functional composition rather than as a temporal sequence model. Because the five POI categories do not constitute a natural sequential structure, the predefined order was used only as a consistent input arrangement and was not assigned a temporal interpretation. A standard multilayer perceptron (MLP) was additionally evaluated as an order-independent comparison model [4]. The resulting output proceeded through fully connected layers (32 → 16) to produce a 16-dimensional semantic feature F_{sem}. For spatial feature extraction, the centroid longitude and latitude coordinates of each block were taken as input. These coordinates passed through two fully connected layers (64 → 32 → 4) to yield a 4-dimensional spatial feature F_{spa}. Concatenating F_{sem} and F_{spa} produced a 20-dimensional latent POI representation $F_{POI}^{(base)}$. The dual-branch representation learning framework was optimized using Mean Squared Error (MSE) loss and the Adam optimizer. The learning rate was 0.001, the batch size was 32, and the model was trained for 10 epochs with a 7:3 training-validation split. The mean cosine similarity between the reconstructed frequency vector and the original vector reached 0.962 on the validation set, indicating that the compressed representation retained most of the original functional composition information after dimensionality reduction [24].

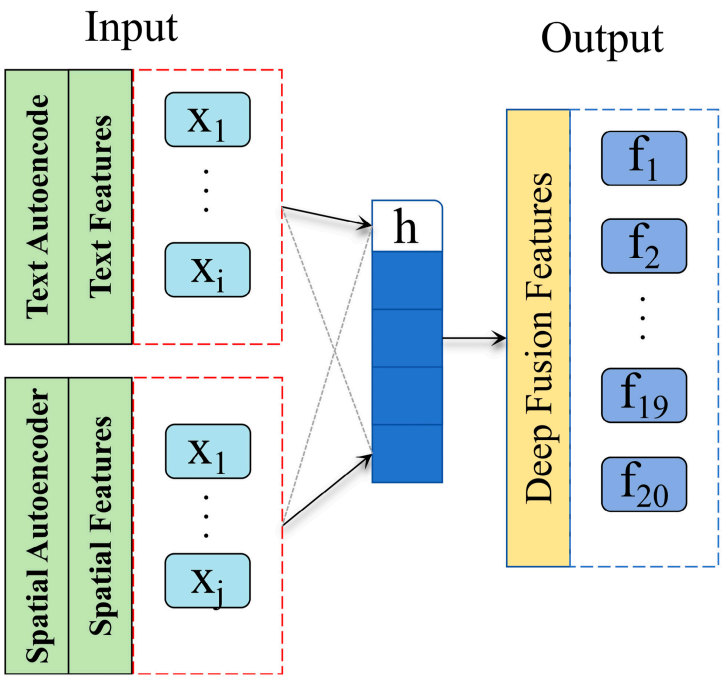


Figure 5. Architecture of the semantic–spatial dual-branch autoencoder.

Functional zone identification employed a two-stage method combining K-Means pre-clustering with DBSCAN fine clustering. The first stage applied K-Means to the 20-dimensional fused features, with the optimal number of clusters determined by the elbow method and silhouette coefficient [25,26]. Second, DBSCAN was applied within each initial cluster to refine cluster boundaries and identify noise points, with Euclidean distance and a Ball-Tree index used for efficient neighborhood search. The DBSCAN parameters were selected via a grid search over eps ∈ [0.1, 1.0] and min_samples ∈ [2, 10]. The final values were eps = 0.5 and min_samples = 3 [27]. The final functional-zone label of each block was assigned according to the block's own dominant POI category, defined as the category with the highest normalized frequency across the five reclassified POI

types listed in Table 1. For blocks without valid POI records, the functional-zone label was adopted from the existing functional-zone attribute in the basic block dataset. This attribute was determined during the initial block data preparation by integrating the original POI distribution, planning land-use information, and the functional characteristics of adjacent blocks. Such blocks were retained for subsequent spatial and risk analyses but were not included in the POI-based clustering consistency evaluation. The two-stage clustering was performed based on the semantic–spatial POI features of the blocks and was used to characterize the latent spatial-functional organization of urban blocks, as well as to examine whether blocks with similar learned representations exhibited coherent functional patterns. The agreement between block-level functional-zone labels and cluster assignments was evaluated using Cohen's kappa. Functional labels and clustering results therefore serve complementary purposes. Functional labels indicate the dominant function of each spatial unit, while clustering reveals the underlying spatial-functional structure of the urban area. Functional-zone labels were determined by the dominant POI category of each block. The clustering results were used to characterize latent spatial-functional organization and to check consistency with these labels, rather than to generate the labels directly. An independent manual validation was conducted for a stratified random sample of 50 blocks, with 10 blocks selected from each of the five functional-zone categories. The functional labels were independently assessed by considering urban planning layouts, satellite imagery and street-view imagery from online map services such as Amap and other available visual or land-use information. Each sampled block was assigned to the functional category that best represented its predominant observable urban use. Agreement between the automated block-level labels and the manual reference labels was evaluated using overall agreement and Cohen's kappa. All clustering results reported in this manuscript are based on the final POI preprocessing and feature extraction pipeline. The corresponding metrics were recalculated using this final pipeline.

2.3.2. POI and Flood Risk Feature Fusion for Function–Risk Characterization

Based on the functional zone delineation, flood risk feature systems were established for block and road units. Multi-source data fusion approaches have been increasingly adopted for urban spatial analysis, particularly integrating POI data with diverse spatial datasets to improve characterization accuracy [28]. For block units, eight indicators were selected across three dimensions: hazard sensitivity, covering elevation, mean slope, distance to water system buffer, and impervious surface ratio; exposure vulnerability, covering population density, underground space ratio, and functional agglomeration degree; and coping capacity, represented by accessibility to emergency shelters [29,30]. For road units, four indicators were extracted: road grade, mean elevation, mean slope, and low-lying area flag. After min-max normalization, the CRITIC method was employed to calculate the composite flood risk index for each unit [31]. The CRITIC method rests on two criteria. Contrast intensity is measured by the standard deviation and reflects the indicator's ability to distinguish risk levels among blocks. Conflict is quantified by the differences among indicators based on pairwise correlation coefficients and reduces information redundancy caused by highly correlated indicators. The CRITIC method was selected over the entropy weighting method because it simultaneously considers both indicator contrast intensity and inter-indicator conflict. This is particularly relevant for this study because several indicators, such as elevation, slope, and distance to water bodies, are physically correlated. The entropy weighting method, by contrast, assigns weights solely on the basis of indicator variability and therefore tends to place excessive weight on a single variable with high dispersion, such as underground space ratio, while nearly ignoring physically important variables such as distance to water bodies. Table 2 compares the weights obtained by the

CRITIC and entropy weighting methods, showing that the entropy method assigned a weight of 0.521 to underground space ratio but only 0.011 to distance to water bodies, whereas the CRITIC method provided a more balanced distribution. All eight indicators were incorporated into the core flood risk feature set F^*. The composite information amount was defined as the product of standard deviation and conflict, with the weight obtained after normalization (Table 3). A weighting sensitivity analysis was also conducted using the entropy weights under the same high-risk identification and validation procedures.

Table 2. Weight comparison of the two weighting methods.

Indicator	CRITIC Weight	Entropy Weight
Elevation	0.097	0.002
Mean Slope	0.084	0.002
Distance to Water System Buffer	0.198	0.011
Impervious Surface Ratio	0.204	0.046
Population Density	0.074	0.200
Underground Space Ratio	0.105	0.521
Functional Agglomeration Degree	0.078	0.209
Accessibility to Emergency Shelters	0.160	0.010

Table 3. Flood risk feature indicator system.

Feature Dimension	Core Indicator	Indicator Attribute	Flood Risk Connotation	CRITIC Weight
Hazard Sensitivity	Elevation	Negative	Lower elevation implies higher waterlogging susceptibility	0.097
	Mean Slope	Negative	Smaller slope leads to slower surface runoff and higher ponding risk	0.084
	Distance to Water System Buffer	Negative	Closer to water bodies elevates flood exposure risk	0.198
	Impervious Surface Ratio	Positive	Higher ratio enhances runoff generation and reduces infiltration	0.204
Exposure Vulnerability	Population Density	Positive	Higher density increases casualty and evacuation pressure	0.074
	Underground Space Ratio	Positive	Higher ratio raises underground flooding and facility damage risk	0.105
	Functional Agglomeration Degree	Positive	Higher agglomeration implies greater economic loss from flood damage	0.078
Coping Capacity	Accessibility to Emergency Shelters	Negative	Closer distance improves emergency evacuation and shelter efficiency	0.160

To evaluate the fusion potential between POI and flood risk features, an attention-based fusion model using a multilayer perceptron was constructed for an ablation study [32,33]. The 20-dimensional latent POI representations and the 8-dimensional flood risk features

were adaptively weighted and fused to generate a 24-dimensional fused feature F_{fusion}, with the formulation as follows:

$$\begin{aligned} F'_{POI} &= \mathrm{ReLU}\Big(\mathrm{BN}(W_{POI}F^{(base)}_{POI} + b_{POI})\Big) \\ F'_{risk} &= \mathrm{ReLU}\Big(\mathrm{BN}(W_{risk}F^{*} + b_{risk})\Big) \\ \alpha &= \sigma\big(W_{\alpha}F'_{POI} + b_{\alpha}\big) \\ \beta &= \sigma\big(W_{\beta}F'_{risk} + b_{\beta}\big) \\ F_{fusion} &= \mathrm{MLP}\Big(\mathrm{BN}\big(\lambda{\cdot}\alpha \odot F'_{POI} + (1-\lambda)\cdot\beta \odot F'_{risk}\big)\Big) \end{aligned} \tag{1}$$

where $F^{(base)}_{POI} \in R^{20}$ and $F^{*} \in R^{8}$ are the latent POI representations and flood risk core features, respectively; W_{POI}, W_{risk}, W_{α}, W_{β} and b_{POI}, b_{risk}, b_{α}, b_{β} denote the corresponding weight matrices and bias terms; σ represents the sigmoid activation function; BN denotes batch normalization; and ⊙ indicates the Hadamard product. The global fusion weight λ was set to 0.5 following sensitivity analysis, during which the downstream clustering silhouette coefficient fluctuated by less than 0.02 over the interval [0.3, 0.7], indicating model insensitivity to λ. Five feature combination strategies were compared through an ablation experiment [34] using K-Means clustering performance as the evaluation criterion, with cluster numbers ranging from 2 to 30 [25]. The relative information contribution of POI and flood risk features was examined to assess potential feature imbalance. Consequently, a cascaded feature organization strategy was adopted. At the first level, the 20-dimensional latent POI representations were used to divide the study area into five urban functional zones [4]. At the second level, each block was assigned to a higher-risk or lower-risk group using the median of the composite flood risk index as the threshold, and these categories were subsequently crossed with the five functional zones to form a 10-class function–risk classification system. This strategy maintained the independence of POI features and flood risk features. It also avoided the double-counting issue arising from flood risk indicators that already contain POI-derived metrics, specifically underground space ratio and functional agglomeration degree.

Before identifying extreme high-risk blocks, the median of the composite flood risk index was used to assign blocks to higher-risk and lower-risk groups for the function–risk classification. After calculating the composite flood risk index, a combined strategy of statistical quantiles and physical constraints was employed to identify high-risk blocks. The statistical threshold was set at the 85th percentile of the composite index. The physical threshold was defined as units simultaneously meeting an elevation below the 10th percentile of the entire domain and a distance to water buffer of less than 500 m. The final high-risk blocks were determined by the union of both criteria. The detailed threshold selection is provided in Appendix A, and the distribution of the composite flood risk index with the 85th percentile threshold is shown in Figure A1. Validation was conducted using observed waterlogging records from the "7·20" extreme rainfall event, with detailed accuracy metrics reported in Section 3.2. The random permutation baseline was evaluated over 1000 independent permutations. In each permutation, the predicted high-risk labels were randomly reassigned while preserving the number of high-risk predictions. Validation metrics were calculated using the same adjusted precision, recall, and F1 procedure as the integrated index. Sensitivity analyses were conducted by varying the spatial matching tolerance from 30 m to 150 m and the high-risk statistical threshold from the 80th to the 90th percentile while keeping the physical constraint unchanged. In addition, a weighting sensitivity analysis was performed by replacing the CRITIC weights with entropy weights and reapplying the same high-risk identification and validation procedures.

For the historical land-cover analysis, the CLCD land-cover classes occurring in the study area were grouped into four broad categories: built-up land (impervious surface), cultivated land (cropland), ecological land (forest, grassland, and water), and other land (barren land). The same grouping scheme was applied to the 1985, 1995, and 2023 maps to ensure temporal comparability. The reclassified CLCD raster maps were then overlaid with the current block boundaries to determine the dominant land-cover category of each block. For each year, the dominant category was identified according to the largest proportion of block area occupied by each land-cover category. Blocks whose dominant land-cover category was built-up were classified as built-up for that year. Blocks classified as non-built-up in 1995 but built-up in 2023 were identified as associated with post-1995 built-up expansion. These blocks were then linked to the current functional-zone labels and high-risk classification to examine the relationship between historical urban expansion and present-day flood risk. For each functional category, the reported proportion of high-risk blocks was calculated as the number of high-risk blocks within the corresponding post-1995 built-up group divided by the total number of blocks of the same functional category within that group.

For road units, flood-prone sections were identified using elevation and slope as discriminant factors. The interquartile range of elevation was calculated for all road segments [35,36]. Segments with elevations below the 25th percentile were marked as relatively low-lying [37,38]. From these, segments with a slope lower than the domain median were retained as the final flood-prone road sections, which were classified into four categories: main roads, secondary roads, branch roads, and intersections [39,40].

2.3.3. Multi-Path Flood Risk Connectivity Analysis

Potential flood risk connectivity networks were constructed to represent possible interaction structures among urban units, rather than directly observed flood propagation. The network construction considered topographic-based hydrological connectivity, inferred functional association, and static road topology with cascading failure simulation as three types of spatial relationships.

For topographic-based hydrological connectivity, a directed network was approximated using a topographic proxy based on block centroid distances and mean elevations [41]. A hydrological upstream–downstream relationship was established when the centroid distance between two blocks did not exceed 500 m and the elevation difference was greater than 1 m. Under these conditions, a hydrological connectivity edge was formed, pointing from the higher-altitude block to the lower-altitude one. This directed network represents topographic-based potential hydrological connectivity rather than actual surface flow paths.

For inferred functional association, a gravity model was employed to estimate inferred functional associations between blocks based on POI counts and spatial proximity [13,42]. The model takes the form $F_{ij} = M_i M_j / d_{ij}^2$, where M_i and M_j represent the total POI counts of blocks i and j, and d_{ij} is the centroid Euclidean distance serving as a distance decay factor. Block pairs with the highest 5% of inferred association values were retained as functional coupling edges. This network reflects potential functional interactions rather than observed functional dependency.

For the road network component, road segments and adjacent blocks were represented as a static topological layer, where an edge indicates spatial adjacency and potential connectivity between two blocks. A cascading failure model was introduced to simulate the dynamic redistribution of risk along the road network under flood shock [43,44]. The model designated the 10 blocks with the highest risk indices as initial failure nodes to trigger the cascading process. Each node i was assigned an initial load L_i equal to its normalized

composite flood risk index. Node capacity was defined as $C_i = (1 + \alpha)L_i$, with $\alpha = 0.3$, representing a moderate tolerance level. When node i failed, its load was redistributed to its direct road-network neighbors $j \in N_i$ according to their remaining capacity:

$$\Delta L_j = \frac{C_j - L_j}{\sum_{k \in N_i}(C_k - L_k)} L_i \quad (2)$$

The load of each neighbor was updated as $L_j \leftarrow L_j + \Delta L_j$. A neighboring node j failed if $L_j > C_j$. The process continued until no new failure occurred. Static road edges were included in the coupled network, whereas failure-propagation edges were recorded separately during the simulation. This iterative mechanism captures the reconfiguration and risk diffusion effects of road network connectivity after local damage.

The ZY-3 DEM used in this study has a nominal vertical accuracy of approximately 1.6 m under sparse ground control points [45], and block-level elevation was averaged from all DEM pixels within each block to reduce pixel-scale uncertainty. Assuming independent elevation errors among DEM pixels and treating the reported vertical accuracy as a conservative estimate of pixel-level uncertainty, the standard error of block-mean elevation can be approximated as $1.6/\sqrt{N}$ m, where N is the number of DEM pixels within a block. For the blocks used in this study, N ranged from approximately 500 to over 20,000, yielding standard errors from about 0.07 m to less than 0.01 m, which are substantially smaller than the 1 m elevation-difference threshold used for hydrological connectivity identification. Hydrological edges were constructed using a centroid distance threshold of 500 m and an elevation difference threshold of 1 m, with sensitivity analyses performed at 300, 500, and 700 m and 0.5, 1.0, and 2.0 m, respectively. The gravity model used a distance decay exponent of 2 [46] and POI count as the mass indicator. The top 5% strongest functional connections were retained as the backbone network [47], with sensitivity tested at exponents of 1.5 and 2.5 and cutoffs of 3% and 10%. For each parameter setting, edge counts, mean degree, top-20 degree and betweenness centrality nodes, functional zone rankings, and cascading failure scale were compared with the baseline.

