

Article

Coupling Delphi-Driven Expert Elicitation with Bayesian Networks in GIS: An Advanced Approach to Quantifying and Mapping River Flood Risk

Bingyu Zhang ^{1,2} , Jing Qin ^{1,2,*}, Zhen Wang ^{1,2}, Lingyun Zhao ^{1,2}, Lu Wang ^{1,2}  and Wencai Ma ³

¹ China Institute of Water Resources and Hydropower Research (IWHR), Beijing 100048, China; bingyu_93@163.com (B.Z.)

² Beijing China Institute of Water Resources and Hydropower Research Engineering Group Co., Ltd., Beijing 100048, China

³ Qinghai Provincial Flood and Drought Disaster Prevention Service Center, Xining 810001, China

* Correspondence: qinjing@iwhr.com

Abstract

Flood disaster risk assessment serves as an important foundation for formulating regional sustainable development strategies. This study establishes a risk assessment model for flood disasters in small and medium-sized rivers based on a theoretical framework integrating Geographic Information Systems (GIS), the Delphi method, and Bayesian networks (Delphi–BNs). An indicator system for the assessment was developed from three dimensions: hazard, vulnerability, and exposure. Hazard is represented by flood inundation area and depth; vulnerability is indicated by population distribution and economic layout; and exposure is reflected by road accessibility. By constructing a GIS-based Bayesian network and employing the Delphi method to create a probabilistic and spatially explicit model, this approach quantifies various sources of uncertainty in the assessment process, enabling a probabilistic expression of risk. Based on the risk assessment results, a stratified, phased flood emergency rescue and personnel transfer plan was established, designating extremely high-risk areas as the core zones for the first phase of personnel transfer, high-risk areas as the second-phase rescue zones, and medium-risk areas as the third-phase rescue zones, thereby providing clear operational guidance for flood emergency response in the basin. The Delphi–BN assessment framework developed in this study focuses on the core elements of flood disaster risk formation, organically integrates expert experience with spatial big data, and effectively overcomes the limitations of traditional assessment methods, such as strong subjectivity, insufficient accuracy, and poor quantification. It achieves a refined and quantitative assessment of flood risk in small and medium-sized river basins in semi-arid regions. The outcomes of this research contribute to a clearer understanding of both the driving mechanisms and the spatial patterns of regional flood risk. Furthermore, they establish a scientifically credible and operationally relevant foundation for key disaster-response decisions, encompassing timely emergency actions, phased population transfers, and the optimized deployment of limited emergency resources.

Keywords: Delphi–BNs; flood disaster risk assessment; hazard; exposure; vulnerability; flood emergency response



Academic Editor: Maria Mimikou

Received: 17 July 2026

Revised: 21 August 2026

Accepted: 21 August 2026

Published: 23 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The increasing frequency of extreme precipitation events, attributed to global climate change, has rendered flood disasters a critical natural threat to the safety of human lives

and property worldwide [1,2]. Although Northwest China is generally arid or semi-arid, the flash floods and debris flows triggered by localized heavy rainfall in the region cannot be ignored. Located within the Huangshui River basin, an important tributary of the upper Yellow River, the Xina Chuan basin is characterized by a combination of valley plains and hilly mountainous terrain. Heavy summer rains often cause rapid river rises, flooding of farmland, and traffic disruptions, highlighting an urgent need for refined flood risk assessment in this area [3,4]. Therefore, local flood control authorities in Qinghai Province place great emphasis on flood disaster prevention efforts. However, how to conduct flood disaster assessment remains an unresolved challenge in the field of disaster research, primarily because the academic community has yet to establish a unified standard for quantifying flood risk. Flood disaster assessment is currently an active research topic [5].

The core link in flood disaster assessment lies in accurately identifying and prioritizing the plots within the affected area that suffer the greatest losses and have the most prominent impacts, thereby enabling efficient coordination and precise allocation of limited rescue resources to maximize the effectiveness of emergency response and disaster reduction [6,7]. Flood disaster assessment itself involves uncertainty issues [8]; therefore, it is necessary to adopt scenario analysis methods that can encompass various types of uncertainty, such as those arising from model structure or input data [9]. However, in flood disaster assessment itself, the communication of uncertainty information is often overlooked, as this is not an easy task. Furthermore, data scarcity constrains the development of flood disaster assessment. Commonly used flood models require multiple types of high-resolution data, such as DEM, DOM, and land use data, which are often difficult to obtain. With rapid economic development, existing local data quickly become outdated, which not only hinders the assessment of current flood risk conditions but also makes it impossible to predict future trends [10].

Early warning system cases are representative [11]. Flood warnings rely on risk prediction to achieve disaster prevention effects. In addition to managing tangible losses through off-site asset relocation and control, they offer three types of additional benefits: ① ensuring life safety and reducing direct intangible losses caused by casualties; ② regulating public evacuation behavior, preventing long-term psychological damage left by disasters, and reducing indirect intangible losses; and ③ optimizing evacuation organization, reducing post-disaster evacuation support costs, and saving indirect tangible costs [12]. The most widely applied framework in the field of disaster risk reduction (DRR) typically defines flood disaster risk as a function of the interaction between hazard and the physical vulnerability and exposure level of the elements at risk [13].

Although traditional flood risk assessment methods, such as hydrological and hydrodynamic models, statistical regression, and multi-criteria decision analysis, can provide some reference, they have limitations in handling nonlinear interactions among multiple factors, data uncertainty, and spatial heterogeneity [14]. Bayesian networks (BNs), with their advantage in probabilistic reasoning and uncertainty handling, have been widely used for environmental risk modeling [15–17]. The robust spatial data management and visualization capabilities of geographic information systems (GIS), when combined with BNs, make it possible to capture intricate dependencies among risk factors and to present risk assessment results in a spatially explicit manner [18].

Recent studies have explored the integration of BNs and GIS for flood risk assessment, and these efforts can be broadly categorized into three main streams. The first stream adopts data-driven approaches, where BN parameters are learned directly from historical flood records and spatial datasets. For instance, GIS and Bayesian belief networks have been used to assess urban pluvial flood vulnerability [19]; a BN model was trained using eight historical flood events in Houston and validated against traditional indicator-based meth-

ods [20]; another BN model was trained using the 2014 Secchia River flood event to predict multi-sectoral losses [21]; and a BNs-GIS framework was developed for flood disaster risk assessment in Yinchuan, China, synthesizing diverse data sources, including hydrological models, infrastructure details, use patterns, remote sensing data, and ecological indicators, to produce a holistic risk assessment [22]. However, the effectiveness of these data-driven approaches is highly dependent on the availability and quality of high-quality historical data, making them difficult to apply in data-scarce regions. The second stream relies on expert knowledge to define BN structures and populate conditional probability tables (CPTs). For example, a spatial BN model was constructed through expert questionnaires to assess urban flood risk to people [12]. While this approach can compensate for data scarcity, existing studies often lack a structured, transparent, and replicable protocol for eliciting and aggregating expert opinions. The third stream treats GIS spatial analysis and BN probabilistic inference as separate functional components rather than integrating them methodologically. In practice, GIS is used for data preprocessing and visualization, while BNs are employed for probabilistic computation, resulting in limited synergy between the two techniques. For instance, a BN model and GIS were applied to assess urban flood disaster risk in Zhengzhou, where GIS was primarily used for spatial data extraction and processing while the BN handled probabilistic inference [18]. Similarly, a spatial framework integrating naïve Bayes and GIS was developed for flooding hazard assessment in Australia's Bowen Basin, employing ArcGIS (v10.8) for indicator derivation and naïve Bayes for probability inference on a cell-by-cell basis [23]. GIS-derived exposure and susceptibility indicators were combined with BN-based probabilistic modelling for agricultural flood risk assessment, demonstrating a similar functional division where GIS provides inputs and BN performs probabilistic reasoning [24]. These approaches, while effective for their specific purposes, primarily chain GIS and BNs as separate modules rather than achieving a deep methodological integration where the two techniques mutually inform and constrain each other throughout the modelling process.