The three types of edges were merged to build a unified coupled adjacency matrix. Based on this matrix, the global Moran's I statistic was calculated to assess the spatial autocorrelation of flood risk under the coupled connectivity structure. This statistic tested whether the risk distribution exhibited significant spatial clustering characteristics [48]. Quantitative evaluation was conducted from three dimensions: connectivity intensity, representing the strength of structural connections associated with potential flood risk interactions; connectivity range, depicting the number of blocks and spatial span reachable within the network; and connectivity efficiency, focusing on the structural propagation capacity under cascading failure scenarios. This evaluation framework integrates the topographic constraints represented by the hydrological layer with the topological accessibility of the road network, enabling an assessment of potential flood risk connectivity and propagation patterns.

To examine whether connectivity metrics differed among functional zones and whether any observed differences were independent of network position, statistical analyses were conducted. The Kruskal–Wallis test and Bonferroni-corrected Mann–Whitney U tests were used to compare connectivity intensity and connectivity range across the five functional zones [18]. Effect sizes were estimated using rank-biserial correlations. In addition, ordinary least squares regressions were performed with connectivity metrics as dependent variables and functional zone dummy variables and network degree centrality as predictors [49]. Degree centrality was standardized using z-score transformation before regression analysis, while connectivity intensity and connectivity range were retained in their original scales.

2.3.4. Coupling Analysis of Functional Mixture Degree and Flood Risk

An improved entropy-based mixed degree (EMD) model was adopted to quantify functional mixture [50]. To address the limitation of traditional entropy when POI type proportions are zero, a subset of effective POI types T'_u was defined, and the proportions were re-normalized [51]. The improved entropy-based mixture degree is defined as:

$$H(u) = -\sum_{t_i \in T'_u} p'_{u,i} \ln(p'_{u,i}) \quad (3)$$

where $p'_{u,i}$ represents the re-normalized proportion of the i-th POI type. $H(u)$ ranges from 0 to ln(ku), with $k_u = \mid T'_u \mid$ being the number of effective POI types within the unit. A value of $H(u) = 0$ indicates that the unit contains only one effective POI type, corresponding to a completely mono-functional state. A value of $H(u) = \ln(k_u)$ means that all effective POI types are equally proportioned, representing the maximum equilibrium degree of functional mixture for that unit. Intermediate values reflect the balance and richness of functional mixture, with higher values indicating greater functional diversity. The standardized mixture degree is defined as $H_{norm}(u) = H(u)/\ln(k_u) \in [0, 1]$. The functional deviation degree $D(u)$ is defined as the standardized absolute deviation between the unit's mixture degree and the typical entropy value $\overline{H}(F)$ of its functional zone:

$$D(u) = \frac{|H(u) - \overline{H}(F)|}{\max|H(u) - \overline{H}(F)| - \min|H(u) - \overline{H}(F)|} \quad (4)$$

Coupling analysis was conducted at three levels: attribute, spatial, and heterogeneity. At the attribute level, the Pearson and Spearman correlation coefficients were calculated between the composite flood risk index and both the standardized mixture degree $H_{norm}(u)$ and the functional deviation degree $D(u)$. These coefficients tested the strength and direction of linear and monotonic associations, respectively. Locally weighted scatterplot smoothing was subsequently applied for nonparametric fitting of the bivariate relationship to identify potential nonlinear trends. At the spatial level, local Moran's I was used to identify the spatial clustering types of mixture degree [52,53]. A KNN (k = 8) spatial weight matrix was adopted to ensure each block had the same number of spatial neighbors, thereby avoiding the "island" problem caused by uneven block area distribution in the study area [54]. At the heterogeneity level, GWR was employed [55,56]. To avoid mathematical coupling between the dependent and explanatory variables, a physical flood susceptibility index was constructed from elevation, slope, and distance to the water system using normalized CRITIC weights. These indicators represent relatively independent physical susceptibility factors and exclude exposure, functional, and POI-derived components contained in the original composite risk index. This physical flood susceptibility index served as the dependent variable, while $H_{norm}(u)$ and $D(u)$ were the sole explanatory variables. Before modeling, a variance inflation factor (VIF) test was applied to the explanatory variables to check multicollinearity [57]. Variables with VIF greater than 10 were excluded, following the commonly used threshold. An adaptive bisquare kernel was used for spatial weighting, and the bandwidth was optimized by minimizing the corrected Akaike Information Criterion (AICc) [58]. Local regressions were estimated using weighted least squares [59]. Model performance was evaluated using R^2 and AICc. Residual spatial autocorrelation was examined using global Moran's I with permutation-based p-values. A Monte Carlo permutation test was conducted to evaluate the statistical significance of spatial variability in the local coefficients.

3. Results

3.1. Urban Functional Zone Identification and Spatial Patterns

Figure 6 presents the spatial distribution of the five urban functional zones across the 1474 blocks. Residential zones comprised 794 blocks, accounting for 53.9% of the total. Industrial zones contained 209 blocks (14.2%), while commercial zones numbered only 32 blocks (2.2%). Ecological and leisure zones included 201 blocks (13.6%), and public service zones consisted of 238 blocks (16.1%). Spatially, the five functional types exhibited distinct organizational patterns. Residential zones were predominantly clustered in the urban core and extended outward to the periphery. Industrial zones formed band-like agglomerations along major transportation corridors on the fringe of the main urban area. Commercial zones were highly concentrated within a 2 km radius of district centers. Ecological and leisure zones displayed a ring-like and patchy distribution along water systems and urban edges. Public service zones showed a combined pattern of core concentration and multi-point dispersion. At the block scale, functional types differed in both area and POI carrying capacity. Industrial zones had the largest average block area (15.8 $\pm$ 12.4 ha), whereas commercial zones had the smallest (5.3 $\pm$ 4.1 ha). Regarding POI density, commercial zones recorded the highest value (186.4 $\pm$ 112.3 POIs·ha^{-1}), and ecological and leisure zones the lowest (21.3 $\pm$ 15.8 POIs·ha^{-1}). Although residential POIs accounted for only 6.73% of all POI records, residential zones comprised 53.9% of blocks, reflecting their extensive spatial coverage and relatively low POI density.

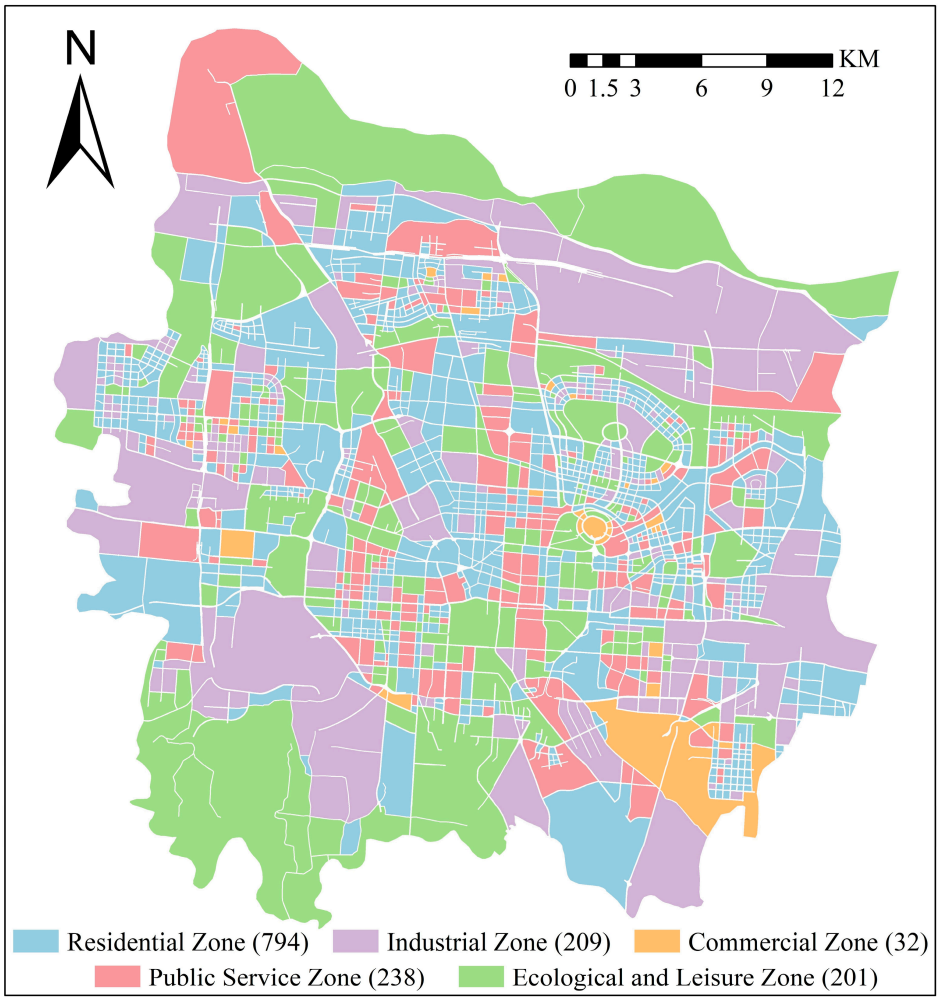


Figure 6. Spatial distribution of the five functional zones.

Figure 7 displays the loss curves of the semantic and spatial autoencoders over the 10 training epochs, and Figure 8 shows the PCA explained variance ratio of the 20-dimensional fused POI features. The training loss of the semantic autoencoder decreased from 0.048 at epoch 1 to 0.009 at epoch 10 (an 81.2% reduction), and the validation loss dropped from 0.041 to 0.008 (an 80.5% reduction), with a train–validation gap of merely 0.001 at the final epoch. The spatial autoencoder's training loss fell from 6748.4 to 0.002 (>99.9% reduction), and its validation loss from 5663.6 to 0.002 (>99.9% reduction). After epoch 3, the spatial loss reached the same order of magnitude as the semantic loss (10^{-2}–10^{-3}). Both autoencoders experienced rapid loss decreases in the first four epochs (72.5% for semantic, >99.9% for spatial). The decline slowed after epoch 5; from epoch 6 to 10, the semantic training loss cumulatively dropped by 48.1%, while the spatial training loss remained within 0.002–0.220 and further stabilized at 0.002. PCA results indicated that the first principal component (PC1) alone explained 31.2% of the total variance, followed by PC2 (18.4%), PC3 (10.2%), PC4 (8.2%), and PC5 (6.1%). The contribution of PC6–PC10 ranged between 2.3% and 5.0%, and each of PC11–PC20 accounted for less than 2.0%. The first five components cumulatively explained 74.1% of the variance; the first ten components reached 92.0%, and fifteen components captured 98.3%. Achieving 95% cumulative explanation required twelve components. PC1 explained 1.7 times the variance of PC2, 3.1 times that of PC3, and 13.6 times that of PC10.

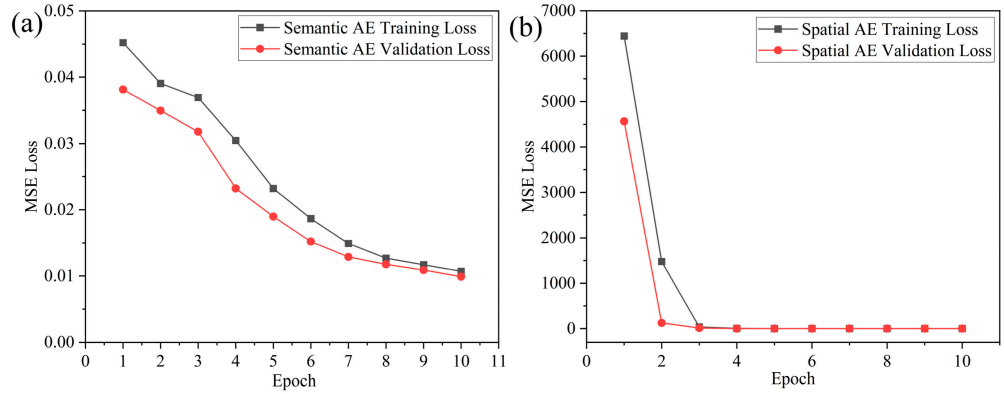


Figure 7. Training and validation loss curves of the autoencoders. (**a**) Semantic autoencoder; (**b**) spatial autoencoder.

Figure 9 presents the clustering results. The highest silhouette coefficient among K values from 2 to 15 was observed at K = 5 (0.568, Figure 9a). Table 4 compares the clustering performance of direct K-Means and the two-stage clustering method. Table 5 summarizes the silhouette coefficients under different feature configurations. The two-stage clustering procedure yielded a silhouette coefficient of 0.564, close to the 0.568 obtained by direct K-Means with K = 5 (Table 4). For the permutation-based analysis across alternative POI-category orderings, the silhouette coefficients ranged from 0.46 to 0.58, with a mean of 0.51. The average cluster purity reached 0.596 for the two-stage method and 0.591 for direct K-Means, and the corresponding Cohen's kappa values relative to the direct dominant-POI classification were 0.609 and 0.598. DBSCAN refinement classified only 2 blocks (0.2%) as noise. The cross-tabulation between final functional-zone labels and cluster assignments is presented in Table 6. Residential blocks were present in all five clusters, with cluster-wise proportions ranging from 48.1% to 59.2%. In terms of dominant POI composition, Cluster 0 was characterized by industrial-related POIs (0.381), Cluster 1 by public-service POIs (0.857), Cluster 2 by commercial POIs (0.475), Cluster 3 by ecological and leisure POIs (0.826), and Cluster 4 by public-service POIs (0.417) (Figure 9b). For the semantic-only feature set (16 dimensions), the clustering silhouette coefficient was 0.473, and replacing the

LSTM component in the semantic branch with a standard multilayer perceptron reduced the coefficient to 0.467 (Table 5). The independent manual validation based on 50 stratified sampled blocks showed an overall agreement of 84.0% (42/50) between the block-level functional labels and the manually interpreted reference labels, with a Cohen's kappa of 0.80. Among the eight misclassified blocks, three were ecological and leisure blocks, two were commercial blocks, and one block each was from the residential, industrial, and public service categories.

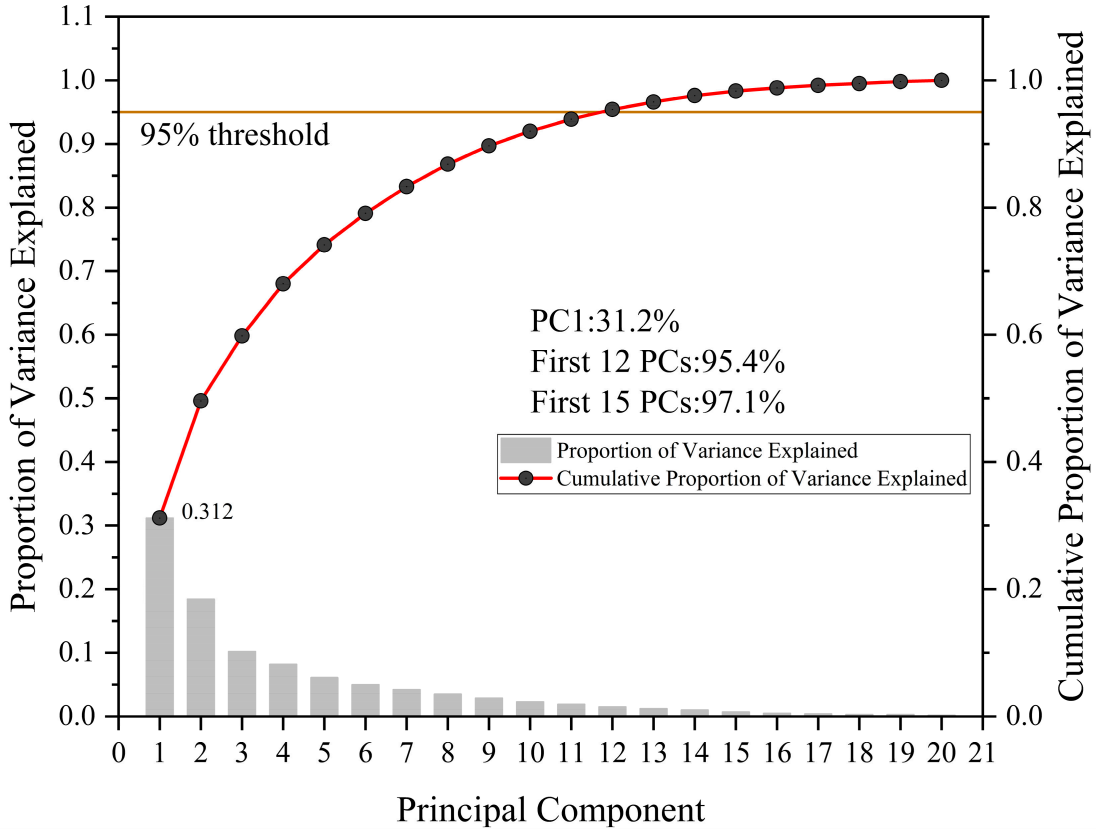


Figure 8. Scree plot of PCA explained variance ratio for the 20-dimensional fused POI features.

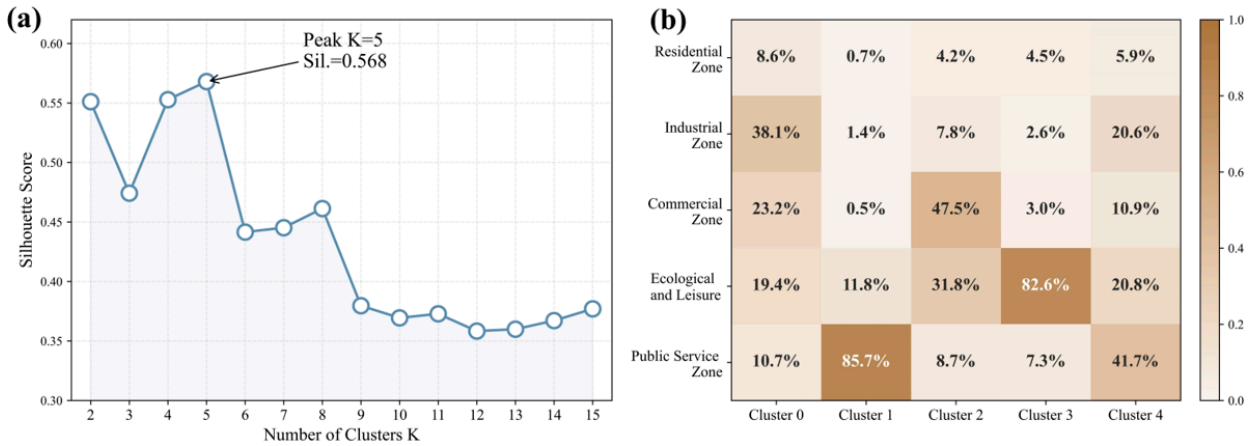


Figure 9. Clustering performance and functional composition of each cluster. (**a**) Variation in the silhouette coefficient with cluster number K (K = 2–15), with the optimum at K = 5 (silhouette coefficient = 0.568). (**b**) Heatmap of the proportion of each functional zone within each cluster (rows: functional zone types; columns: cluster numbers). Cell values indicate the percentage of a functional zone type within the corresponding cluster.

Table 4. Comparison of direct K-Means and two-stage clustering for functional zone identification.

Method	Silhouette	Avg. Cluster Purity	Cohen's Kappa	Noise Ratio
Direct K-Means (K = 5)	0.568	0.591	0.598	0%
Two-stage (K-Means + DBSCAN)	0.564	0.596	0.609	0.2%

Table 5. Clustering silhouette coefficients under different feature configurations.

Feature Configuration	Silhouette
Semantic-only (16D)	0.473
Spatial-only (4D)	0.554
Semantic–spatial (LSTM)	0.564
Semantic–spatial (MLP replacing LSTM)	0.467

Table 6. Cross-tabulation between final functional-zone labels and cluster assignments for blocks included in the clustering analysis.

Functional Zone	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total
Residential	127	40	372	71	41	651
Industrial	53	10	98	17	12	190
Commercial	4	5	15	2	2	28
Ecological and Leisure	32	6	131	15	13	197
Public Service	32	10	157	15	9	223
Total	248	71	773	120	77	1289

3.2. Flood Risk Identification and Function–Risk Classification

Figure 10 presents the spatial distribution of the composite flood risk index and the identified high-risk blocks. Table 7 summarizes the number and proportion of high-risk blocks identified by the 85th percentile and physical threshold criteria across the five functional zones. The mean, median, and standard deviation of the composite flood risk index were 0.508, 0.516, and 0.072, respectively, with a left-skewed distribution (skewness = −0.735) (Figure 10a). The mean index within 500 m river buffers, core areas with an impervious surface ratio above 80%, and blocks in the top 25% by underground space ratio were 0.536, 0.614, and 0.519, respectively, exceeding the domain average by 5.5%, 20.8%, and 2.1%. Using the 85th percentile (0.576) and the physical constraint, 221 statistical and 97 physical high-risk blocks were identified, with 16 overlapping blocks. Their union yielded 302 high-risk blocks (20.5%) (Figure 10b). Residential zones contained the largest number of high-risk blocks (162, 53.6%), followed by public service (51, 16.9%), industrial (44, 14.6%), ecological and leisure (38, 12.6%), and commercial zones (7, 2.3%). The within-zone proportion was highest in commercial zones (21.9%) and lowest in ecological and leisure zones (18.9%).

Figure 11 displays the distribution of the composite flood risk index across the five functional zones. The domain-wide mean and median were 0.508 and 0.516, respectively. The mean values across zones ranged from 0.503 to 0.526, with public service zones showing the highest mean (0.526) and ecological and leisure zones the lowest (0.503), a difference of 0.023. Public service zones recorded the highest median (0.534), compared with 0.497 for commercial zones. All zones except commercial displayed negative skewness, ranging from −0.641 to −1.050; commercial zones, by contrast, exhibited the only positive skewness (0.665). The interquartile range was largest in industrial zones (0.093) and smallest in public service zones (0.071), a ratio of 1.31. Residential zones had the largest standard deviation

(0.075) and interquartile range (0.080) among the five functional types. For public service zones, the corresponding values were 0.064 and 0.071 (Figure 11b).

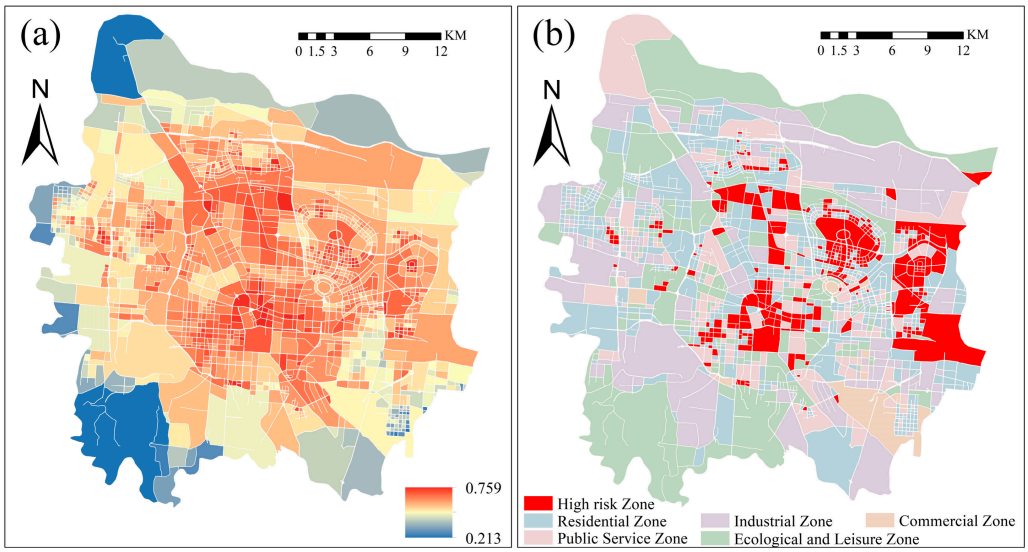


Figure 10. Spatial distribution of the composite flood risk index and identified high-risk blocks. (**a**) Composite flood risk index; (**b**) high-risk blocks.

Table 7. High-risk blocks identified by the 85th percentile and physical threshold criteria by functional zone.

Functional Zone	Total Blocks	High-Risk Blocks	Percentage of All High-Risk Blocks (%)	Within-Zone Proportion (%)
Residential	794	162	53.6	20.4
Industrial	209	44	14.6	21.1
Commercial	32	7	2.3	21.9
Ecological and Leisure	201	38	12.6	18.9
Public Service	238	51	16.9	21.4
Total	1474	302	100.0	20.5

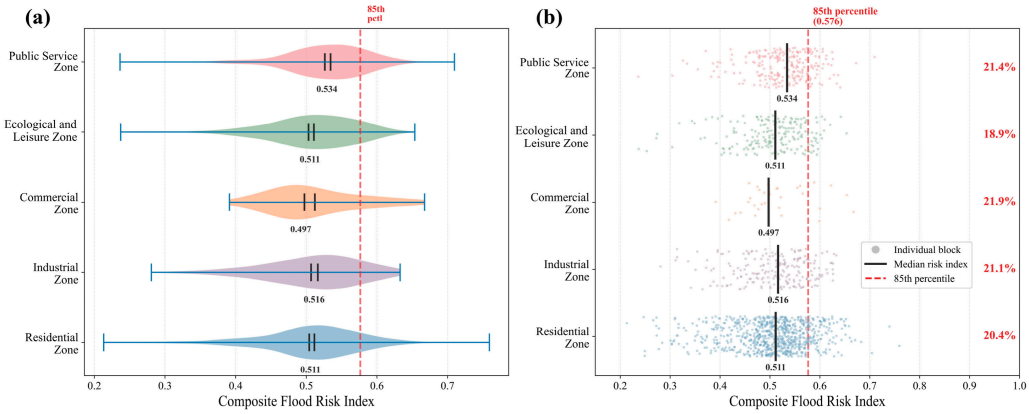


Figure 11. Distribution of the composite flood risk index across the five functional zones. (**a**) Violin plot; (**b**) scatter plot.

Figure 12 presents the spatial distribution of the 10 function–risk combinations formed by crossing the five functional zones with the higher-risk and lower-risk groups defined by the median split of the composite flood risk index, and Table 8 lists the corresponding block counts. Using the median risk index (0.516) as the threshold, the blocks were divided

into a higher-risk group (737 blocks) and a lower-risk group (737 blocks). Residential zones had the largest number of blocks in the higher-risk group (376 blocks, 25.5% of all blocks), whereas commercial zones had the fewest (12 blocks, 0.8%) (Figure 12a). Within each functional zone, the proportion of blocks in the higher-risk group varied considerably: public service zones had the highest proportion (61.3%), followed by industrial (50.7%), ecological and leisure (48.3%), residential (47.4%), and commercial zones (37.5%), the last being the only type where blocks in the lower-risk group predominated (62.5%) (Figure 12b). The range of these proportions was 23.8 percentage points. Residential zones accounted for 51.0% of all blocks in the higher-risk group and 56.7% of all blocks in the lower-risk group, a difference of 5.7 percentage points. Public service zones constituted 19.8% of the higher-risk group, which was 1.6 times their share of the lower-risk group (12.5%).

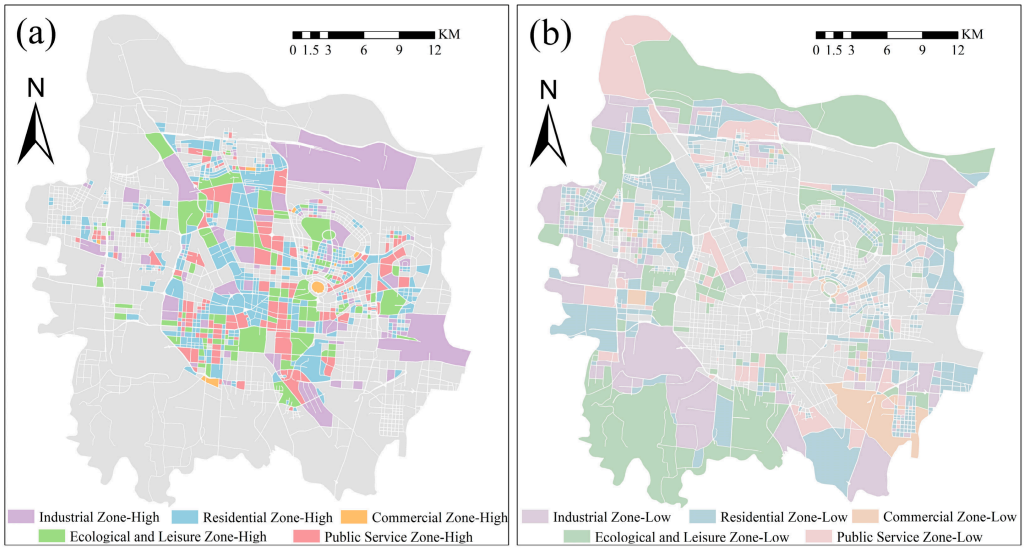


Figure 12. Spatial distribution of the two-dimensional function–risk classification based on the median split of the composite flood risk index. (**a**) Higher-risk group; (**b**) lower-risk group.

Table 8. Distribution of higher-risk and lower-risk groups (median split) by functional zone.

Functional Zone	Higher-Risk Group	Lower-Risk Group	Within-Zone Higher-Risk Proportion (%)
Residential	376	418	47.4
Industrial	106	103	50.7
Commercial	12	20	37.5
Ecological and Leisure	97	104	48.3
Public Service	146	92	61.3
Total	737	737	50.0

Figure 13 illustrates the spatial distribution of flood-prone road sections, and Table 9 summarizes their statistics by administrative district. The study area contained 38,512 valid road segments with a total length of 3702.0 km. Among them, 466.4 km were identified as flood-prone, representing 12.6% of the entire road network length. The length and proportion of flood-prone sections exhibited gradient differences across districts. Jinshui District recorded the largest flood-prone length (165.1 km), representing 13.9% of its total road length—1.3 percentage points above the domain-wide average. Guancheng District followed with 101.2 km (13.7%), Huiji District had 77.2 km (13.2%), Zhongyuan District had 101.8 km (12.5%), close to the domain average, and Erqi District recorded the smallest flood-prone length (21.0 km), accounting for only 5.7% of its total road length and falling

6.9 percentage points below the domain average. Its flood-prone length was merely 12.7% of that in Jinshui District. The ranking of flood-prone length and proportion was consistent across districts, with Jinshui recording the highest values, Guancheng and Zhongyuan showing comparable intermediate levels, followed by Huiji, and Erqi the lowest.

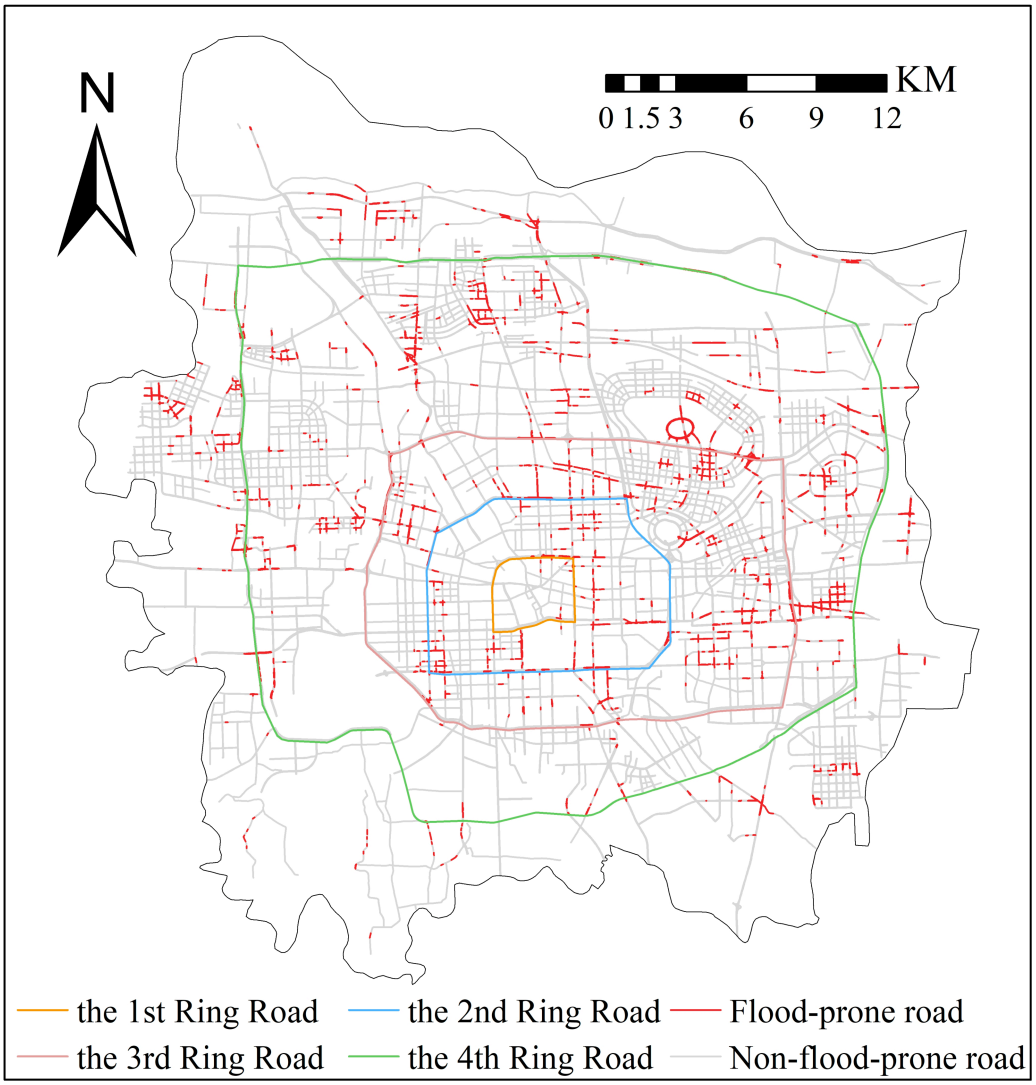


Figure 13. Spatial distribution of flood-prone road sections.

Table 9. Statistics of flood-prone road sections by administrative district.