To address these gaps, this study presents a novel methodological framework that systematically integrates a structured multi-round Delphi expert elicitation process with BNs and GIS. Unlike previous data-driven approaches that require extensive historical data, our method is specifically designed for data-scarce regions such as the Xina Chuan basin. Unlike prior expert-based BN applications that rely on unstructured or non-transparent expert input, our framework provides a transparent, replicable, and auditable protocol for translating diverse expert judgments into probabilistic model parameters through a rigorous Delphi process with predefined consensus criteria ($IQR \leq 1.0$ and $C_v \leq 0.25$). Unlike existing GIS-BN combinations that merely couple the two tools, our approach achieves a deep methodological integration, wherein GIS is used for spatial data preprocessing and result visualisation, BNs for probabilistic inference, and Delphi for systematic knowledge elicitation, thereby creating a coherent workflow that is greater than the sum of its parts. Furthermore, while traditional BNs-GIS methods are mostly general-purpose frameworks that overlook the unique characteristics of small and medium-sized river basins in semi-arid regions, such as complex terrain and high flood suddenness, our framework is specifically tailored to the Xina Chuan basin. This customization is achieved by using locally adapted indicators, namely inundation extent, inundation depth, road accessibility, population vulnerability, and economic vulnerability, and by incorporating expert knowledge from local specialists in hydrology, meteorology, emergency management, and urban planning.

Specifically, the Delphi scoring method is used to invite experts in fields such as hydraulic engineering, disaster management, and geographic information science to conduct scientific assessment and scoring on the selection of flood risk assessment indicators, the weighting of indicators, and the initial assignment of CPTs based on the actual characteris-

tics of the Xina Chuan small and medium-sized river basin. This addresses the drawback of high subjectivity in CPT assignment in traditional BN construction and improves the scientific rigor and rationality of the assessment. With Bayesian networks serving as the core of probabilistic inference, the Delphi scoring results are used to quantify the nonlinear coupling relationships among the three dimensions of factors, capture flood risk characteristics under complex scenarios, and simultaneously quantify various uncertainties in the assessment process to achieve a probabilistic expression of risk assessment. Relying on the spatial data processing and visualization capabilities of GIS technology, the Delphi–BN coupled model is provided with precise spatial background data support. At the same time, the risk results calculated by the model are transformed into spatial risk maps, enabling an intuitive representation of risk distribution.

To this end, facing multiple scenarios of climate and socio-economic change, this paper proposes a flood disaster risk assessment method. This method includes a flood risk assessment model based on an indicator system covering three dimensions: hazard, vulnerability, and exposure. Flood hazard is jointly characterized by inundation area and inundation depth; vulnerability is characterized by population density and the distribution of economic output; and exposure is characterized by road accessibility. The method accurately depicts future flood disaster scenarios in small and medium-sized river basins and is suitable for spatial assessment of flood disaster risk in data-scarce regions. The study integrates multi-type heterogeneous data based on a Bayesian network, relying on risk impact factors to achieve spatial representation of flood risk in the absence of measured flood inundation data. During model construction, it fully leverages the advantages of Bayesian networks in probabilistic inference and uncertainty analysis to accurately quantify the causal relationships and interaction mechanisms among various risk factors. At the same time, relying on the powerful spatial data processing, spatial analysis, and visualization capabilities of Geographic Information System technology, it accurately characterizes the spatial distribution patterns of flood risk within the basin. By incorporating regional actual disaster conditions and management needs, flood risk levels are clearly classified into four tiers, low, moderate, high, and extremely high, providing a clear hierarchical basis for subsequent flood risk early warning and enabling more effective spatial flood risk management.

2. Flood Risk Assessment Methods

2.1. Flood Simulation

This study uses the 100-year design flood of the Xina Chuan River, along with its inundation extent and inundation depth, for flood risk assessment. The design flood is calculated based on the local hydrological manual, using the rational formula to simulate the 100-year flood process of the Xina Chuan River. The results are further validated using field data, and the outcomes derived from the rational formula are ultimately adopted. The application of this method to flood risk assessment faces two major challenges: (1) the setting of subjective weighting criteria [25] and (2) the uncertainty of parameters in the rational formula [26]. The method consists of two parts, runoff generation and runoff concentration, and the specific steps are as follows.

2.1.1. Runoff Generation

The design flood runoff is composed of three components: direct runoff, interflow, and base flow. Direct runoff and interflow are generated by the design storm and constitute the main body of the design flood, collectively referred to as flood runoff, whose magnitude is equal to the depth of “net rain”. Base flow is supplied by groundwater and is not closely related to the design storm.

Among various conceptual hydrological models, the hyperbolic tangent model incorporates variables such as the rainfall amount over the main rain-bearing area, the duration of the main storm, and the maximum possible loss, which reflect the rainfall intensity characteristics and loss variation in arid and semi-arid zones. This model features clear conceptual foundations, a simple structure, minimal parameter requirements, and ease of calibration. Therefore, this study employs the hyperbolic tangent model to compute runoff generation. The watershed runoff generation calculation includes two parts: the calculation of the net rain depth of the design flood and the calculation of the net rain process.

The hyperbolic tangent model is a pure runoff generation model, and its formula is as follows.

$$R_p = H_P(t_z) - F(t_z) \tanh\left(\frac{H_P(t_z)}{F(t_z)}\right) \quad (1)$$

$$F(t_z) = S_r(1 - B_0)t_z^{0.5} + 2K_s t_z \quad (2)$$

where $\tanh(\cdot)$ is the hyperbolic tangent function; $t_z(h)$ represents the length of the main storm event; R_p (mm) is the designed net rain depth simulated by the model; $H_P(t_z)$ is the designed areal rainfall depth corresponding to the main storm duration t_z (mm); $F(t_z)$ (mm) is the potential loss of the watershed during the main storm period; and S_r (mm/h^{1/2}) is the absorption rate of the air-dry watershed soil, denoting the watershed's comprehensive water absorption capacity, which arises from the spatial assemblage of distinct land use categories. For a specific watershed, it is a physical constant, equal to the area-weighted average of the water absorption capacity of each land use type within the watershed; K_s (mm/h) is the hydraulic conductivity of the saturated watershed soil, also a physical constant for a specific watershed, calculated as the area-weighted average of the hydraulic conductivity of each land use type; and B_0 is the soil water holding degree of the vadose zone before the rainstorm, with the physical meaning being the ratio of the water stored in the vadose zone to its water holding capacity I_{cm} .

The regionalized parameters S_r , K_s , and soil water holding degree B_0 for Qinghai Province refer to the parameter tables corresponding to different zones, soil types, and antecedent rainfall conditions (Tables 1 and 2).

Table 1. Schedule for Qinghai Province regionalized parameters S_r and K_s .

Zone	S_r (mm/h ^{1/2})	K_s (mm/h)	Location	S_r (mm/h ^{1/2})	K_s (mm/h)
Qilian Mountain Semi-arid and Semi-humid Zone	10.64~21.9	0.85~1.31	Nanmen Xia	16.21	0.85
			Xina Chuan	14.09	1.31
			Hai Yan	21.9	1.09
			Gang Cha	10.64	0.87
			Heili Lin	14.75	1.27

Table 2. B_0 value of soil water retention capacity in the semi-arid and semi-humid zones of the Qilian Mountains.

Precipitation in the Last 7 Days (mm)	Soil		
	Sandy Loam	Loamy Soil	Clay Soil
5~10	0.13~0.16	0.16~0.19	0.19~0.22
10~25	0.22~0.25	0.25~0.28	0.28~0.31
25~50	0.31~0.37	0.37~0.43	0.43~0.5
50~	0.5~1.0	0.5~1.0	0.5~1.0

2.1.2. Runoff Concentration

The rational formula method is one of the earliest approaches used to indirectly estimate the peak design flood discharge based on storm data. For design flood calculation in small watersheds, the rational formula method is characterized by its simplicity in computation and reasonable accuracy, making it the most commonly used method for small watershed concentration routing in water conservancy and hydropower projects. The rational formula method is generally applicable to flood design calculations for small watersheds (with an area less than 1000 km²). Given the sparse distribution of hydrological stations in Qinghai Province, the area of selected watersheds may slightly exceed 1000 km² in some cases. When applying the rational formula method in data-sparse regions, it is first necessary to derive the flow concentration parameters based on watershed characteristic parameters, precipitation, vegetation, and geomorphology, and then compute the design flood using numerical or graphical methods.

The rational concentration calculation includes solving for the maximum discharge Q_m , the watershed concentration time t , the concentration velocity V , and the flood hydrograph. The basic principle follows the runoff generation convolution formula based on modern watershed concentration theory. By assuming that either the watershed concentration curve or the net rain process is rectangular, and letting the duration of the net rain that generates the peak flow equal the watershed concentration time t , the peak-generating formula under the condition of full-area runoff is derived as follows:

$$Q_m = \begin{cases} 0.278 \frac{h_T}{\tau} A & t_c > \tau \\ 0.278 \frac{h_R}{\tau} A & t_c \leq \tau \end{cases} \quad (3)$$

where A (km²) is the watershed area, τ (h) is the watershed concentration time, h_T (mm) is the net rain depth for peak generation, i.e., the net rain depth within the time period τ (mm); and h_R (mm) is the flood net rain depth. For the design flood, this is the watershed runoff generation result R_p .