Administrative District	Total Road Length (km)	Flood-Prone Length (km)	Proportion of Flood-Prone Length (%)
Jinshui	1191.9	165.1	13.9
Zhongyuan	816.8	101.8	12.5
Guancheng	738.4	101.2	13.7
Huiji	585.3	77.2	13.2
Erqi	369.6	21.0	5.7
Total	3702.0	466.4	12.6

Figure 14 shows the spatial overlay between the identified high-risk blocks and the observed waterlogging points from the "7·20" event. Based on 732 field-measured and 173 multi-source points, a 50 m spatial matching tolerance and a 500 m observation radius yielded 62 true positives, 70 deterministic false positives, 170 potential-risk blocks located

in observation blind zones, and 278 false negatives. Table 10 summarizes the precision, recall, and F1 scores of high-risk identification for each functional zone. The adjusted precision was 46.97%, recall was 18.24%, and the F1 score was 0.263. When potential-risk blocks were treated as false positives, the unadjusted precision decreased to 20.53%. At a 100 m tolerance, the adjusted precision increased to 59.09% and recall to 19.31%, with an F1 score of 0.291. Across matching tolerances of 30, 50, 100, and 150 m, adjusted precision ranged from 37.88% to 64.39%, while recall remained between 17.06% and 19.06%. At the 80th, 85th, and 90th percentile thresholds, recall at 50 m tolerance was 24.41%, 18.24%, and 11.47%, respectively. Functional-zone-specific validation metrics are reported in Table 10. The highest F1 score was observed in public service zones (0.330), followed by industrial (0.286) and commercial (0.286) zones. Residential zones had an F1 score of 0.238, and ecological and leisure zones had 0.220. Table 11 compares the validation performance of the proposed framework with baseline methods, including a physically based susceptibility model and a random permutation baseline. A physically based susceptibility model using only elevation, slope, and distance to water was evaluated under the same high-risk criteria and 50 m matching tolerance. This baseline yielded an adjusted precision of 51.04%, a recall of 14.41%, and an F1 score of 0.225. The random baseline was repeated 1000 times, yielding a mean F1 score of 0.249 (95% range: 0.237–0.261). The integrated index achieved an F1 of 0.263, falling above the 95% range of the random F1 scores and indicating a limited but non-random improvement.

Table 10. Functional-zone-specific validation performance of high-risk identification.

Functional Zone	Reference Samples (TP + FN)	TP	FP	FN	Precision	Recall	F1
Residential	134	24	44	110	0.353	0.179	0.238
Industrial	42	8	6	34	0.571	0.190	0.286
Commercial	12	2	0	10	1.000 *	0.167	0.286
Ecological and Leisure	71	10	10	61	0.500	0.141	0.220
Public Service	81	18	10	63	0.643	0.222	0.330

* The precision value for commercial zones was derived from two true positives and no false positives, reflecting the small number of validation samples in this category.

Table 11. Validation performance of the proposed integrated index and baseline methods.

Method	Adjusted Precision (%)	Recall (%)	F1
Physical susceptibility baseline	51.04	14.41	0.225
CRITIC-based integrated index	46.97	18.24	0.263
Entropy-weighted index	50.36	20.29	0.289
Random baseline (1000 permutations)	—	—	0.249

Figure 15 shows the relationship between the CRITIC and entropy-weighted risk indices. Figure 16 compares the spatial distributions of high-risk blocks identified by the CRITIC and entropy weighting methods under the same 85th percentile and physical constraint criteria. The Spearman rank correlation between the two indices was 0.57 ($p < 0.001$). The CRITIC method identified 302 high-risk blocks, whereas the entropy weighting method identified 310. The Jaccard overlap between the two sets was 0.41, and Cohen's Kappa was 0.47. Adjusted precision was 46.97% for the CRITIC-based index and 50.36% for the entropy-based index. Recall was 18.24% and 20.29%, respectively, and the corresponding F1 scores were 0.263 and 0.289. Public service zones consistently had the highest proportion of higher-risk blocks under both weighting schemes, while the relative ordering of commercial, industrial, and ecological zones varied.

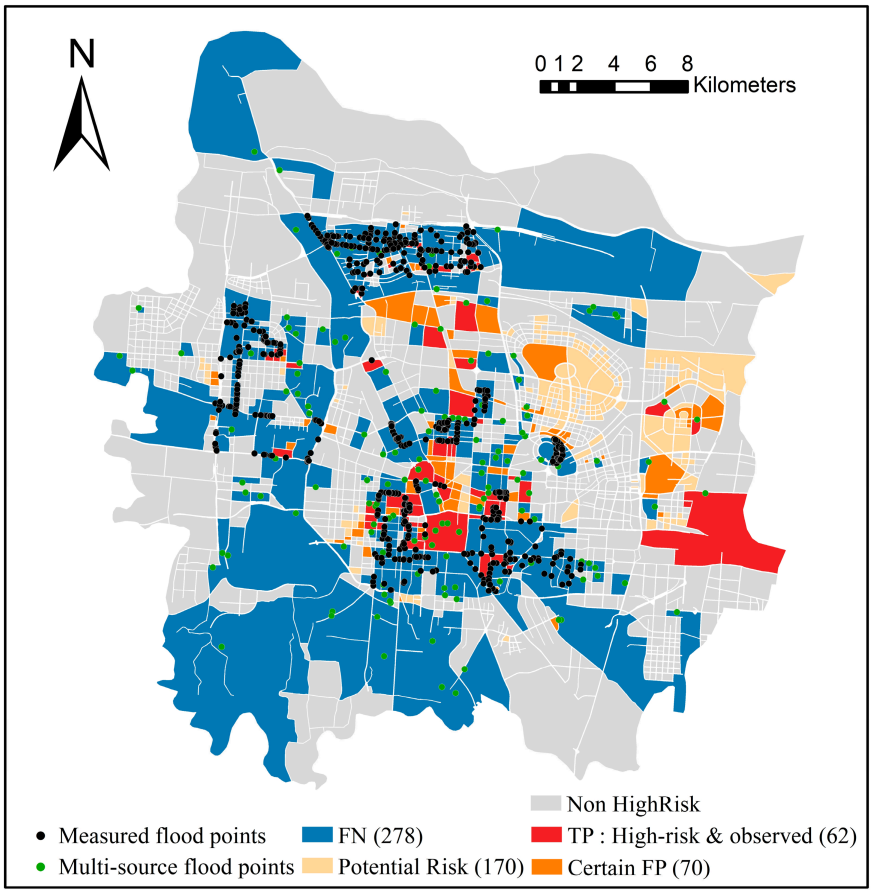


Figure 14. Spatial overlay of model-identified high-risk blocks and observed waterlogging points from the "7·20" event.

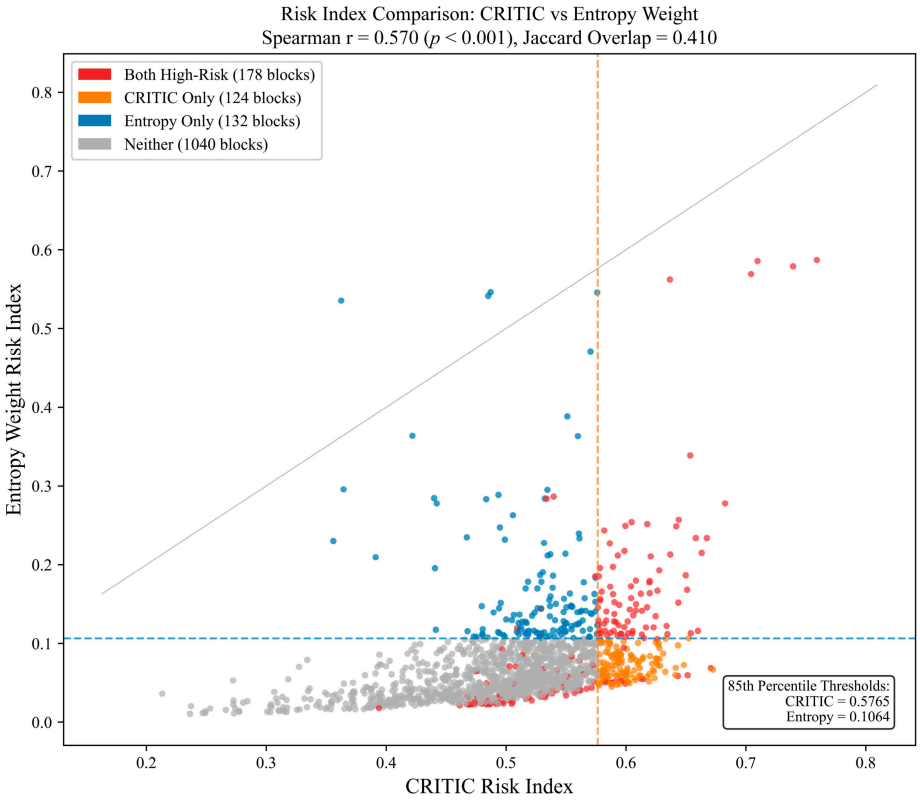


Figure 15. Scatter plot of the CRITIC and entropy-weighted risk indices.

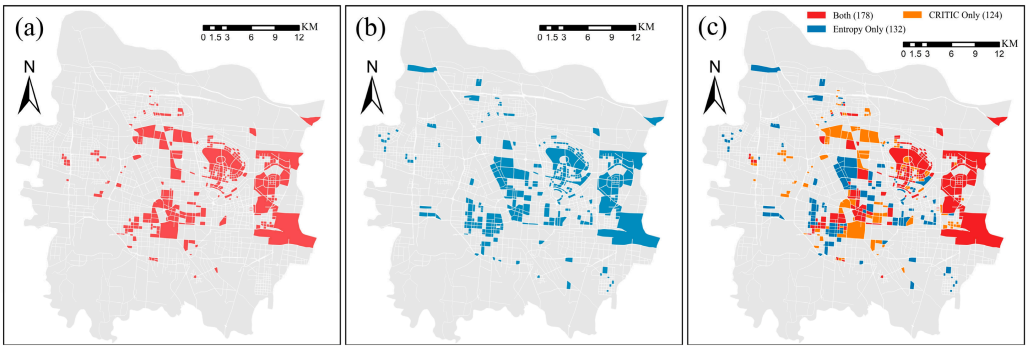


Figure 16. Spatial distributions of high-risk blocks identified by the CRITIC and entropy weighting methods under the same 85th percentile and physical constraint criteria. (**a**) CRITIC method (302 high-risk blocks); (**b**) entropy weighting method (310 high-risk blocks); (**c**) overlap comparison.

Figure 17 presents the land use transition matrix from 1985 to 2023. Built-up land increased from 245.88 km^2 in 1985 to 317.29 km^2 in 1995 and 711.97 km^2 in 2023, representing a net increase of 466.09 km^2. The annual growth rate was 7.14 km^2 yr^{-1} during 1985–1995 and 14.10 km^2 yr^{-1} during 1995–2023, indicating a marked acceleration after the mid-1990s. Built-up land added after 1995 accounted for 84.7% of the total increase. Of the total built-up land added during 1985–2023, 99.0% was converted from cultivated land, while ecological land decreased from 35.34 km^2 to 30.73 km^2. Among built-up land added after 1995, the proportions of high-risk residential and commercial blocks were 10.2% and 4.6%, respectively, compared with domain-wide baseline levels of 20.4% and 21.9%. High-risk commercial and public service blocks remained concentrated in early-built core areas.

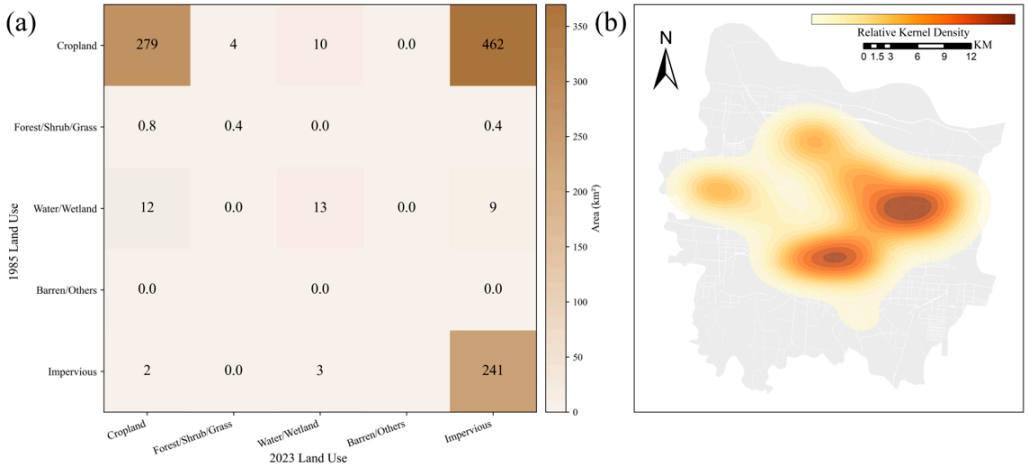


Figure 17. Historic urban land use transfer and flood risk spatial lock-in. (**a**) Land use transition matrix from 1985 to 2023, with cell values representing transfer areas (km^2); (**b**) kernel density heatmap of high-risk commercial and public service blocks, showing their spatial agglomeration in early-built core areas.

3.3. Characteristics of the Multi-Path Flood Risk Connectivity Network

Figure 18 shows the three types of flood risk connectivity networks together with their structural statistics. Before deduplication, the three layers contained 4684 edges in total, including 1355 hydrological connectivity edges, 103 functional coupling edges, and 3226 road network edges. After the three edge sets were combined and deduplicated, the coupled network contained 3621 unique edges. Because individual edges may correspond to multiple connection mechanisms, edge coverage among different layers is not mutually exclusive. Road-related connections were involved in 3226 unique edges (89.1%), hydrological connections in 1355 edges (37.4%), and functional coupling connections in 103 edges

(2.8%). The hydrological connectivity network contained 975 nodes, with a mean out-degree and in-degree both equal to 1.39. 12 blocks (1.2%) had an out-degree greater than 5, and 25 blocks (2.6%) had an in-degree greater than 5. Among the functional coupling edges, residential–residential connections were the most frequent (33 edges), followed by public service–public service connections (13 edges) and public service–residential connections (11 edges). These three types together accounted for 55.3% of all functional coupling edges. The mean functional coupling out-degree was 0.11 for high-risk blocks and 0.03 for low-risk blocks. In the road network cascading failure simulation, the 10 highest-risk blocks were set as initial failure nodes. Under the specified parameters, cascading failure propagated to 1203 nodes. The sensitivity of this result to the tolerance parameter α is summarized in Table 12.

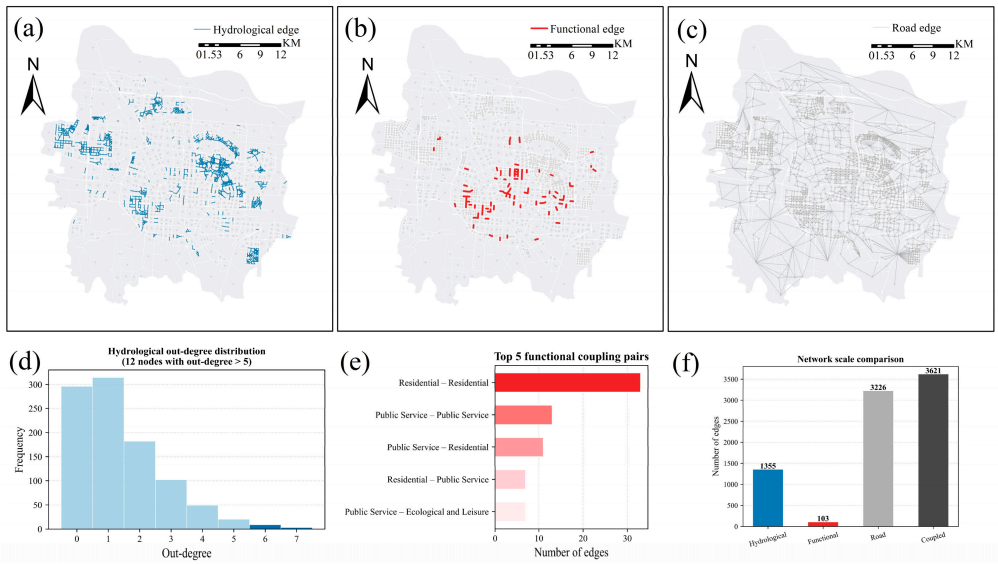


Figure 18. Three types of flood risk connectivity networks and associated structural statistics. (**a**) Hydrological connectivity network; (**b**) functional coupling network; (**c**) road network topology; (**d**) out-degree distribution of the hydrological connectivity network; (**e**) top five functional coupling pairs by functional zone type; (**f**) comparison of edge counts among the hydrological, functional, road, and coupled networks.

Table 12. Sensitivity of cascading failure results to the tolerance parameter α.

α	Final Failed Nodes
0.1	1393
0.3	1203
0.5	879
0.7	10

Figure 19 presents the coupled flood risk connectivity network after superimposing the three types of edges. Following deduplication, 3621 unique coupled edges were obtained, involving 1466 participating nodes. The largest connected component encompassed 1455 blocks, covering 99.2% of all blocks. The network density was 0.0034, the mean degree reached 4.94, the average clustering coefficient was 0.391, and the average path length was 15.01. The global efficiency of the coupled network was 0.089, providing an additional measure of the overall structural connectivity and reachability of the network. Fourteen hub nodes (1.0%) possessed a degree greater than 10, distributed mainly in the urban core and near major river confluences. The top 5% of nodes ranked by betweenness centrality carried 49.3% of the shortest path flow. The global Moran's I was 0.703 ($p < 0.001$), indicat-

ing significant spatial autocorrelation under the coupled connectivity structure rather than dynamic propagation along network links.

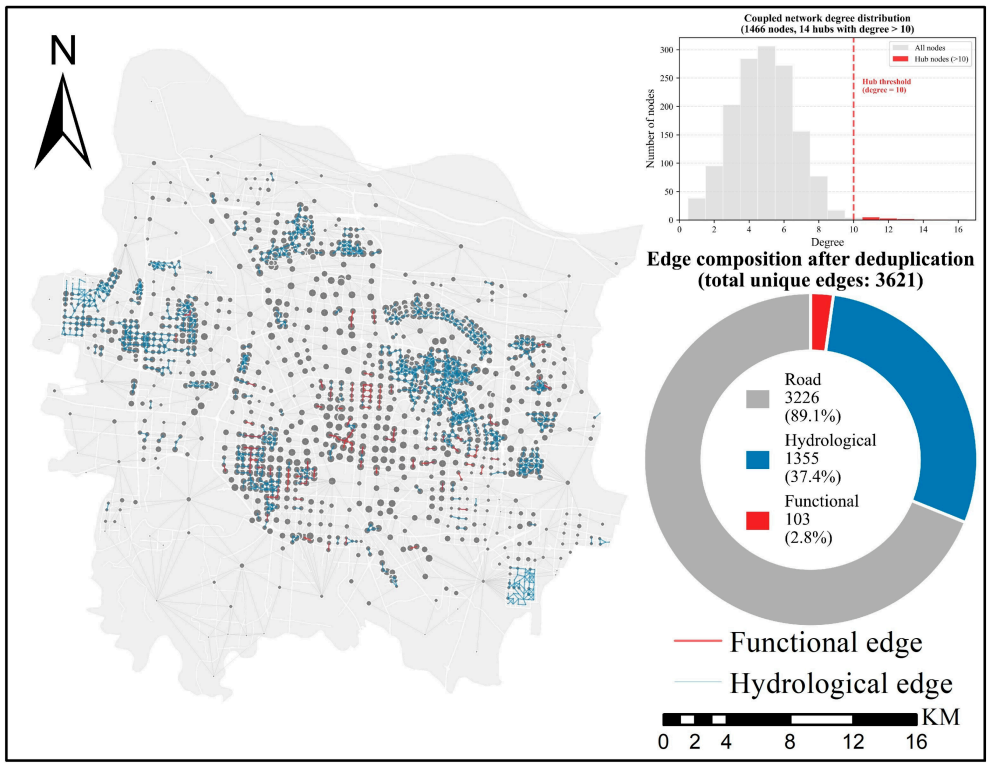


Figure 19. Coupled flood risk connectivity network. Node size represents the block composite flood risk index.

Figure 20 displays the spatial distribution of risk connectivity intensity and range in the coupled network. The domain-wide mean connectivity intensity, measured by weighted out-degree, was 4.51, and the median was 4.29. High-intensity areas were concentrated in the urban core and at major river confluences. The mean connectivity range, defined as the number of reachable nodes within three steps, reached 31.6, with a median of 29.0. Both indicators displayed a pronounced core–periphery gradient. The Pearson correlation coefficient between intensity and range was 0.443 ($p < 0.001$).

Figure 21 compares connectivity intensity and range across the five functional zones. Table 13 summarizes the key pairwise comparisons and highlights the significant differences in connectivity intensity. Commercial zones exhibited the highest mean connectivity intensity (5.72), followed by ecological and leisure (4.84), public service (4.75), industrial (4.36), and residential (4.35) zones. Commercial zones also recorded the largest mean connectivity range (35.0), followed by industrial (33.1) and ecological and leisure (32.4) zones. The Kruskal–Wallis test indicated a significant overall difference in connectivity intensity among functional zones ($H = 13.98$, $p = 0.007$), whereas no significant difference was detected for connectivity range ($H = 5.03$, $p = 0.284$). Bonferroni-corrected Mann–Whitney U pairwise comparisons identified significant differences in connectivity intensity only between public service and residential zones (adjusted $p = 0.016$) and between public service and industrial zones (adjusted $p = 0.029$). All other pairwise comparisons, including those involving commercial zones, were not significant. The corresponding rank-biserial effect sizes ranged from −0.19 to 0.19, with all confidence intervals including zero. Regression analyses further showed that a one-standard-deviation increase in degree centrality was associated with an increase of 2.07 in connectivity intensity ($p < 0.001$) and 6.28 in

connectivity range ($p < 0.001$). Functional zone type did not show significant effects after accounting for network position.

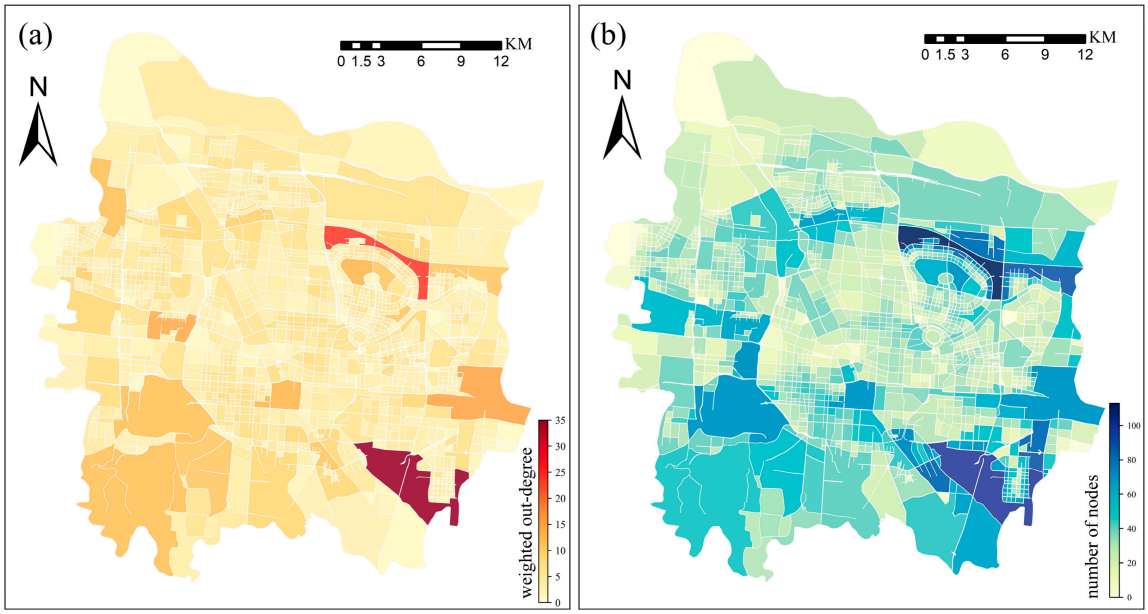


Figure 20. Spatial distribution of risk connectivity intensity and range in the coupled network. (**a**) connectivity intensity (weighted out-degree); (**b**) connectivity range (number of nodes).

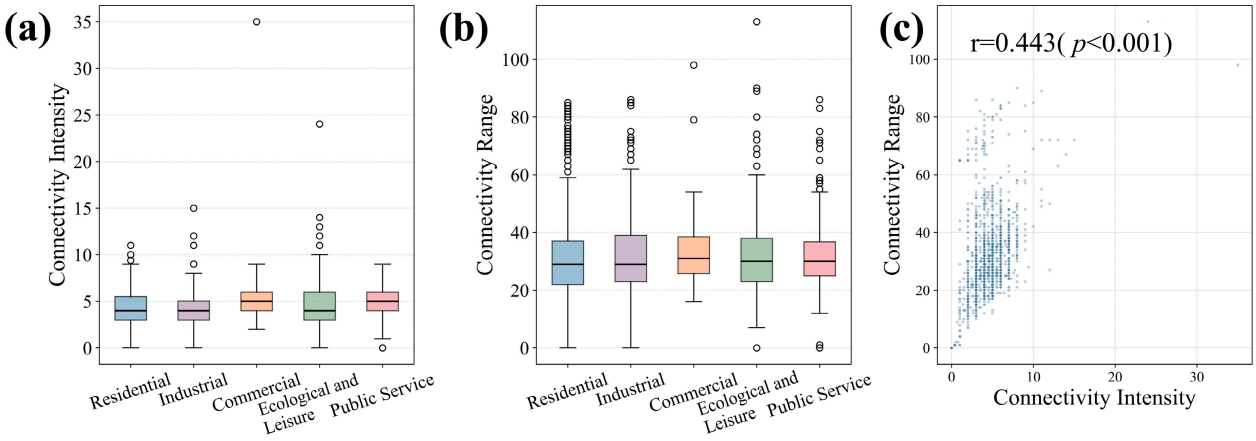


Figure 21. Risk connectivity intensity and range by functional zone. (**a**) connectivity intensity; (**b**) connectivity range; (**c**) scatter plot of intensity versus range. White dots in (**a**,**b**) represent outliers beyond 1.5 times the interquartile range (IQR).

Table 14 summarizes the structural sensitivity of network indicators to hydrological and functional coupling parameters. Hydrological edges were most sensitive to centroid distance, ranging from 264 at 300 m to 2880 at 700 m, a 10.9-fold difference. Relative to the baseline of 1355 edges, the 300 m setting retained only 19.5%, while the 700 m setting produced 2.13 times as many. The elevation difference threshold had a smaller effect, with hydrological edges ranging from 859 at 2.0 m to 1691 at 0.5 m, corresponding to 63.4% and 124.8% of the baseline, respectively. Functional coupling edges were insensitive to the gravity exponent, remaining at 103 under all tested values, but declined from 144 to 62 when the top quantile cutoff changed from 3% to 10%, a reduction of 57.0%. Across all parameter settings, functional-zone-level rankings remained largely stable. Thus, the size of the hydrological connectivity layer was threshold-dependent. Functional-zone-level rankings, however, remained largely stable across the tested parameter settings.

Table 13. Key pairwise comparisons of connectivity intensity among functional zones using Bonferroni-corrected Mann–Whitney U tests.

Comparison	Adjusted p	Rank-Biserial r	95% CI
Public Service vs. Residential	0.016	−0.133	−0.217 to −0.050
Public Service vs. Industrial	0.029	−0.162	−0.269 to −0.054
Commercial vs. Residential	1.000	−0.161	−0.365 to 0.043
Commercial vs. Industrial	0.789	−0.191	−0.407 to 0.024
Commercial vs. Ecological and Leisure	1.000	0.093	−0.122 to 0.309
Commercial vs. Public Service	1.000	0.024	−0.190 to 0.237

Table 14. Sensitivity of structural network indicators to hydrological and functional coupling parameters.

Parameter Setting	Hydrological Edges	Functional Edges	Degree Top 20 Overlap	Betweenness Top 20 Overlap
Baseline (500 m, 1 m, exp 2, 5%)	1355	103	1.00	1.00
Hydro distance 300 m	264	103	0.95	1.00
Hydro distance 700 m	2880	103	0.20	0.85
Elevation difference 0.5 m	1691	103	0.90	0.95
Elevation difference 2.0 m	859	103	1.00	0.95
Gravity exponent 1.5	1355	103	1.00	0.90
Gravity exponent 2.5	1355	103	1.00	0.95
Top cutoff 3%	1355	144	1.00	0.95
Top cutoff 10%	1355	62	1.00	0.95

3.4. Coupling Relationship Between Functional Mixture and Flood Risk

Figure 22 presents the standardized mixture degree (H_norm), functional deviation degree (D), and LISA cluster type composition across the five functional zones. The domain-wide mean of H_norm was 0.629, with a median of 0.773 (Figure 22a). A total of 69.0% of blocks were classified as high-mixture (H_norm > 0.67). Ecological and leisure zones recorded the highest mean H_norm (0.751), and residential zones had the lowest (0.566). The proportion of high-mixture blocks was highest in ecological and leisure zones (86.1%) and lowest in residential zones (61.5%). The mean D value was highest for residential zones (0.458) and lowest for ecological and leisure zones (0.231) (Figure 22b). Among LISA cluster types, 153 blocks (10.4%) were identified as High-High clusters and 185 blocks (12.6%) as Low-Low clusters (Figure 22c). Residential zones accounted for 41.2% of High-High clusters and 75.1% of Low-Low clusters, appearing in both extreme clustering types with a ratio of 0.45 between the two. The proportions of LISA-significant blocks for industrial and commercial functional zones were 30.6% and 25.0%, respectively, which were lower than those of residential (33.4%) and ecological and leisure zones (26.4%).

Figure 23 illustrates the spatial distribution and frequency distributions of GWR local coefficients for the physical flood susceptibility index. The median local coefficient of H_norm was 0.0015, with an interquartile range of 0.091 (Q1 = −0.043, Q3 = 0.049) (Figure 23a). Positive coefficients occurred in 51.3% of blocks, and 75.8% of coefficients were within ±0.1. The median local coefficient of D was 0.0111, with an interquartile range of 0.126 (Q1 = −0.051, Q3 = 0.074) (Figure 23b). Positive coefficients occurred in 56.0% of blocks, and 70.8% of coefficients were within ±0.1. The GWR model produced local coefficient estimates that varied

across space. (Figure 23c,d). With an adaptive bisquare kernel, the optimized bandwidth was 47 nearest neighbors. The GWR model achieved an R^2 of 0.792 and an AICc of −4045.522, compared with an R^2 of 0.011 and an AICc of −2315.494 for the OLS baseline. Local R^2 values ranged from −0.491 to 0.876, with a mean of 0.466. The Monte Carlo test did not provide evidence of statistically significant global spatial non-stationarity (p = 0.960). Residual spatial autocorrelation remained significant (Moran's I = 0.310, p = 0.001).

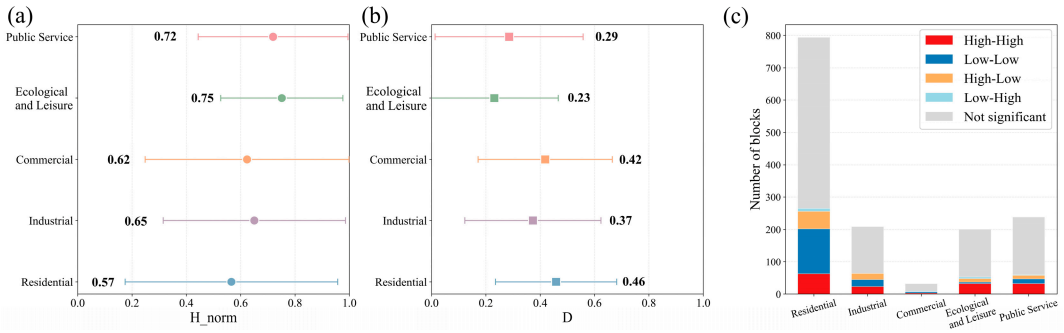


Figure 22. Standardized mixture degree and functional deviation degree across functional zones. (**a**) Mean H_norm; (**b**) mean D; (**c**) LISA cluster type composition.

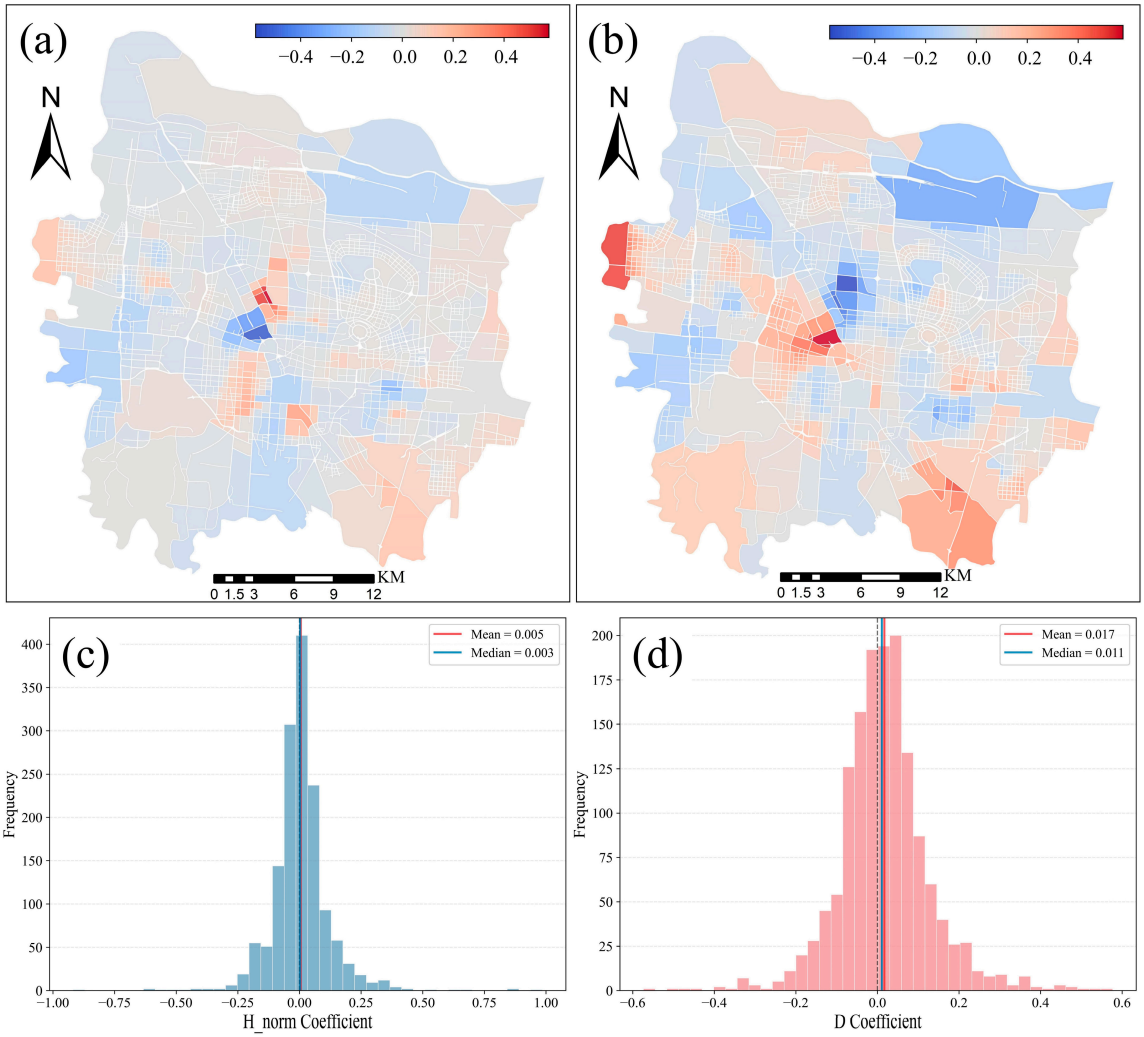


Figure 23. Spatial distribution and frequency distributions of GWR local coefficients for the physical flood susceptibility index. (**a**) H_norm local coefficient; (**b**) D coefficient; (**c**) frequency distribution of H_norm coefficient; (**d**) frequency distribution of D coefficient.