The concentration time of a watershed is defined as the mean duration required for water located at the farthest point from the outlet to traverse the longest flow path L to the outlet section, and is given by the following equation:

$$\tau = 0.278 \frac{L}{V} \quad (4)$$

where L (km) denotes the length of the main river channel, which corresponds to the longest flow path that water traverses. V (m/s) is the watershed concentration velocity, i.e., the average speed at which water flows along the maximum flow path L (m).

2.1.3. Hydrodynamic Simulation

Hydrodynamic simulation was carried out using IFMS (v 4.0.2), the first Chinese general-purpose flood analysis software, which was developed under the leadership of the China Institute of Water Resources and Hydropower Research (IWHR) [27,28]. This software is applicable to almost all types of flood simulation in China, comprising a 1D river network engine capable of handling complex hydraulic engineering operations, a high-resolution 2D flood engine, a fast unstructured mesh generation module, and a pipe network engine. Based on a self-developed GIS platform [29–31], the software has integrated pre- and post-processing functionalities and enables 1D-2D coupling [32–34] as well as coupling between pipe networks and 2D models.

1. 1D Model Principle

The one-dimensional flow motion in river channels is governed by the Saint-Venant equations, which are given as follows:

$$\begin{cases} B \frac{\partial Z}{\partial t} + \frac{\partial Q}{\partial x} = q \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{\alpha Q^2}{A} \right) + gA \frac{\partial Z}{\partial x} + gA \frac{|Q|Q}{K^2} = qV_x \end{cases} \quad (5)$$

where q (m^3/s) is the lateral inflow, Q (m^3/s) stands for the flow rate at the given river cross-section, A (m^2) stands for the cross-sectional area available for flow, B (m) represents the river width, and Z (m) stands for the water level. The component of the lateral inflow velocity V_x (m/s) in the flow direction is generally approximated as zero. K is the conveyance (or modulus of discharge), reflecting the actual flow capacity of the river channel; α denotes the momentum correction coefficient, representing the evenness of the velocity distribution across the cross-section.

2. 2D Model Principle

This study implements the Godunov algorithm for numerical computation in the 2D overland flow model, in which the Riemann problem is resolved using the Roe scheme's approximate Riemann solver. The discretization of the bed slope source term is accomplished via a characteristic decomposition method, which guarantees the conservation property inherent to the model. The friction source term is treated with implicit discretization, a measure taken to strengthen the stability of the model. Second-order accuracy in both time and space is accomplished via the implementation of MUSCL spatial reconstruction and the predictor-corrector method.

The computational domain for flood propagation is complex and may contain various water-related structures. The flow at and around these structures no longer conforms to shallow water flow, and thus cannot be simulated using the shallow water model; such features are commonly referred to as internal boundary conditions. Handling internal boundary conditions is one of the challenging aspects of 2D flood modeling. In this software, by combining empirical formulas with the numerical solution of the model and using a flux-generalized calculation method, the flux across the computational cell edges connected to water-related structures is provided. This flux calculation method not only ensures the well-balanced stability of the model but also accurately computes the flow flux through the structures and approximates the momentum flux.

The depth-averaged 2D shallow water equations can be written in abbreviated form as:

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = S \quad (6)$$

where

$$U = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, F = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{bmatrix}, G = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{bmatrix}, S = \begin{bmatrix} 0 \\ s_x \\ s_y \end{bmatrix}$$

in which h corresponds to the water depth; u and v denote the flow velocity components in the x and y directions, respectively; and s_x and s_y represent the source terms, the expressions for which are provided as below.

$$s_x = -\frac{h}{\rho} \frac{\partial p_a}{\partial x} - gh \frac{\partial z_b}{\partial x} + \frac{\tau_{ax} - \tau_{bx}}{h} + c_x s_y = -\frac{h}{\rho} \frac{\partial p_a}{\partial y} - gh \frac{\partial z_b}{\partial y} + \frac{\tau_{ay} - \tau_{by}}{h} + c_y \quad (7)$$

where P_a denotes the atmospheric pressure imposed upon the water surface; Z_b is the bed elevation; τ_{ax} , τ_{bx} is the wind stress; C_x , C_y is the Coriolis force; and τ_{bx} , τ_{by} is the bottom friction.

3. Conceptual Framework of the 1D-2D Coupled Model

In IFMS (v 4.0.2), the 1D-2D coupled model can be classified into frontal connection and lateral connection, depending on the water exchange mode (see Figure 1).

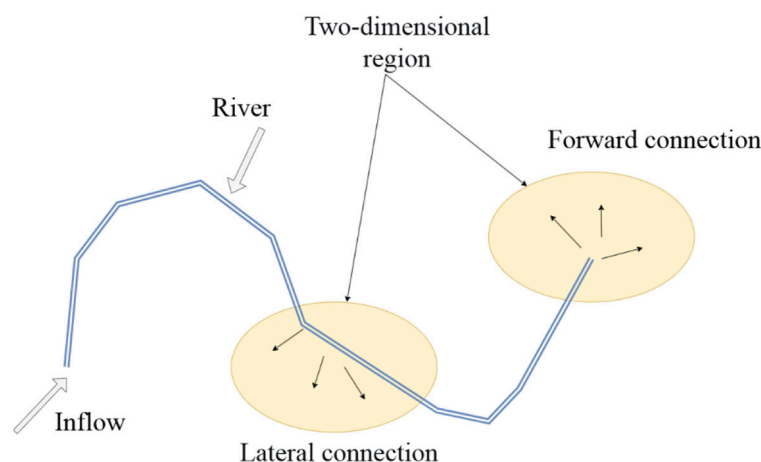


Figure 1. Schematic of two connection types of horizontal water exchange. The arrows indicate the direction of water flow from the 1D model domain to the 2D model domain during model coupling.

Frontal connections are generally established at the start, confluence, and end points of a river channel according to its natural flow direction, thereby specifying the direction of water flow and enabling the calculation of changes in flow rate, water level, etc., within the channel. Lateral connections are defined as the hydraulic interactions between the river channel and adjacent terrain through overbank flow. Such connections typically encompass the exchange of water with the floodplain surface, as well as with flood storage and detention basins. The water exchange at lateral connections is computed using the broad-crested weir flow formula, which quantifies the flow between the river channel and the adjacent terrain.

2.1.4. Performance Evaluation Metrics

The hydraulic model's performance was quantified using three evaluation metrics. The Nash–Sutcliffe efficiency (NSE) was selected to measure the overall fit between simulated and observed hydrographs, calculated as:

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \quad (8)$$

where $Q_{obs,i}$ and $Q_{sim,i}$ are the observed and simulated discharge at time step i , respectively, and $\overline{Q_{obs}}$ denotes the mean observed discharge over the evaluation period.

Furthermore, two flood-specific metrics were used to evaluate peak reproduction capability. The relative peak flow error (R_Q) describes the magnitude deviation between modeled and recorded peak flows:

$$R_Q = \frac{Q_{obs,p} - Q_{sim,p}}{Q_{obs,p}} \quad (9)$$

where $Q_{obs,p}$ and $Q_{sim,p}$ denote the observed and simulated peak discharge values.

The timing error (E_T) was computed as the absolute difference between the observed and simulated peak arrival times:

$$E_T = |T_{obs,p} - T_{sim,p}| \quad (10)$$

where $T_{\text{obs},p}$ and $T_{\text{sim},p}$ are the observed and simulated times of peak discharge, respectively. This metric quantifies the magnitude of phase discrepancy regardless of whether the simulated peak arrives earlier or later than observed.

2.2. Bayesian Networks and Delphi Method for Flood Risk Assessment

2.2.1. Theoretical Foundations of Bayesian Networks

A Bayesian network [35] is a directed acyclic graph (DAG) that captures the probabilistic dependencies among variables. The joint probability distribution over a set of variables $Z_1, Z_2, \dots, Z_n$ can be expressed as the product of conditional probability distributions for each variable. This is expressed as:

$$P(Z_1, Z_2, \dots, Z_n) = \prod_{i=1}^n P(Z_i | \text{Parents}(Z_i)) \quad (11)$$

where $\text{Parents}(Z_i)$ denotes the set of parent nodes of Z_i . In this study, the conditional probability tables (CPTs) are populated through the Delphi expert elicitation procedure, rather than learned from historical data, due to the scarcity of observed flood records in the study area.

2.2.2. Application of Bayesian Networks to Flood Risk

Bayesian networks (BNs) have been widely applied in various environmental modeling contexts, including water quality assessment, ecosystem health analysis, and natural disaster risk prediction, due to their unique ability to handle uncertainty, integrate multiple data sources, and incorporate expert knowledge [36,37]. Traditional assessment methods struggle to characterize multi-factor interactions in flood risk analysis. Bayesian networks, with their flexible structural characteristics, overcome this difficulty by capturing complex relationships among flood risk components based on expert knowledge and existing literature. This provides strong support for research on flood risk assessment, loss quantification, and rescue coordination [38–41]. In the field of flood disaster research, the application of Bayesian networks has gradually expanded from single-event assessments to multi-scenario, multi-scale integrated analyses, encompassing the entire process from risk identification to emergency decision-making [8,42].

Currently, academic research on Bayesian networks is increasingly focusing on core areas such as reliability assessment, risk analysis, and risk management, with a growing number of studies that are increasingly tailored to practical application needs. To address the limitations of traditional assessment methods, one study developed an improved Bayesian network disaster risk assessment model. The key innovations include (1) optimizing the grey relational analysis method, and (2) integrating a genetic algorithm to optimize node probability distributions. These enhancements enable accurate characterization of optimal correlations among risk factors and effective handling of incomplete data and correlated variables, thereby significantly improving the model's scientific rigor and practical applicability [43,44]. Another study confirmed that combining historical disaster statistics with Bayesian network models can significantly improve the identification accuracy of high-risk floodplains. For instance, researchers at the Chinese Research Academy of Environmental Sciences employed Bayesian networks to integrate climate model outputs with historical hydrological observations [45]. This led to the development of a flood cascade risk modeling framework for the Pearl River Delta, which attained an overall floodplain identification accuracy of over 80%. Consequently, this approach offers more precise decision support for disaster prevention and mitigation efforts. Furthermore, to address the challenge of flood loss assessment in data-scarce regions, a study proposed the SMOTE-BN-FLA framework, which integrates the synthetic minority oversampling

technique with Bayesian networks to effectively address the sample imbalance problem and improve the assessment accuracy in high-loss areas [46].

Overall, Bayesian networks can not only effectively handle various uncertainties in the risk assessment process, integrate multiple data sources and expert knowledge, and enhance the robustness of assessment results, but also provide a structured analytical framework that clearly disentangles the mechanisms of flood hazard and loss chains. This lays a solid foundation for the systematic development of research on flood disaster risk assessment, quantitative loss calculation, and rescue coordination [12,47,48]. As research continues to advance, the integrated application of Bayesian networks with geographic information systems and synthetic oversampling techniques is providing a reliable technical pathway for coping with flood risk in the context of climate change [13,37]. This study employs an integration of the Delphi scoring method and Bayesian networks (Delphi-BNs), comprehensively considering the influence of three dimensions on flood risk, which are flood threat, vulnerability, and exposure.

2.2.3. Delphi Method

The structure of the Bayesian network consists of three main components (hazard, exposure, and vulnerability), each represented by a set of input variables (nodes) and their probabilistic dependencies (edges) (Figure 2). The conditional probability tables (CPTs) are generated using the Delphi method. This technique is a structured communication method that systematically collects expert judgment through multiple rounds of anonymous surveys to achieve group consensus. It is a typical “soft evaluation method,” i.e., it relies on the knowledge and experience of experts to address complex problems lacking historical data. The core concept is that the judgment of an organized group of experts is generally more accurate than that of an unorganized group or individuals [49,50]. The implementation of the Delphi method usually follows a rigorous process, mainly consisting of the following steps:

- Define the problem and form an expert panel: Clearly define the problem to be studied, and select 5 to 20 experts with profound knowledge and diverse perspectives in the fields of hydrology, meteorology, emergency management, water resources, etc.
- Design and distribute the first-round questionnaire: Carefully design the questionnaire to present the issues requiring expert judgment, accompanied by background materials. Experts are asked to independently and anonymously provide their opinions and the rationale behind their judgments.
- Collect and aggregate expert opinions (statistical analysis): Perform statistical analysis on the aggregated results (e.g., calculate the median, upper and lower quartiles) to produce summary feedback.
- Conduct a new round of survey (feedback and revision): Provide each expert with the aggregated statistical results from the previous round. Experts are then asked to independently re-evaluate and revise their views, fully considering the panel opinion. This process is repeated for several rounds (typically 3–4 rounds) until expert opinions converge.
- Reach consensus and produce conclusions: Based on the experts’ final round of judgments, collate the results to form a consistent conclusion. At this point, the final consensus can be summarized and obtained.

Based on the CPT parameters established through the Delphi method, the developed BNs-GIS model generates probabilistic predictions of flood risk. Subsequently, the flood risk associated with each assessment unit is categorized into one of four levels, ranging from Level 1 to Level 4.

$$E(Z)_{\text{flood}} = \sum_{k=1}^{\infty} z_i \Pr(Z = z_i) \quad (12)$$

Within this framework, Z_i signifies the risk coefficient score associated with distinct flood risk levels. The term $\Pr$ represents the conditional probability of flood risk occurrence, conditional upon a set of determinants: inundation extent, inundation depth risk, road accessibility, population vulnerability, and economic vulnerability. For the Xina Chuan River basin, the CPTs governing the flood risk Bayesian networks were established using the Delphi method.

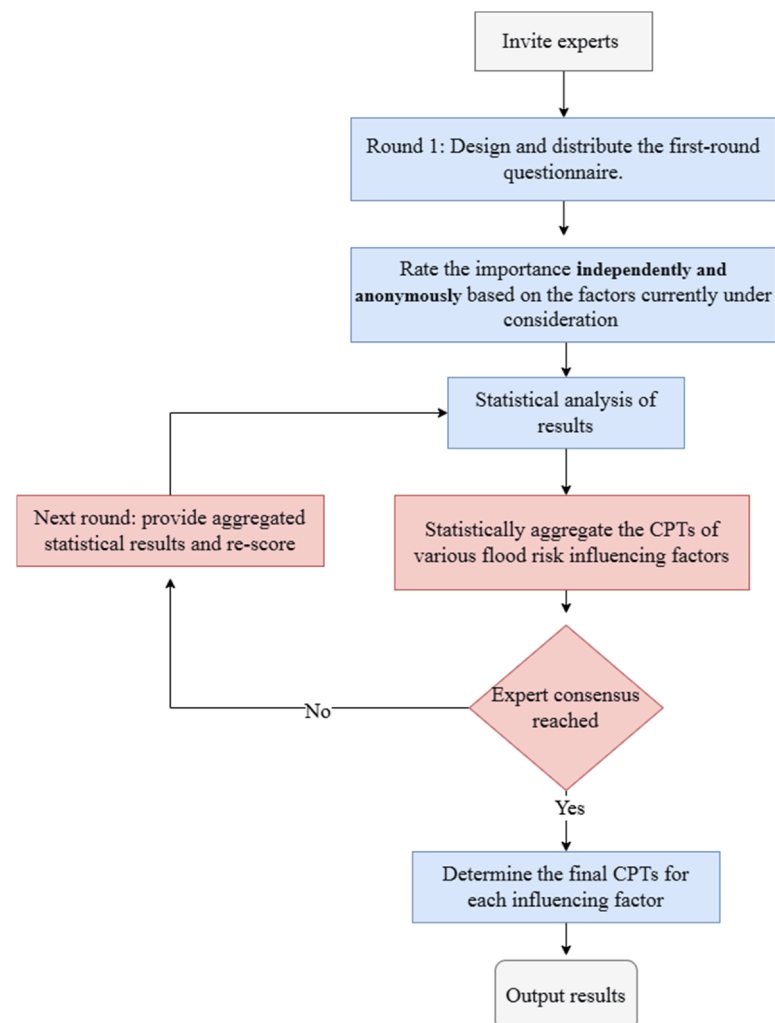


Figure 2. The flowchart of the Delphi method framework.

2.3. Sensitivity Analysis Method

The objective of sensitivity analysis is to evaluate the degree to which assessment indicators and parameters affect the assumptions and outputs of a mathematical model. This approach enables the effective identification of the key indicators exerting the most substantial influence [51,52]. In flood disaster risk assessment studies that integrate Bayesian networks and geographic information systems, the constructed conditional probability tables (CPTs) have a decisive impact on the results. Therefore, a sensitivity analysis is required, and the underlying principle is as follows.