Figure 24 presents the LISA spatial clustering map of H_norm, and Figure 25 shows the percentage composition of LISA cluster types within each functional zone as a heatmap. Among the 1474 blocks, 450 (30.5%) were identified as LISA-significant. High-High clusters accounted for 153 blocks (10.4%) and Low-Low clusters for 185 blocks (12.6%), representing the two dominant types. High-Low clusters comprised 96 blocks (6.5%), and Low-High clusters were the least common, with only 16 blocks (1.1%). The remaining 1024 blocks (69.5%) showed no significant spatial association. Across functional zones, the High-High proportion was highest in ecological and leisure zones (15.92%), followed by public service zones (13.45%), industrial zones (11.00%), and residential zones (7.93%). The Low-Low proportion was highest in residential zones (17.51%), exceeding their High-High proportion (7.93%), while ecological and leisure zones had the lowest Low-Low proportion (2.49%) and commercial zones recorded 12.50%. High-Low clusters were most prevalent in industrial zones (9.09%) and residential zones (6.80%). Low-High clusters were generally rare and absent in industrial and commercial zones. The proportion of non-significant blocks was highest in commercial zones (75.00%) and lowest in residential zones (66.62%).

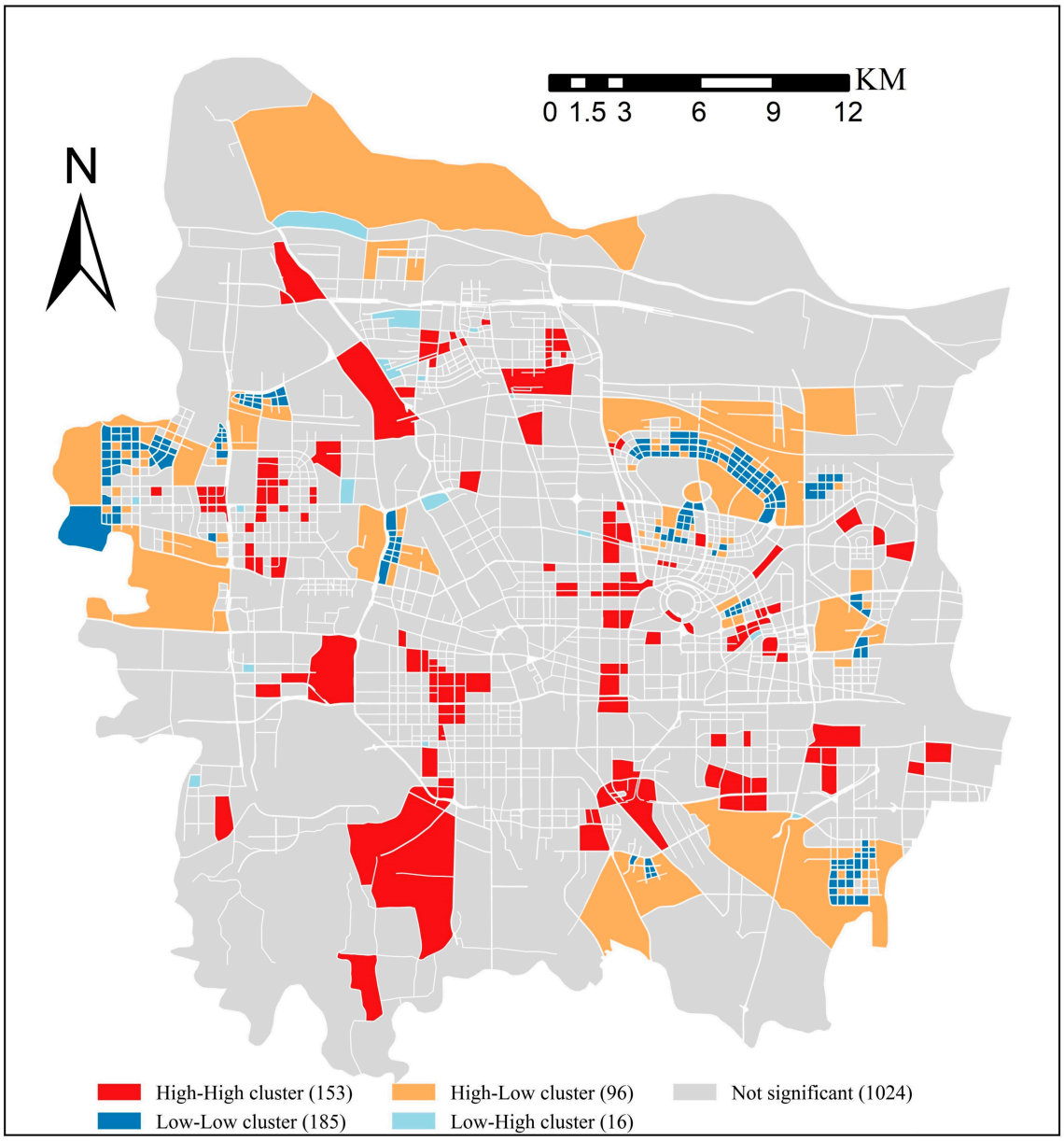


Figure 24. LISA spatial clustering map of H_norm.

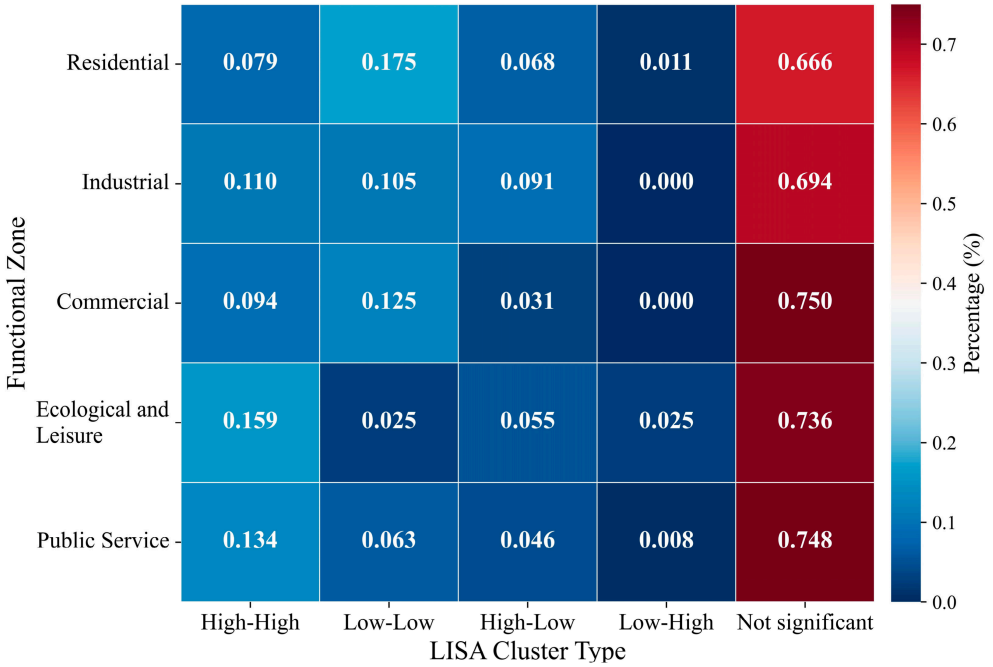


Figure 25. Heatmap of LISA cluster type percentages within each functional zone.

4. Discussion

4.1. Urban Expansion and the Spatial Differentiation of Flood Risk

Urban expansion has been associated with increased impervious surfaces and reduced natural regulation capacity, which may alter the hydrological conditions underlying the spatial distribution of flood risk. Impervious surface expansion is an important factor contributing to urban waterlogging risk, as urban development alters land surface characteristics and affects watershed hydrological responses [60]. The long-term land use evolution analyzed in this study provides spatial evidence of urban development trajectories [19]. As reported in Section 3.2, built-up land increased substantially from 1985 to 2023, mainly through the conversion of cultivated land, whereas ecological land decreased by approximately 13%. This trend indicates a substantial decline in natural hydrological regulation capacity and increasing dependence on artificial drainage systems.

The spatial distribution of flood risk reflects a development-stage effect, with high-risk areas concentrated in early-built urban regions rather than showing a consistent increase in newly developed areas. Urban expansion has been reported to increase flood exposure in newly developed areas [61]; yet newly developed blocks in this study show a different pattern. As reported in Section 3.2, high-risk commercial and public service areas remained concentrated in early-built core areas, while the proportions of high-risk residential and commercial blocks in post-1995 built-up land were lower than the domain-wide baselines (Figure 17b). This pattern can be interpreted as a form of spatial lock-in associated with historical development trajectories. Existing drainage infrastructure may constrain flood adaptation because of historical design conditions and retrofit limitations [62]. Core areas exhibited an impervious surface ratio above 80%, a composite risk index of 0.614 (20.8% above the domain average), and a POI density of 186.4 POIs·ha^{-1}, which may reflect a persistent high-risk pattern associated with intensive historical development. This interpretation is consistent with the broader concept of infrastructure path dependency [63]. The difficulty of large-scale renewal perpetuates high-risk conditions in old urban cores, whereas newly developed districts may benefit from improved drainage and greater blue-green space availability, resulting in comparatively lower flood risk.

Residential areas remain highly exposed to flood risk because both old and newly developed settlements face different forms of pressure. This result aligns with the broader observation that residential blocks dominate the high-risk categories under both the median-based and the 85th-percentile definitions. Previous studies have emphasized the vulnerability of older residential communities during rapid urbanization [64]. This study extends that understanding by revealing that residential risk is not limited to aging settlements. Older communities are likely constrained by historically limited drainage capacity, while newly developed residential areas may experience comparable runoff pressure due to high-density construction and limited ecological buffering space. This comparable risk pattern across different development stages highlights the temporal characteristics of residential expansion during rapid urbanization. The contrast between the low proportion of residential POIs and the large share of residential functional blocks also suggests that POI-based functional identification should not be interpreted directly from category proportions. Urban functions differ considerably in POI density and spatial coverage.

Based on the differentiated risk characteristics of the five functional zones, targeted flood prevention strategies can be considered. Residential zones require drainage improvement and emergency access enhancement, especially in densely built older communities. Industrial zones should strengthen protection of critical facilities. Commercial zones require enhanced protection of underground spaces and highly connected nodes. Ecological and leisure zones can improve stormwater regulation through blue-green infrastructure [65]. Public service zones should prioritize emergency response capacity and essential facility protection. These differentiated strategies provide practical implications for functional-zone-based flood resilience planning.

4.2. Multi-Path Connectivity of Flood Risk and Spatial Effects of Functional Mixing

Urban functional zones are often conceptualized as discrete spatial units, but in practice many blocks contain mixed POI categories that blur the boundaries between land-use types [10,50]. Direct clustering methods may assign heterogeneous blocks to a single cluster according to feature similarity, which could obscure mixed functional characteristics and introduce uncertainty into subsequent analyses. To address this issue, a density-based refinement step was introduced to identify a small proportion of low-density blocks whose functional characteristics were difficult to represent by dominant cluster patterns [27]. This refinement produced comparable clustering quality to direct K-Means while slightly improving cluster purity and enabling the identification of mixed-function blocks (Table 4). For this study, maintaining functional coherence of spatial units is considered more relevant than maximizing cluster compactness alone, because the subsequent flood risk and network analyses rely on unit-level functional characteristics rather than cluster geometry. Therefore, the two-stage clustering strategy provides a practical balance between clustering performance and functional interpretability. The final functional labels were assigned at the block level, not the cluster level. The clustering results therefore serve as a spatial consistency check.

The ablation results in Table 5 provide further insight into the two feature branches. The spatial-only configuration achieved a silhouette coefficient of 0.554. This was close to the full semantic–spatial representation (0.564). The semantic-only configuration achieved only 0.473. These results indicate that clustering is primarily driven by spatial proximity. This is consistent with the geographic continuity of urban functional patterns. At the same time, the semantic branch provides complementary information on POI composition. It allows the clustering to account for functional differences beyond spatial proximity alone. However, the increase from 0.554 to 0.564 is small. The semantic contribution to global cluster separation is therefore limited but complementary. Thus, the combined

representation should be interpreted as spatially dominant but semantically enriched, rather than purely geographic. The manual validation showed that the block-level functional labels were generally consistent with independent reference information (overall agreement = 84.0%, Cohen's kappa = 0.80). The misclassified blocks were primarily those with mixed residential and commercial functions. Some ecological and leisure blocks also contained large industrial or public service land uses. These cases reflect the challenge of assigning a single dominant category to complex urban blocks. Because the study area is a highly urbanized region, the five-category classification is inevitably coarse; future research could refine the categories or adopt multi-label indicators to better represent mixed blocks. The permutation analysis further suggests that the extracted semantic features may be influenced by the predefined POI-category ordering. However, the overall clustering structure remained identifiable across the tested orderings. Accordingly, the LSTM is used here as a nonlinear encoder under a predefined input arrangement, rather than as a model of an intrinsic sequential relationship among POI categories. The MLP comparison further supports this interpretation, as the order-independent alternative did not substantially improve clustering performance. Future work could explore order-invariant encoders or more systematic permutation analyses.

Flood risk connectivity in urban areas is shaped by the interaction of multiple pathways rather than by isolated spatial exposure. Previous studies have mainly examined individual connectivity mechanisms, including hydrological connectivity [66], infrastructure networks [67], and socioeconomic connectivity reflected in population mobility and exposure [13]. This study integrates these pathways into a coupled network framework to examine their combined effects. In the functional coupling component, a gravity model with a distance decay exponent of 2 was used. However, reported exponents in interaction models vary considerably across studies, which motivated the sensitivity analysis of this parameter [68]. Road-related connections were involved in 89.1% of the unique network edges, representing the most prevalent connection type in the coupled network (Figure 19). The proportion of flood-prone road segments in the core circle (13.9%) is significantly higher than that in peripheral areas (5.7%), showing a clear spatial gradient (Figure 13). Dense road networks may facilitate rapid redistribution of flood impacts during extreme rainfall events, and their connectivity efficiency is closely related to network density [69]. The sensitivity analysis showed that network properties were not uniformly robust. Functional-zone rankings remained relatively stable, but hydrological edge counts and degree centrality were highly sensitive to the distance threshold. Increasing the threshold from 300 m to 700 m raised hydrological edges from 264 to 2880. Degree top-20 overlap dropped to 0.20 at 700 m. This stability should not be generalized to the entire network. The 500 m threshold was selected as an intermediate baseline for block-scale connectivity. It balances local detail against network densification. This choice is a modeling convention rather than a universal value. Future work could use flow-direction models to reduce this uncertainty.

Commercial functional zones comprised 32 blocks (2.2% of all blocks), and their within-zone proportion of high-risk blocks was the highest among the five functional types under the adopted definition (21.9%, Table 7). Their descriptive connectivity intensity and range were 5.72 and 35.0, whereas the residential-zone values were 4.35 and 30.6 (Figure 21). However, these connectivity differences were not statistically significant after Bonferroni correction. Commercial nodes also exhibited relatively high betweenness centrality, indicating their relatively central position in the network. Regression analysis further showed that connectivity intensity and range were more strongly associated with network degree centrality than with functional type. Thus, the relatively high descriptive connectivity observed in commercial zones should not be attributed to commercial function alone. Given their high POI density (186.4 POIs·ha^{-1}), commercial areas may warrant

particular planning attention where concentrated urban activity coincides with central network positions.

At the baseline setting of $\alpha = 0.3$, the cascading failure simulation resulted in 1203 failed nodes. Compared with the earlier specification, the revised model explicitly incorporates load redistribution and node capacity constraints, allowing simulated cascading effects to extend beyond the initial failure nodes. The simulated outcome was highly sensitive to the assumed tolerance parameter, with the number of failed nodes ranging from 1393 at $\alpha = 0.1$ to 10 at $\alpha = 0.7$. The value $\alpha = 0.3$ was selected as an illustrative intermediate setting rather than an empirically calibrated parameter. The cascading simulation is therefore treated as a stress-testing scenario to examine the potential sensitivity of network failure propagation to assumed capacity conditions, rather than to predict actual flood-induced failures. Nevertheless, the possibility of failure propagation under more severe disturbances cannot be excluded [70]. The spatial overlap between highly connected road nodes and high-risk functional areas further indicates locations that may warrant attention under multi-node failure scenarios.