Denote the parent nodes of random variables x and y as u and v , respectively. Under this configuration, the joint probability $Pr(e)$ can be expressed as a second-order polynomial of the CPT parameters:

$$Pr(e) = C + \sum_u C_u \theta_{x|u} + \sum_v C_v \theta_{y|v} + \sum_{u,v} C_{u,v} \theta_{x|u} \theta_{y|v} \quad (13)$$

The first-order partial derivatives are:

$$\frac{\partial Pr(e)}{\partial \theta_{x|u}} = C_u + \sum_v C_{u,v} \theta_{y|v} \quad (14)$$

$$\frac{\partial Pr(e)}{\partial \theta_{y|v}} = C_v + \sum_u C_{u,v} \theta_{x|u} \quad (15)$$

The mixed second-order partial derivative is:

$$\frac{\partial^2 Pr(e)}{\partial \theta_{x|u} \partial \theta_{y|v}} = C_{u,v} \quad (16)$$

Finally, the function $\alpha(\theta_{x|u}, \theta_{y|v})$ is defined as:

$$\alpha(\theta_{x|u}, \theta_{y|v}) = \frac{\partial^2 Pr(z, e)}{\partial \theta_{x|u} \partial \theta_{y|v}} - p \frac{\partial^2 Pr(e)}{\partial \theta_{x|u} \partial \theta_{y|v}} \quad (17)$$

where C denotes a constant. The second-order partial derivatives of $Pr(e)$ are subsequently evaluated with respect to $\theta_{x|u}$ and, thereby spanning the complete set of combinations over u and v .

The aforementioned calculation can be generalized to scenarios involving parameter changes across more than two CPTs. For example, when three CPTs are modified concurrently, it necessitates the computation of third-order partial derivatives with respect to the associated parameters. This methodology quantifies the dispersion of the belief distribution, thereby enabling the measurement of the target node's (i.e., flood risk) sensitivity to alterations in model outputs or parameters. Consequently, a greater belief difference derived from the CPTs corresponds to a heightened sensitivity of the target node to the respective parameter changes. By using CPT sensitivity analysis, a refined and spatially explicit flood risk assessment of the Xina Chuan River Basin in Qinghai Province is achieved, comprehensively characterizing the complex interactions among hazard, exposure, and vulnerability.

3. Application of Delphi-BNs

3.1. Description of the Study Area

The Xina Chuan River represents a first-order tributary along the left bank of the Huangshui River. It is located in Hai Yan and Huang Zhong counties in eastern Qinghai Province and is also known as the Shuixia River (Figure 3). The river originates 2 km northwest of Hongshanzhang in the eastern part of Haiyan County. The source elevation is 4039 m above sea level. The main stream flows from northwest to southeast. The section upstream of the Shuixia Gorge outlet is called the Shuixia River, and the section downstream is called the Xina Chuan River. It flows into the Huangshui River at Gaolengkan Village, Huangzhong County. The Xina Chuan River has a length of 88 km and a basin area of 957 km², of which 790.72 km² is located within the administrative boundaries of Huangzhong County. The elevation at the mouth is 2353 m, with a total drop of 1686 m and an average channel slope of 2.1‰. The river alternates between gorges and valleys, with a

channel width of about 20 m and a gravel bed. In the upper reaches, there are extensive swamps and grasslands; the middle reaches are characterized by rolling mountains and dense forests; the lower reaches consist of intermountain basins. The drainage system is well developed, with a dense network of gullies. Runoff is mainly supplied by precipitation. From the river mouth to its source, the mean annual precipitation varies between 400 and 650 mm. Additionally, the mean annual flow rate is recorded at $5.18 \text{ m}^3/\text{s}$, corresponding to a mean annual runoff of 163 million m^3 .

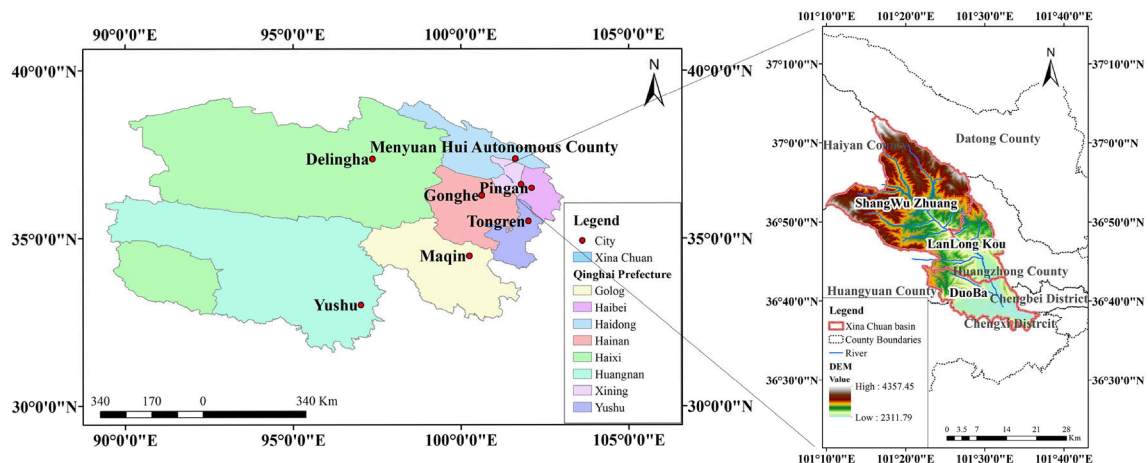


Figure 3. Location diagram of Xina Chuan River, Xining, Qinghai.

In recent years, extreme rainfall events have become frequent, with particularly high concentrations during July and August. The heightened flood risk posed by this trend to the villages along the Xina Chuan River could lead to a range of consequences, including casualties, economic losses, as well as extensive damage to infrastructure. This study utilized data sourced from the Qinghai Provincial Department of Water Resources, which included the hydro-meteorological database, census records, remote sensing imagery, and the Xina Chuan River channel management report. The study area was discretized into $30 \text{ m} \times 30 \text{ m}$ grid cells, yielding a total of 6,805,893 cells. This grid-based discretization improves spatial resolution and enables a thorough evaluation of flood risk across individual administrative regions.

3.2. Hydrodynamic Model Simulation Results

3.2.1. Construction of the One-Dimensional (1D) Model

The one-dimensional (1D) model of the Xina Chuan River includes the river's main channel along with its two principal tributaries. The total length of the modeled river network is 36 km, and the average channel slope is 1.7% . In this project, 11 cross-sections were measured in the field. Combined with high-precision DEM data, an additional 122 cross-sections were interpolated at an average interval of 250 m, resulting in a total of 133 river cross-sections, including 6 water-related structures.

3.2.2. Construction of the Two-Dimensional (2D) Model

The Mesh2D software (v 4.0.2) was used to generate the mesh for the computational domain of Xina Chuan River. The average mesh size near the river channel and around key protected towns and villages was controlled at approximately 10 m, while the average mesh size for other areas was about 60 m. A total of 37,151 mesh elements and 40,629 nodes were generated. Grid elevation interpolation was performed using high-resolution DEM elevation data (5 m resolution).

3.2.3. Construction of the 1D-2D Coupled Model

Leveraging the embankments along both banks of the Xina Chuan River, a lateral coupling scheme was implemented to link the 1D hydrodynamic model with the 2D overland flow model. This approach achieves complete integration of the main stream and its tributaries with the adjacent two-dimensional overland surface.

3.2.4. Model Parameter Determination and Calibration

Based on field investigations (Figure 4) and relevant design data, the roughness coefficient of the 1D river channel was in the range of 0.022~0.04. In the 2D model, surface grid cells were assigned roughness coefficients corresponding to their respective land use types. Additionally, the processing of grid cells was guided by the flood protection standards applicable to existing embankment sections.



Figure 4. Xina Chuan River channel site photographs: (a) Xina Chuan River upstream; (b) Xina Chuan River downstream.

Two typical floods were selected for model calibration in the Xina Chuan River basin: the flood event of 22 August 2013 and the flood event of 21 August 2022. The measured flood of 22 August 2013 was used for model validation. The main model parameter is the roughness coefficient. Based on the calibration results, the Manning's roughness coefficient for the main channel of the Xina Chuan River was determined to be in the range of 0.022–0.04, whereas the floodplain roughness coefficient was found to vary between 0.025 and 0.45. Roughness values were assigned to the surface grid cells of the 2D model in accordance with land use types, and the cells were subsequently processed following the flood protection standards of the existing embankment sections. The results demonstrate that at the downstream hydrometric station (i.e., the Xina Chuan River (II) station), the relative difference in peak discharge is within 10%, and the maximum discrepancy in water level is no greater than 20 cm. The model results show strong agreement with the design values, demonstrating that the model possesses satisfactory accuracy and fulfills the precision criteria required for hydrodynamic modeling [37]. Therefore, the model can be applied to subsequent flood control engineering planning.

3.2.5. Results of Model Simulation

The 100-year flood scenario was simulated for the Xina Chuan River flood-prone area (Figure 5). The total inundation extent was 12.25 km², with water depth distribution as follows: depths below 0.3 m accounted for the largest proportion at 7.98 km², followed by 0.3–0.5 m (1.64 km²), 0.5–1.0 m (1.49 km²), 1.0–2.0 m (0.66 km²), and >2.0 m (0.48 km²).

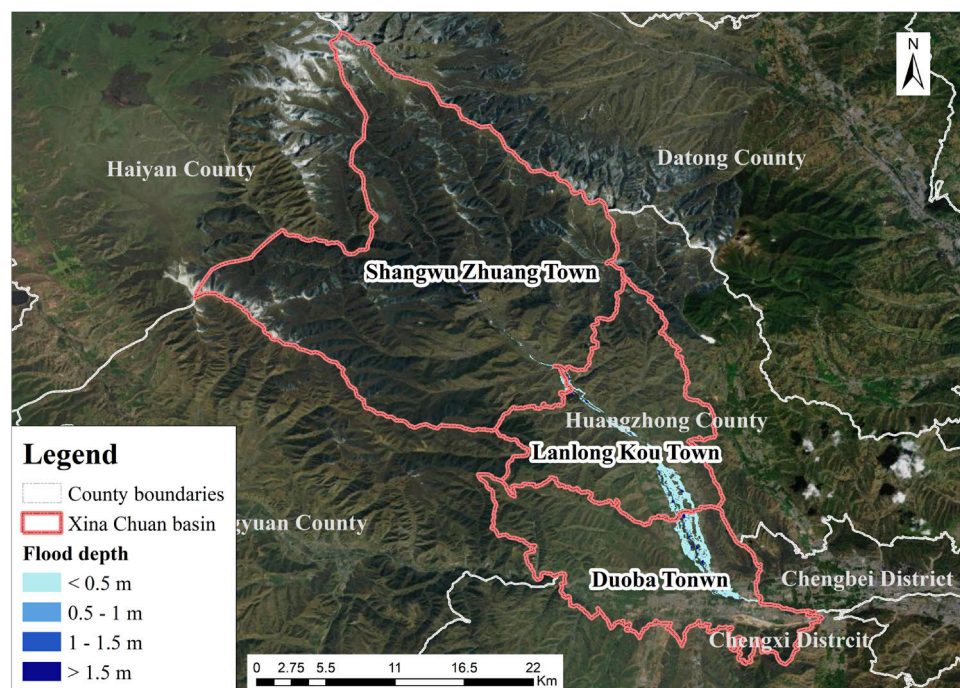


Figure 5. Result map of the inundation extent of Xina Chuan River. The remote sensing images are sourced from the National Geographic Information Public Service Platform of China, Tianditu (Map World). For more information, visit <http://lbs.tianditu.gov.cn/server/MapService.html> (accessed on 17 August 2026).

3.3. Flood Risk Assessment Framework

This study makes a key contribution by developing an integrated Delphi–BN framework for flood risk assessment (Figure 6). This framework was implemented in the Xina Chuan River Basin, integrating three key dimensions: hazard (H), exposure (E), and vulnerability (V). The process comprises the following five steps:

1. Spatial data extraction and preprocessing.

Based on the GIS spatial database of the Xina Chuan River basin, multiple types of spatial data are systematically extracted, including topography, meteorology and hydrology, land use, population distribution, and transportation networks. Through operations such as standardization, outlier removal, and data normalization, the various spatial data are categorized and organized to form a standardized dataset that meets the assessment requirements, laying a solid data foundation for Delphi expert scoring, BNs model construction, and GIS spatial analysis.

2. Delphi scoring and indicator system optimization.

A panel of experts from diverse disciplines was convened to assess and assign scores to both the selection of assessment indicators and the determination of indicator weights, grounded in the formation mechanisms of flood risk in small and medium-sized rivers. Through the Delphi method, the assessment indicator system covering the three dimensions of hazard, exposure, and vulnerability is finalized. At the same time, based on the expert scoring results, the causal relationships among nodes in the BNs model and the initial assignment of conditional probability tables (CPTs) are preliminarily determined, thereby optimizing the model structure and enhancing the scientific rigor and adaptability of the model.

3. Delphi–BN model construction and probabilistic inference.

Using the Delphi scoring results, a targeted Bayesian network model is constructed, clarifying the causal relationships among the assessment indicators. Case-based learning techniques are employed to refine the CPT assignments (correcting subjective biases in expert scoring). Subsequently, Bayesian network probabilistic inference is applied. It quantifies the internal relationships among flood hazard, regional vulnerability, and exposure parameters. As a result, the probabilistic distribution characteristics of flood risk are obtained, enabling a quantitative description of flood risk for small and medium rivers.

4. Flood risk map generation and visualization.

The flood risk probability results calculated by the BNs were integrated with watershed spatial background data. Using GIS spatial visualization techniques, a flood risk level map for the Xina Chuan River Basin was then generated. Providing intuitive visual support for flood risk identification and control in small and medium rivers, this map clearly shows the zonation of four risk levels: low, moderate, high, and extremely high.

5. Sensitivity analysis.

Sensitivity analysis of the flood risk map is conducted using belief variance, incorporating three indicators: Node Mutual Information, Proportion, and Belief Difference. These form a complete sensitivity analysis framework. Node Mutual Information defines the extent to which one node (e.g., the node representing flood risk) is dependent on another node, thereby characterizing their statistical interdependence. Proportion, within belief function theory, refers to the “focal elements” and their basic probability assignments (BPA, also known as mass functions) in evidence theory. It reflects the probability that all current evidence supports the conclusion of “flood risk.” Belief Difference is the most important output of sensitivity analysis. It measures the degree of fluctuation in the final belief value when small changes occur in the input parameters.

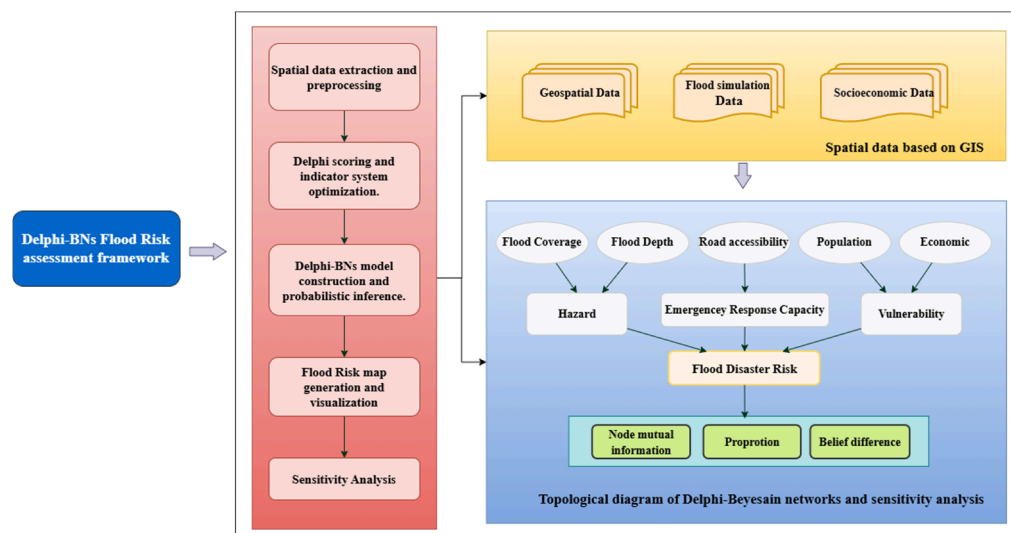


Figure 6. A Delphi-Bayesian network probabilistic framework for flood risk assessment.