Functional mixture is associated with flood risk in spatially heterogeneous ways rather than through a uniform relationship. Previous studies have reported contrasting effects of functional mixing, with some suggesting that mixed land use can improve resilience [64], while others emphasize the additional pressure caused by intensified development [65]. After removing POI-derived indicators from the dependent variable, the association between functional mixture and flood hazard weakened substantially and became spatially inconsistent. This indicates that functional structure alone does not appear to be a dominant determinant of physical flood hazard formation. Instead, these associations may reflect differences in exposure, functional dependence, and the position of functional zones within the flood risk connectivity network. The GWR results produced spatially varying local coefficient estimates, but the Monte Carlo test did not confirm statistically significant global non-stationarity. The observed spatial variation in local coefficients should therefore be interpreted descriptively rather than as evidence of systematic non-stationary processes. The significant residual spatial autocorrelation (Moran's I = 0.310, $p = 0.001$) further suggests that some spatially structured factors, possibly including drainage infrastructure or micro-topographic features, remain unaccounted for in the current model.

The CRITIC method was retained for constructing the composite flood risk index because it considers both indicator contrast intensity and inter-indicator conflict. This is particularly relevant in Zhengzhou, where low-lying areas tend to be close to water bodies and have gentle slopes, so several physical indicators are correlated. The entropy weighting method, by contrast, assigns weights only according to indicator dispersion and does not account for such correlations. In the present dataset, underground space ratio received an entropy weight of 0.521, whereas distance to water bodies received only 0.011. This produces an index dominated by a single exposure-related variable and largely ignores the topographic and hydrological factors that are central to flood generation in Zhengzhou. Although the entropy-weighted index achieved slightly better validation scores than the CRITIC-based index, the difference was not substantial. The CRITIC method was therefore preferred because it provides a more balanced representation of the physical conditions relevant to flood susceptibility. Nevertheless, the exact ordering of some functional zones and the selection of initial failure nodes were sensitive to the weighting method, indicating that these specific rankings should be interpreted with caution.

4.3. Limitations and Future Perspectives

The analytical framework adopted in this study does not fully capture the dynamic processes of rainfall, runoff, and drainage system interactions. Flood risk is evaluated

using a static composite index based on terrain, impervious surfaces, and functional characteristics. This approach effectively reveals spatial differences in vulnerability but does not explicitly simulate rainfall-runoff processes or underground drainage dynamics. Compared with hydrodynamic models used in previous studies [71], the current framework has limited capability in representing temporal flood evolution. The validation results also show a relatively high false-negative rate, which can be attributed to several factors. First, the static risk index does not incorporate rainfall intensity, drainage network capacity, or real-time flood dynamics, and therefore cannot fully reproduce event-specific waterlogging locations. Second, the available waterlogging observations are predominantly roadside records, while systematic records within urban parcels are limited, which may overestimate false negatives. Third, several predictors were temporally inconsistent with the 2021 validation event. POI records, the 2023 land cover product, and building footprints reflect more recent conditions, and underground space ratio and functional agglomeration degree may have changed during this interval. This mismatch may contribute to false negatives, although its independent effect cannot be quantified with the available data. Consequently, the identified high-risk areas should be interpreted as potential flood-sensitive zones rather than as deterministic predictions of flood occurrence. The physically based baseline yielded an F1 score of 0.225, compared with 0.263 for the CRITIC-based integrated index, indicating a modest improvement when functional and exposure-related information was incorporated. The random permutation baseline was repeated 1000 times, yielding a mean F1 score of 0.249 (95% range: 0.237–0.261). The integrated index achieved an F1 of 0.263, exceeding the 97.5th percentile of the empirical random distribution. Moreover, the residential and ecological and leisure zones yielded F1 scores of 0.238 and 0.220, respectively, both below the random mean. These results suggest that the integrated index provides additional information beyond the physical baseline, but its advantage for matching event-specific waterlogging locations remains limited. The relatively low recall is consistent with the static nature of the risk indicators, the spatial incompleteness of the observed waterlogging records, and the temporal mismatch between some predictor variables and the 2021 validation event. Functional-zone-specific validation further showed that identification performance varied across functional types, with relatively lower recall in several zones. In addition, the use of urban blocks as assessment units reduces spatial noise and facilitates functional comparison.

However, this aggregation may overlook fine-scale factors such as local terrain variation and the precise locations of underground space entrances [72]. A supplementary sensitivity check was performed by aggregating high-risk blocks into regular grids of 500 m, 1 km, and 2 km. The proportion of grid cells containing high-risk blocks ranged from 2.9% to 16.7%, compared with 14.99% under the traffic-block unit. This variation indicates that the spatial representation of high-risk areas is sensitive to the choice of analytical unit. Traffic blocks were retained because they are widely used in urban planning and flood management, and because they are necessary for functional zone identification and road-network analysis. Multi-scale validation should be considered in future work. Future research could integrate high-resolution two-dimensional hydrodynamic models to examine how flood risk networks develop under different rainfall scenarios [73]. Incorporating LiDAR-derived elevation models and Building Information Models could extend the current two-dimensional spatial analysis to three-dimensional urban spaces.

Another limitation is that the indicator system did not include social vulnerability indicators such as age structure or income because consistent block-level data were not available, which particularly limits the interpretation of residential high-risk areas [74]. The choice of weighting scheme represents another source of uncertainty in composite flood risk assessment. In this study, the CRITIC method was adopted because it simultaneously

accounts for indicator variability and inter-indicator correlations, thereby providing a more balanced representation of hazard, exposure, and functional dimensions when some physical indicators are correlated. However, comparison with the entropy weighting method demonstrated that the weighting scheme can influence the relative contribution of individual indicators and the resulting risk classification. The entropy-based approach assigned greater weights to highly variable indicators, such as underground space ratio, whereas CRITIC provided a more distributed contribution among the three dimensions. Despite these differences, the two weighting schemes showed comparable performance in external validation, suggesting that the general spatial patterns of flood-sensitive areas were not solely determined by the weighting choice. Nevertheless, specific interpretations, such as the relative ranking of functional zones and the selection of initial failure nodes, remained sensitive to weighting uncertainty and should therefore be interpreted with caution.

The representation of risk connectivity mechanisms requires further development to incorporate dynamic human and traffic responses. The current cascading failure analysis is based on static network topology and does not capture real-time changes in traffic flows, evacuation behavior, or adaptive responses during flood events. Compared with studies incorporating mobile phone signaling data for mobility simulation [13], this framework has limitations in representing behavioral dynamics. In addition to mobility dynamics, the present framework primarily addresses physical-spatial flood risk and has limited representation of public risk perception and behavioural responses during flood events. Social-media-sourced data could provide temporally dynamic information on public risk perception, emergency needs, and public response information, offering a complementary perspective to mobility data. Integrating such information may support a more human–land coupled flood-risk assessment that accounts for both physical exposure and societal response [75]. Future work could integrate mobility data and agent-based models to investigate how disruptions at critical nodes cascade through urban activities and lifeline infrastructure. This study also does not simulate how flood risk networks may evolve under future urban expansion scenarios. Coupling the framework with cellular automata or other spatial simulation models could enable multi-scenario assessments of future risk dynamics [61]. The GWR analysis was also designed to explore spatially varying relationships rather than to completely eliminate residual spatial autocorrelation. The residual Moran's I was 0.310 (p = 0.001), and the Monte Carlo test did not indicate statistically significant global non-stationarity. The observed spatial variation in local coefficients should therefore be interpreted descriptively rather than as evidence of statistically significant global non-stationarity. The significant residual spatial autocorrelation may indicate that additional spatially structured factors, such as drainage infrastructure or micro-topographic features, are not fully captured by the current explanatory variables. Future work could incorporate more detailed drainage network data or spatial econometric approaches to further reduce residual dependence.

The transferability of the proposed framework depends on the terrain, POI characteristics, and drainage infrastructure of the target city. In plain cities such as Zhengzhou, the terrain-related assumptions are relatively consistent with the framework design, although city-specific indicators and parameters still require recalibration. In mountainous cities, elevation-based connectivity may be insufficient to represent complex flow pathways, and DEM-derived flow direction and drainage infrastructure data may be needed [71]. In coastal cities, tidal flooding and storm surge introduce additional hazards, and POI composition and drainage infrastructure may differ from inland cities. The framework should therefore be adapted rather than directly transferred. A preliminary cross-city validation could compare functional-zone patterns, high-risk screening, and network statistics across

representative plain, mountainous, and coastal cities using a common analytical workflow with city-specific recalibration [76].

5. Conclusions

Taking Zhengzhou as a case study, this study developed an integrated framework combining POI semantic–spatial representation, multidimensional flood risk characterization, and a multi-path flood risk connectivity network to reveal how urban functional structures are associated with the spatial distribution and connectivity of flood risk. The main conclusions are as follows.

(1) The proposed semantic–spatial dual-branch autoencoder combined with two-stage clustering revealed the latent spatial-functional organization of urban blocks. The resulting functional categories were interpreted according to dominant POI compositions. Residential areas accounted for 53.9% of all blocks, and commercial zones exhibited the highest POI density (186.4 POIs·ha^{-1}). The two-stage clustering achieved a silhouette coefficient of 0.564 with only 0.2% noise blocks, and the comparison with the dominant POI-based labels yielded a Cohen's kappa of 0.609.

(2) A total of 302 high-risk blocks were identified (20.5% of all blocks). Spatial validation using observed waterlogging points yielded an adjusted precision of 46.97%, a recall of 18.24%, and an F1 score of 0.263 under a 50 m tolerance. A random permutation baseline (1000 permutations) produced a mean F1 of 0.249, indicating limited improvement over random selection. Public service zones had the highest proportion in the higher-risk group (61.3%), while commercial zones had the highest within-zone high-risk proportion (21.9%). The flood-prone road length proportion was 13.9% in the core circle versus 5.7% in peripheral areas.

(3) The flood risk connectivity network covered 99.2% of all blocks (global Moran's I = 0.703), with road-related connections involved in 89.1% of unique network edges. Connectivity intensity differed significantly between public service and residential zones (adjusted $p = 0.016$) and between public service and industrial zones (adjusted $p = 0.029$). Commercial zones showed the highest descriptive connectivity intensity and range, but these differences were not statistically significant after Bonferroni correction. Network degree centrality was significantly associated with connectivity intensity and range ($p < 0.001$).

(4) The mean standardized functional mixture degree was 0.629. In the GWR model, the median local coefficients of H_norm and D were 0.0015 and 0.0111, respectively, with positive coefficients in 51.3% and 56.0% of blocks. The GWR model achieved an R^2 of 0.792, but the Monte Carlo test indicated no significant global non-stationarity, and the residual Moran's I was 0.310 ($p = 0.001$).

Overall, the findings indicate that urban functions are associated with flood-risk patterns, with these associations reflected through local exposure characteristics and spatial connectivity. By integrating functional-zone identification, block-scale risk characterization, and multi-path connectivity analysis, the proposed framework contributes to refined cognition and management of urban flood risk and provides evidence to support differentiated flood-risk management in urban planning.

Author Contributions: Methodology, Meiyan Gao; data curation, Meiyan Gao; writing—original draft preparation, Meiyan Gao; writing—review & editing, Meiyan Gao and Zongmin Wang; supervision, Zongmin Wang. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The model code, network parameter settings, and derived block-level indicator table are available at https://doi.org/10.5281/zenodo.22073742. The underlying POI and waterlogging datasets are not publicly available due to access restrictions, but they can be requested from the corresponding author.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT-5.6 for language polishing and translation purposes. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

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

Appendix A

Threshold Selection

The composite flood risk index exhibited a left-skewed distribution (skewness = −0.735; Figure A1), with most values concentrated at relatively high levels and a small number of lower-value observations forming the left tail. Following the IPCC definition of extreme climate events, the 85th percentile (threshold = 0.5765) was adopted as the objective statistical threshold for high-risk units, focusing on tail-end extreme risks. A total of 221 high-risk blocks were identified among the 1474 blocks in the study area. A physical threshold was simultaneously introduced as a supplementary criterion, defining physically high-risk areas as units simultaneously meeting an elevation below the 10th percentile of the entire domain and a distance to water buffer of less than 500 m. The final high-risk units were determined by the union of the statistical and physical thresholds, ensuring that topographically low-lying areas with absolute hazard were not overlooked.

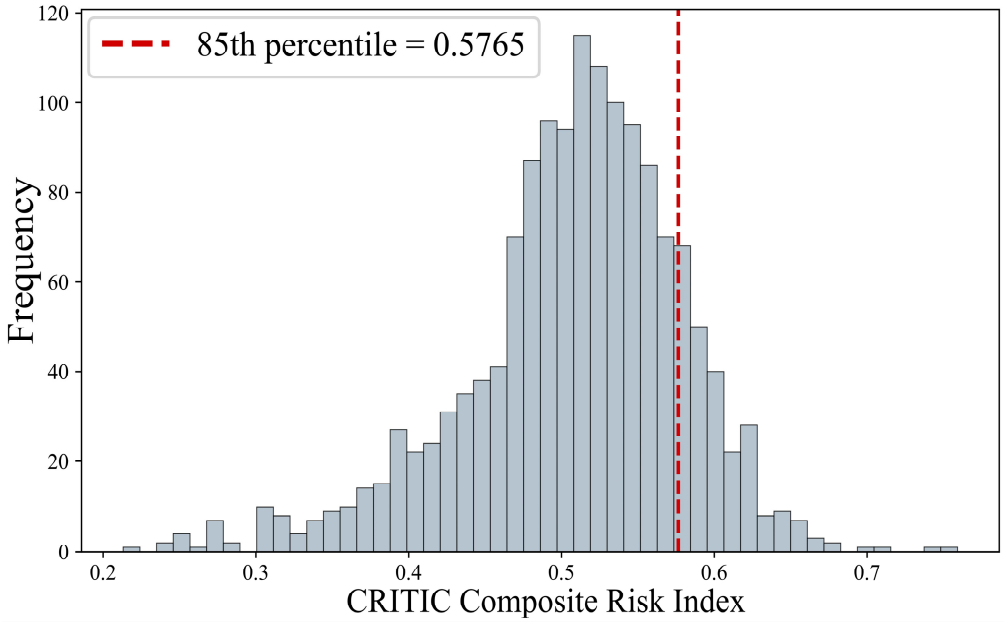


Figure A1. Distribution histogram of the composite flood risk index with the 85th percentile threshold.

References

1. Zhou, L.L.; Shi, Y.S.; Xie, M.Q. Urban Complexity and the Dynamic Evolution of Urban Land Functions in Yiwu City: A Micro-Analysis with Multi-Source Big Data. *Land* **2024**, *13*, 312. [CrossRef]
2. He, X.; Kozlowski, M.; Ujang, N.B.; Ma, Y. Defining Inner-City Transitional Street Typology Using Point of Interest (PoI) Data in Hillside Cities of China. *Sustainability* **2024**, *16*, 4690. [CrossRef]
3. Gao, M.X.; Guo, H.J.; Liu, L.W.; Zeng, Y.Y.; Liu, W.K.; Liu, Y.F.; Xing, H.F. Integrating street view imagery and taxi trajectory for identifying urban function of street space. *Geo-Spat. Inf. Sci.* **2025**, *28*, 1085–1107. [CrossRef]