3.3.1. Flood Risk Assessment Indicators

Informed by the flood risk formation mechanisms in the Xina Chuan River basin and the needs of model construction, this study proposes a scientific and systematic indicator system for flood risk assessment, structured around the three core pillars of hazard (H), exposure (E), and vulnerability (V). The selection of assessment dimensions, specific indicators, classification, and description for each dimension are shown in Table 3.

Table 3. Assessment framework and classification criteria for flood risk indicators.

Variable	Classification	Description
Flood coverage	Uninundated/inundated area	Binary classification
Flood depth	Low/Moderate/High/Extremely High	Depth ranges: <0.5 m/0.5–1 m/ 1–1.5 m/>1.5 m
Road accessibility	Inaccessible/Accessible	Based on road elevation and flood depth
Population vulnerability	Low/Moderate/High/Extremely High	Based on population density
GDP vulnerability	Low/Moderate/High/Extremely High	Based on economic output per unit area
Hazard	Low/Moderate/High/Extremely High	Hazard is the child node of flood coverage and flood depth
Exposure	Low/Moderate/High/Extremely High	Exposure is determined by road accessibility
Vulnerability	Low/Moderate/High/Extremely High	Vulnerability is determined by population and GDP vulnerability
Flood Risk	Low/Moderate/High/Extremely High	Flood risk is the child node of hazard, exposure, and vulnerability

1. Hazard Factor Indicators.

The primary objective of flood hazard assessment is to provide a quantitative measure of the intrinsic intensity of flood hazards. The assessment outcome is primarily determined by two key indicators: flood inundation extent and inundation depth, which jointly reflect the impact range and destructive potential of a flood event. To accurately obtain the data for these two indicators, this study uses the rational formula method from the hydrological manual, combined with basic data such as topography, meteorology, and soil of the Xina Chuan River Basin, to calculate the runoff corresponding to the basin area. Based on these calculations, the 100-year design flood data were derived from the corresponding 100-year design storm parameters. These data served as the primary input to the hydrodynamic model for simulating the spatial distribution of inundation extent and depth across the Xina Chuan River Basin. The hydrodynamic simulation employs IFMS (v 4.0.2), a hydrodynamic model, which has high accuracy and stability for flood simulation in small and medium rivers and can accurately reproduce the flood propagation process, providing reliable support for the quantification of hazard indicators. The hydrodynamic model was forced by a 100-year return period flood event, and the simulation results were calibrated and validated against observed flood records. The resulting flood depth and inundation extent were extracted for each grid cell (30 m × 30 m) and subsequently classified into four hazard levels (low, moderate, high, extremely high) using the natural breaks scheme in ArcGIS.

2. Exposure Factor Indicators.

Exposure assessment focuses on the exposure characteristics of the elements at risk that may be affected by flood disasters. In this study, road accessibility is selected as the core exposure indicator because it is directly related to the feasibility and efficiency of emergency rescue during floods and is an important factor affecting the actual impact level of flood risk. Studies [53,54] have provided a mature analytical framework for understanding road accessibility features and their role in influencing flood risk, providing a theoretical reference for the selection and analysis of this indicator. Relevant research confirms that the level of road accessibility significantly affects the timeliness and effectiveness of flood rescue efforts [55]. Road data used in this study were derived from the Third National Land Survey data provided by the Department of Natural Resources of Qinghai Province. Road accessibility was assessed based on a flood depth threshold, where depths exceeding 0.3 m were considered impassable. For the current situation of the Xina Chuan River basin, this study takes the entire basin as the study area, uses a 30 m × 30 m grid resolution,

and rasterizes all roads within the basin, clearly distinguishing between passable and impassable grid cells. These elements are adopted into the flood risk assessment framework, serving to quantify the exposure level of the road system during flood disasters. To ensure standardization and consistency of the assessment, all exposure indicators are categorized into four tiers (low, moderate, high, and extremely high). The categorization is performed with the natural breaks scheme available in ArcGIS, guided by the observed numerical distribution of the indicators. This yields standardized data for the subsequent probabilistic assessment.

3. Vulnerability Factor Indicators.

Vulnerability assessment mainly reflects the capacity and susceptibility of the elements at risk within the basin to withstand flood disasters. This study selects indicators from two core dimensions, namely population vulnerability and economic vulnerability, to comprehensively quantify the vulnerability level of regional flood risk. Population density and GDP spatial distribution data were obtained from the Qinghai Provincial Statistical Yearbook (2024) and spatially interpolated to a $30\text{ m} \times 30\text{ m}$ grid resolution. Population vulnerability is primarily assessed by the spatial distribution of population density. A positive correlation exists between population density and disaster losses: denser populations lead to more potential casualties and a wider impact range during flood disasters [56]. Economic vulnerability is primarily assessed by the spatial distribution of GDP: the higher the GDP level, the denser the economic assets such as regional industries and infrastructure, and the more substantial the direct and indirect economic losses attributable to flood events. Consistent with the exposure indicators, all vulnerability indicators are categorized into four levels (low, moderate, high, and extremely high). This categorization is performed via the natural breaks algorithm within ArcGIS, based on the natural distribution patterns of the indicator values. This guarantees both the scientific rigor of indicator quantification and the uniformity of assessment criteria, thereby supplying standardized input parameters for probabilistic inference based on a Bayesian network.

3.3.2. A Delphi–BN Approach to Flood Risk Quantification

The core idea of applying the Delphi–BNs flood risk assessment method is to utilize the Delphi method to populate the Conditional Probability Tables (CPTs) corresponding to all nodes in the Bayesian network. Using a complete node system of flood risk influencing factors, the parent–child relationships and topological structure of each node are clarified. An expert panel was assembled to score the CPTs of each node.

A panel of 18 experts was invited to participate in the Delphi survey. The panel comprised six full professors, five associate professors, four senior engineers, and three researchers from universities, research institutes, and government agencies, specializing in hydrology, meteorology, emergency management, water resources engineering, and urban planning. Each expert possessed a minimum of 10 years of professional experience in flood risk-related fields, ensuring a high level of domain knowledge and practical expertise.

The Delphi process was conducted over three rounds. In the first round, experts were asked to independently evaluate the candidate indicators and provide open-ended suggestions for refinement. In the second round, the aggregated statistical feedback from the first round was provided to each expert, who was then requested to re-evaluate their judgments. The third round followed the same procedure to further narrow the divergence of opinions. The response rates for the three rounds were 94.4% (17/18), 88.9% (16/18), and 94.4% (17/18), respectively, indicating a consistently high level of expert engagement throughout the process.

To synthesize expert opinions, the median and interquartile range (IQR) were calculated for each indicator across all expert responses in each round. The median was used

as the central tendency measure to represent the collective judgment of the panel, while the IQR served as an indicator of the dispersion of opinions. The coefficient of variation (C_v) was additionally computed as a supplementary measure of consensus. The consensus criterion was predefined as an $IQR \leq 1.0$ on a 5-point Likert scale and a $C_v \leq 0.25$, indicating that the majority of expert ratings were closely concentrated around the median value. Indicators meeting these criteria after the third round were retained in the final assessment framework, while those failing to reach consensus were either revised or eliminated based on expert feedback. Table 4 summarizes the expert-derived weights for all assessment indicators after three Delphi rounds.

Table 4. Summary of expert-derived weights for assessment indicators.

Indicator Layer	Sub-Indicator	Median Weight	IQR	C_v	Consensus Achieved
Hazard	Flood Depth	0.38	0.09	0.11	Yes
Hazard	Flood Coverage	0.22	0.10	0.13	Yes
Vulnerability	Access Road	0.18	0.12	0.14	Yes
Exposure	Population	0.12	0.11	0.15	Yes
Exposure	Economic	0.10	0.10	0.16	Yes
Total		1.00	—	—	—

Based on the finalized expert consensus, the assessment indicator system covering the three dimensions of hazard, exposure, and vulnerability was established. The qualitative structure of the BN was established through literature review and expert consultation. This structure is a directed acyclic graph (DAG) that defines the causal dependencies among variables. Following the hazard-exposure-vulnerability framework, the target node (Flood Disaster Risk) was decomposed into three intermediate nodes: Hazard, Exposure, and Vulnerability. Each intermediate node was then linked to its respective leaf nodes based on documented causal relationships and validated through the Delphi panel (Figure 7). The resulting DAG encodes the conditional independence assumptions that allow the joint probability distribution over all nodes to be factorized as shown in Equation (11).