4. Sun, Z.H.; Jiao, H.Z.; Wu, H.; Peng, Z.H.; Liu, L.B. Block2vec: An Approach for Identifying Urban Functional Regions by Integrating Sentence Embedding Model and Points of Interest. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 339. [CrossRef]
5. Chen, S.L.; Tao, H.Y.; Li, X.L.; Zhuo, L. Detecting urban commercial patterns using a latent semantic information model: A case study of spatial-temporal evolution in Guangzhou, China. *PLoS ONE* **2018**, *13*, e0202162. [CrossRef] [PubMed]
6. Tang, L.Y.; He, J.G.; Peng, W.; Huang, H.Y.; Chen, C.C.; Yu, C. Assessing the visibility of urban greenery using MLS LiDAR data. *Landsc. Urban Plan.* **2023**, *232*, 104662. [CrossRef]
7. Yan, J.F.; Feng, P.F.; Jia, F.X.; Su, F.Z.; Wang, J.; Wang, N.H. Identification of secondary functional areas and functional structure analysis based on multisource geographic data. *Geocarto Int.* **2023**, *38*, 2191995. [CrossRef]
8. Bhere, S.; Reddy, M.J. Flood risk assessment of suburban region in India by incorporating flood hazard, vulnerability and exposure. *Nat. Hazards* **2025**, *121*, 6625–6649. [CrossRef]
9. Zhu, J.L.; Zhou, W.Q.; Yu, W.J.; Wang, W.M. Block-level spatial integration of population density, social vulnerability, and heavy precipitation reveals intensified urban flooding risk. *Sustain. Cities Soc.* **2024**, *117*, 105984. [CrossRef]
10. Chen, X.C.; Wang, Z.M.; Yang, H.B.; Liang, Q.H.; Li, J.H.; Cai, Y.C. Assessing dynamic flood vulnerability variations in urban functional zones using dynamic population data and a PSO-based weighting approach. *Int. J. Disaster Risk Reduct.* **2025**, *116*, 105154. [CrossRef]
11. Wei, N.; Lin, Y.; Zheng, H. Prediction of the flood distribution caused by returning cropland to forest based on Generative Adversarial Network and multi-source remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *143*, 104790. [CrossRef]
12. Du, J.K.; Sha, H.J.; Wang, Z.L.; Zeng, Z.Y.; Tan, C.; Zhao, B.K.; Hu, N.; Lai, C.G. Refined impact assessment of rainfall-induced urban flood in a densely populated and built-up watershed. *J. Hydrol.-Reg. Stud.* **2025**, *62*, 102927. [CrossRef]
13. Liu, X.Y.; Yang, S.I.; Ye, T.; An, R.; Chen, C.Z. A new approach to estimating flood-affected populations by combining mobility patterns with multi-source data: A case study of Wuhan, China. *Int. J. Disaster Risk Reduct.* **2021**, *55*, 102106. [CrossRef]
14. Qin, X.L.; Wang, S.F.; Meng, M.; Long, H.Y.; Zhang, H.L.; Shi, H.C. Enhancing urban resilience through machine learning-supported flood risk assessment: Integrating flood susceptibility with building function vulnerability. *npj Urban Sustain.* **2025**, *5*, 19. [CrossRef]
15. Kim, J.; Kim, H.; Yang, J. Vertical Urban Functional Pattern Analysis Based on Multi-Dimensional Geo Data Cube. *ISPRS Int. J. Geo-Inf.* **2026**, *15*, 47. [CrossRef]
16. Huang, K.; Li, J.Q.; Meng, X.; Zhang, G.H.; Wu, L. Modeling Urban Flood Disaster Chain Propagation Based on Coupling Pathways of Exposed Units. *Int. J. Disaster Risk Sci.* 2026, *advance online publication*. [CrossRef]
17. Li, X.F.; Qiu, F.; Li, K.; Jia, Y.C.; Xia, J.N.; Aishanjian, J. Structural-Functional Suitability Assessment of Yangtze River Waterfront in the Yichang Section: A Three-Zone Spatial and POI-Based Approach. *Land* **2026**, *15*, 91. [CrossRef]
18. Dhungana, B.; Liu, W.B. Urban-Rural Exposure to Flood Hazard and Social Vulnerability in the Conterminous United States. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 339. [CrossRef]
19. Yang, J.; Huang, X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 3907–3925. [CrossRef]
20. Li, Z.H.; He, W.; Cheng, M.F.; Hu, J.X.; Yang, G.Y.; Zhang, H.Y. SinoLC-1: The first 1 m resolution national-scale land-cover map of China created with a deep learning framework and open-access data. *Earth Syst. Sci. Data* **2023**, *15*, 4749–4780. [CrossRef]
21. Zhang, F.; Zu, J.Y.; Hu, M.Y.; Zhu, D.; Kang, Y.H.; Gao, S.; Zhang, Y.; Huang, Z. Uncovering inconspicuous places using social media check-ins and street view images. *Comput. Environ. Urban Syst.* **2020**, *81*, 101478. [CrossRef]
22. Xu, Z.Z.; Cui, G.; Zhong, M.; Wang, X. Anomalous Urban Mobility Pattern Detection Based on GPS Trajectories and POI Data. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 308. [CrossRef]
23. Cao, S.; Zhao, X.S.; Du, S.H. Multi-type and fine-grained urban green space function mapping based on BERT model and multi-source data fusion. *Int. J. Digit. Earth* **2024**, *17*, 2308723. [CrossRef]
24. Ning, F.W.; Li, Z.R.; Lu, J.X.; Wang, Y.X.; Niu, Y.X.; Shi, Y. 3D CAD model dynamic clustering based on inertial feature encoder. *Appl. Soft Comput.* **2025**, *182*, 113627. [CrossRef]
25. Bagirov, A.M.; Aliguliyev, R.M.; Sultanova, N. Finding compact and well-separated clusters: Clustering using silhouette coefficients. *Pattern Recognit.* **2023**, *135*, 109144. [CrossRef]
26. de Jeneret, P.D.; Diamantis, I. Topology of currencies: Persistent homology for FX Co-movements: A Comparative clustering study. *Data Sci. Finance Econ.* **2026**, *6*, 183–211. [CrossRef]
27. Tran, T.N.; Drab, K.; Daszykowski, M. Revised DBSCAN algorithm to cluster data with dense adjacent clusters. *Chemom. Intell. Lab. Syst.* **2013**, *120*, 92–96. [CrossRef]
28. Yuan, X.D.; Dai, X.J.; Zou, Z.D.; He, X.; Sun, Y.C.; Zhou, C.S. Multi-Source Data-Driven Extraction of Urban Residential Space: A Case Study of the Guangdong-Hong Kong-Macao Greater Bay Area Urban Agglomeration. *Remote Sens.* **2024**, *16*, 3631. [CrossRef]
29. Park, K.; Choi, S.H. Analysis of urban flood risk for the implementation of sustainable land use measures. *Clim. Serv.* **2025**, *38*, 100568. [CrossRef]

30. Almeida, M.D.; Telhado, M.J.; Morais, M.; Barreiro, J. Multisector Risk Identification to Assess Resilience to Flooding. *Climate* **2021**, *9*, 73. [CrossRef]
31. Sun, J.; Zhai, N.N.; Mu, H.R.; Miao, J.C.; Li, W.X.; Li, M.F. Assessment of urban resilience and subsystem coupling coordination in the Beijing-Tianjin-Hebei urban agglomeration. *Sustain. Cities Soc.* **2024**, *100*, 105058. [CrossRef]
32. Di, Z.G.; Li, J.M.; Tao, S.Y.; Jia, C.R. Multi-Scale Dynamic Fusion Classification Network Based on Mamba. *Laser Optoelectron. Prog.* **2026**, *63*, 0228007. [CrossRef]
33. He, Q.P.; Wu, M.; Zhang, P.C.; Wang, L.; Shi, Q.B. Multimodal Image Segmentation with Dynamic Adaptive Window and Cross-Scale Fusion for Heterogeneous Data Environments. *Appl. Sci.* **2025**, *15*, 10813. [CrossRef]
34. Zhi, J.J.; Ge, L.W.; Geng, T.; Zhang, Z.H.; Li, L.; Zhu, H.; Zhou, Z.Q.; Jiang, W.; Qu, L.A.; Su, Y.; et al. Accurately mapping social functional zones of urban green spaces by integrating remote sensing images and crowd-sourced geospatial data. *Int. J. Digit. Earth* **2024**, *17*, 1–26. [CrossRef]
35. Wen, R.K.; Sun, J.J.; Tao, C.L.; Tao, H.; N'Tani, C.; Yang, L. Simulation of spatial flooding disaster on urban roads and analysis of influencing factors: Taking main city of Hangzhou as an example. *Stoch. Environ. Res. Risk Assess.* **2024**, *38*, 4151–4168. [CrossRef]
36. Zhang, X.L.; Kang, A.Q.; Ye, M.; Song, Q.X.; Lei, X.H.; Wang, H. Influence of Terrain Factors on Urban Pluvial Flooding Characteristics: A Case Study of a Small Watershed in Guangzhou, China. *Water* **2023**, *15*, 2261. [CrossRef]
37. Kalantari, Z.; Nickman, A.; Lyon, S.W.; Olofsson, B.; Folkeson, L. A method for mapping flood hazard along roads. *J. Environ. Manag.* **2014**, *133*, 69–77. [CrossRef]
38. Luo, Z.Y.; Tian, J.; Zeng, J.; Pilla, F. Flood risk evaluation of the coastal city by the EWM-TOPSIS and machine learning hybrid method. *Int. J. Disaster Risk Reduct.* **2024**, *106*, 104435. [CrossRef]
39. Lang, Y.; Jiang, Z.; Wu, X. Investigating the Linkage between Extreme Rainstorms and Concurrent Synoptic Features: A Case Study in Henan, Central China. *Water* **2022**, *14*, 1065. [CrossRef]
40. Lv, H.; Wu, Z.N.; Zheng, X.K.; Yan, D.M.; Yu, Z.L.; Shang, W.X. Multi-attribute diagnosis of urban flood-bearing bodies based on integrated learning with Stacking-GPR-QPSO coupling. *J. Hydrol.* **2024**, *645*, 132222. [CrossRef]
41. Bhaumik, K.; Das, S. Computational hydrodynamics based flow accumulation models to identify urban waterlogging at deltaic plain using GIS. *Environ. Earth Sci.* **2022**, *81*, 201. [CrossRef]
42. Chang, H.S.; Su, Q.M.; Jheng, D. Comparing the spatial patterns of flooding and individual risk perception: A case study of Yunlin, Taiwan. *Urban Clim.* **2022**, *45*, 101260. [CrossRef]
43. Zhang, K.H.; Mei, C.; Liu, J.H.; Wang, J.; Song, T.X.; Shi, H.Y.; Zhang, A.Q.; Cao, X.J. Spatiotemporal evolution analysis of urban lifeline infrastructure disaster chain network based on numerical simulation and cascading effects. *Sustain. Cities Soc.* **2025**, *133*, 106880. [CrossRef]
44. Chen, Z.C.; Zheng, C.J.; Xu, M.; Wu, Z.L.; Zheng, S.K.; Ma, G.H. Quantitative assessment of congestion diffusion and cascading effect under rainfall-flood disasters: A case study of Nanjing, China. *Int. J. Disaster Risk Reduct.* **2024**, *114*, 104915. [CrossRef]
45. Zhou, P.; Tang, X.M.; Wang, Z.M.; Cao, N.; Wang, X. Vertical Accuracy Effect Verification for Satellite Imagery with Different GCPs. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1268–1272. [CrossRef]
46. Chen, Y.G. The distance-decay function of geographical gravity model: Power law or exponential law? *Chaos Solit. Fractals* **2015**, *77*, 174–189. [CrossRef]
47. Ghalmane, Z.; Cherifi, C.; Cherifi, H.; El Hassouni, M. Extracting modular-based backbones in weighted networks. *Inf. Sci.* **2021**, *576*, 454–474. [CrossRef]
48. Kays, H.M.I.; Sadri, A.M.; Muraleetharan, K.K.; Harvey, P.S.; Miller, G.A. Exploring the Interdependencies Between Transportation and Stormwater Networks: The Case of Norman, Oklahoma. *Transp. Res. Rec.* **2024**, *2678*, 491–513. [CrossRef]
49. Forsberg, S.; Thomas, K.; Bergkvist, M. Power grid vulnerability analysis using complex network theory: A topological study of the Nordic transmission grid. *Phys. A Stat. Mech. Appl.* **2023**, *626*, 129072. [CrossRef]
50. Yue, Y.; Zhuang, Y.; Yeh, A.G.O.; Xie, J.Y.; Ma, C.L.; Li, Q.Q. Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 658–675. [CrossRef]
51. Wang, M.; Yuan, M.C.; Han, P.Y.; Wang, D. Assessing sustainable urban development based on functional spatial differentiation of urban agriculture in Wuhan, China. *Land Use Policy* **2022**, *115*, 105999. [CrossRef]
52. Zhou, J.J.; Zhang, S.X.; Wang, H.; Ding, Y. Spatial Coupling Analysis of Urban Waterlogging Depth and Value Based on Land Use: Case Study of Beijing. *Water* **2025**, *17*, 529. [CrossRef]
53. Yu, J.N.; Zhang, H.C.; Wang, P.X.; Wang, J.Z.; Lu, F. Sequence analysis of local indicators of spatio-temporal association for evolutionary pattern discovery. *GISci. Remote Sens.* **2025**, *62*, 2487292. [CrossRef]
54. Sun, H.; Wang, Q.; Wang, G.X.; Lin, H.; Luo, P.; Li, J.P.; Zeng, S.Q.; Xu, X.Y.; Ren, L.X. Optimizing kNN for Mapping Vegetation Cover of Arid and Semi-Arid Areas Using Landsat Images. *Remote Sens.* **2018**, *10*, 1248. [CrossRef]
55. Li, Z.Q.; Fotheringham, A.S. Computational improvements to multi-scale geographically weighted regression. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 1378–1397. [CrossRef]

56. Geniaux, G. Top-down scale approaches for multiscale GWR with locally adaptive bandwidths. *J. Geogr. Syst.* **2026**, *28*, 27–76. [CrossRef]
57. Li, Y.G.; Liu, X.T.; Han, Z.; Dou, J. Spatial Proximity-Based Geographically Weighted Regression Model for Landslide Susceptibility Assessment: A Case Study of Qingchuan Area, China. *Appl. Sci.* **2020**, *10*, 1107. [CrossRef]
58. Oshan, T.M.; Li, Z.Q.; Kang, W.; Wolf, L.J.; Fotheringham, A.S. MGWR: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 269. [CrossRef]
59. Lu, B.B.; Hu, Y.G.; Murakami, D.; Brunsdon, C.; Comber, A.; Charlton, M.; Harris, P. High-performance solutions of geographically weighted regression in R. *Geo-Spat. Inf. Sci.* **2022**, *25*, 536–549. [CrossRef]
60. Zhong, M.; Chen, T.L.; Zhuo, L.; Wang, Z.Q.; Ling, F.; Han, D.W. Exploring characteristics and drivers of flood hazard under different urban development patterns. *Environ. Sustain. Indic.* **2025**, *28*, 100955. [CrossRef]
61. Al Rifat, S.A.; Liu, W.B. Predicting future urban growth scenarios and potential urban flood exposure using Artificial Neural Network-Markov Chain model in Miami Metropolitan Area. *Land Use Policy* **2022**, *114*, 105994. [CrossRef]
62. Kwon, S.H.; Lee, W.J.; Kang, J.H.; Jun, H.D. Incorporating Pipe Age and Sizes into Pipe Roughness Coefficient Estimation for Urban Flood Modeling: A Scenario-Based Roughness Approach. *Sustainability* **2025**, *17*, 7989. [CrossRef]
63. O'Neill, B.F.; Boyer, A.L. 'Locking in' Desalination in the U.S.-Mexico Borderlands: Path Dependency, Techno-Optimism and Climate Adaptation. *Water Altern. Interdiscip. J. Water Polit. Dev.* **2023**, *16*, 480–508.
64. Fatemi, M.N.; Okyere, S.A.; Diko, S.K.; Kita, M.; Shimoda, M.; Matsubara, S. Physical Vulnerability and Local Responses to Flood Damage in Peri-Urban Areas of Dhaka, Bangladesh. *Sustainability* **2020**, *12*, 3957. [CrossRef]
65. Wang, M.; Zhao, J.Y.; Xiong, Z.H.; Zhang, M.H.; Wang, L.; Tan, S.K. Strategic deployment of nature-based solutions for urban flood management in high-density urban landscapes. *Ecol. Indic.* **2025**, *176*, 113681. [CrossRef]
66. Lu, Q.O.; Bahramloo, R.; Rodrigo-Comino, J.; Wang, J.; Talebi, A.; Tran, Q.T.P.; Ghahramani, A.; Sepehri, M. Evaluating the impact of roof rainwater harvesting on hydrological connectivity and urban flood mitigation. *Results Eng.* **2025**, *25*, 104022. [CrossRef]
67. Liu, R.; Dai, Q.; Xiao, Y.Y.; Lai, X.Y.; Zhang, J.; Wang, H.M. Designing Flood-Resilient Cities: How Urban Morphology Shapes Population Exposure Risk. *Risk Anal.* **2026**, *46*, e70218. [CrossRef] [PubMed]
68. Wang, Z.Y.; Ye, X.Y.; Lee, J.; Chang, X.M.; Liu, H.M.; Li, Q.Q. A spatial econometric modeling of online social interactions using microblogs. *Comput. Environ. Urban Syst.* **2018**, *70*, 53–58. [CrossRef]
69. Zhou, R.T.; Zheng, H.; Liu, Y.Y.; Xie, G.T.; Wan, W.H. Flood impacts on urban road connectivity in southern China. *Sci. Rep.* **2022**, *12*, 16866. [CrossRef] [PubMed]
70. Ishtiaque, A.; Sangwan, N.; Yu, D.J. Robust-yet-fragile nature of partly engineered social-ecological systems: A case study of coastal Bangladesh. *Ecol. Soc.* **2017**, *22*, 5. [CrossRef]
71. Samadi, A.; Jafarzadegan, K.; Moradkhani, H. DEM-based pluvial flood inundation modeling at a metropolitan scale. *Environ. Model. Softw.* **2025**, *183*, 106226. [CrossRef]
72. Wang, Z.Y.; Liu, G.J.; Hu, Y.; Sun, L. Evaluating Urban Underground Space Supply-Demand Imbalances Based on Remote Sensing and POI Data: Evidence from Nanjing, China. *Land* **2025**, *14*, 1671. [CrossRef]
73. Bakhtiari, V.; Piadeh, F.; Behzadian, K.; Kapelan, Z. A critical review for the application of cutting-edge digital visualisation technologies for effective urban flood risk management. *Sustain. Cities Soc.* **2023**, *99*, 104958. [CrossRef]
74. Chakraborty, L.; Thistlethwaite, J.; Scott, D.; Henstra, D.; Minano, A.; Rus, H. Assessing social vulnerability and identifying spatial hotspots of flood risk to inform socially just flood management policy. *Risk Anal.* **2023**, *43*, 1058–1078. [CrossRef] [PubMed]
75. Zhang, Y.; Chen, Z.Q.; Zheng, X.; Chen, N.C.; Wang, Y.Q. Extracting the location of flooding events in urban systems and analyzing the semantic risk using social sensing data. *J. Hydrol.* **2021**, *603*, 127053. [CrossRef]
76. Yan, Q.X.; Yuan, J.C.; Wu, D.; Lin, Y.F.; Lian, Z. Spatiotemporal Assessment and Obstacle Factor Analysis of Urban Flood Resilience in the Shenyang Metropolitan Area Based on an LSTM-Attention Model. *Sustainability* **2026**, *18*, 50. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.