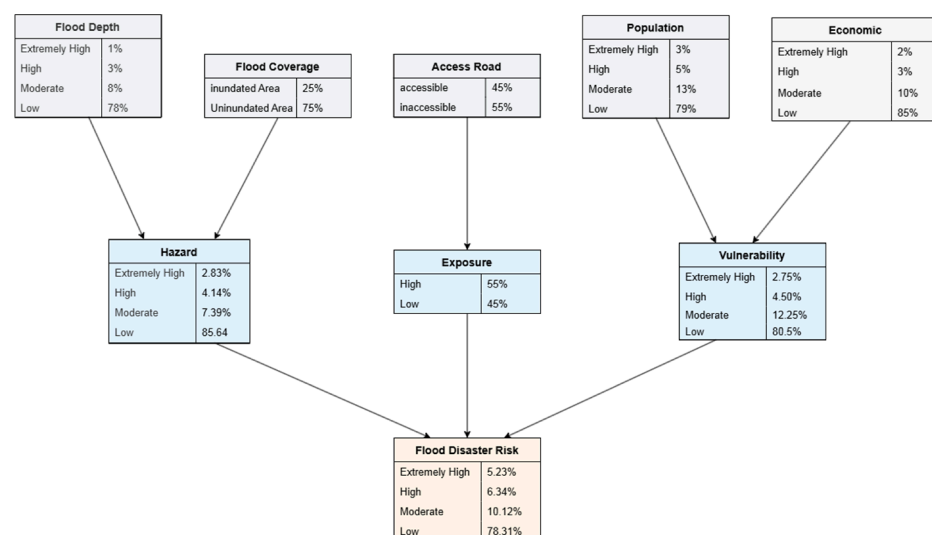


Figure 7. Topology diagram and conditional probability tables (CPTs) of the Delphi-BN model for flood risk assessment.

The expert consensus scores were transformed into Conditional Probability Tables (CPTs) through a three-step procedure. First, for each parent node (indicator), the expert

panel assessed two parameters: the weight w_i , representing the relative importance of the indicator to flood risk, and the state contribution coefficient c_{ij} , representing the degree to which the j state of indicator i contributes to flood risk. Second, for each possible combination of parent node states, an aggregated severity score S was computed as the weighted sum:

$$S = \sum_{i=1}^n w_i \times c_{ij} \quad (18)$$

where n is the number of parent nodes. Third, the aggregated score S was mapped to the conditional probability distribution over the child node states using a predefined piecewise linear mapping function. The mapping thresholds were established in alignment with the indicator discretization criteria and validated through expert consultation. Table 5 presents the state contribution coefficients for each indicator, which quantify the degree to which a given state of an indicator contributes to overall flood risk. For instance, flood depth in the “extremely high” state is assigned a coefficient of 1.00, indicating the maximum contribution, while “low” is assigned 0.10, reflecting a minimal contribution. Binary indicators such as flood coverage and access road are assigned coefficients of 1.00 for the more severe state and 0.00 for the less severe state.

Table 5. State contribution coefficients for each indicator.

Indicator	State	State Contribution Coefficient c_{ij}
Flood Depth	Extremely High	1.00
	High	0.75
	Moderate	0.45
	Low	0.10
Flood Coverage	Inundated	1.00
	Uninundated	0.00
Population	Extremely High	1.00
	High	0.70
	Moderate	0.40
	Low	0.10
Economic	Extremely High	1.00
	High	0.65
	Moderate	0.35
	Low	0.05
Access Road	Inaccessible	1.00
	Accessible	0.00

Using the expert-derived weights (Table 4) and state contribution coefficients (Table 5), the aggregated severity score S was computed for each possible parent–state combination as the weighted sum S . Table 6 provides four illustrative examples of S calculations for selected parent–state combinations, ranging from a low-risk scenario to an extreme-risk scenario. These examples demonstrate the monotonic increase in the aggregated score as indicator states become more severe, confirming the logical consistency of the aggregation scheme.

The final step involved mapping each aggregated score S to a conditional probability distribution over the four risk states (low, moderate, high, and extremely high) using the piecewise linear function defined in Table 7. For example, an S value of 0.073 (Scenario 1) falls within the interval $0.00 \leq S < 0.15$ and is therefore mapped to probabilities of 0.85 (low), 0.10 (moderate), 0.03 (high), and 0.02 (extremely high). Conversely, an S value of 1.000 (Scenario 4) falls within $0.85 \leq S \leq 1.00$ and is mapped to 0.02 (low), 0.08 (moderate), 0.25 (high), and 0.65 (extremely high). This monotonic mapping ensures that higher aggregated

severity scores consistently translate into higher probabilities of more severe risk states, thereby preserving the intuitive interpretation of the risk continuum.

Table 6. Illustrative examples of aggregated severity scores and corresponding CPT probabilities for selected parent–state combinations.

Scenario	Flood Depth	Flood Coverage	Population	Economic	Access Road	Aggregated Severity Scores S
1	Low (0.10)	Uninundated (0)	Low (0.10)	Low (0.05)	Accessible (0)	0.073
2	Moderate (0.45)	Inundated (1.0)	Moderate (0.40)	Moderate (0.35)	Accessible (0)	0.494
3	High (0.75)	Inundated (1.0)	High (0.70)	High (0.65)	Inaccessible (1.0)	0.834
4	Extremely High (1.0)	Inundated (1.0)	Extremely High (1.0)	Extremely High (1.0)	Inaccessible (1.0)	1

Table 7. Piecewise linear probability mapping function.

Aggregated Severity Scores S Interval	Risk—Low	Risk—Moderate	Risk—High	Risk—Extremely High
$0.00 \leq S < 0.15$	0.85	0.10	0.03	0.02
$0.15 \leq S < 0.40$	0.55	0.30	0.10	0.05
$0.40 \leq S < 0.65$	0.25	0.45	0.20	0.10
$0.65 \leq S < 0.85$	0.10	0.25	0.40	0.25
$0.85 \leq S \leq 1.00$	0.02	0.08	0.25	0.65

After the CPTs for all nodes were populated using the above mapping procedure, the Bayesian network was applied to each spatial unit in the study area. For every grid cell, the BN performed probabilistic inference based on the observed states of the leaf nodes, yielding a posterior probability distribution over the flood disaster risk states. By aggregating these cell-specific risk probability distributions across the entire study area, the marginal probability of each risk state at the network level, the overall percentage of cells classified into each risk category, was obtained. These aggregated probabilities correspond to the node probability summaries displayed in the BN model.

Concurrently, the causal relationships among nodes in the BN model and the initial assignments of conditional probability tables (CPTs) were preliminarily determined (Figure 7).

4. Results

4.1. Model Calibration and Validation Results

4.1.1. Hydrological Calibration and Validation

The Xina Chuan hydrological station, situated near the outlet of the Xina Chuan River basin, has collected continuous discharge observations over the period 1957–2023. A total of 15 flood events from 1957 to 2010 were used for model calibration, while three flood events with diverse magnitudes between 2011 and 2023 were reserved for model validation. The evaluation yielded the following results.

The performance of the hydrological model during the calibration period was assessed using three statistical indices: NSE, R_Q , and E_T , based on 15 observed flood events (Table 8). Among the 15 calibration events, the flood of 29 July 1974 was selected as the representative case for hydrograph visualization (Figure 8) because it represents both the largest flood in the calibration record and the best-performing event (NSE = 0.90, E_T = 0 h). This selection provides the most compelling visual demonstration of the model's ability to capture flood timing, peak magnitude, and overall hydrograph shape under extreme flow

conditions. The corresponding rainfall hyetograph is also included to illustrate the clear rainfall–runoff response.

Table 8. Model calibration performance metrics.

Flood Event Identifier	Measured Peak Flow	Simulated Peak Flow	Nash–Sutcliffe Efficiency (NSE)	Relative Peak Flow Error (R_Q)	Timing Error (E_T)
19610809	79.3	72.6	0.81	8.4%	1
19630723	47.4	43	0.78	9.3%	2
19640721	44.5	45.5	0.87	−2.2%	0
19670802	140	146.5	0.84	−4.6%	0
19700807	91.1	82.3	0.73	9.7%	3
19710902	44.5	46.2	0.82	−3.8%	0
19740729	174	161	0.9	7.5%	0
19790810	155	140	0.76	9.7%	2
19830813	74.3	87.1	0.69	−17.2%	4
19860702	66.9	60.4	0.72	9.7%	2
19890505	41.4	37.4	0.75	9.7%	3
19900501	45.3	46.5	0.87	−2.6%	0
19940819	40.1	38.1	0.84	5.0%	1
19950906	64.2	58.7	0.89	8.6%	0
19990806	52.2	46.2	0.79	11.5%	2

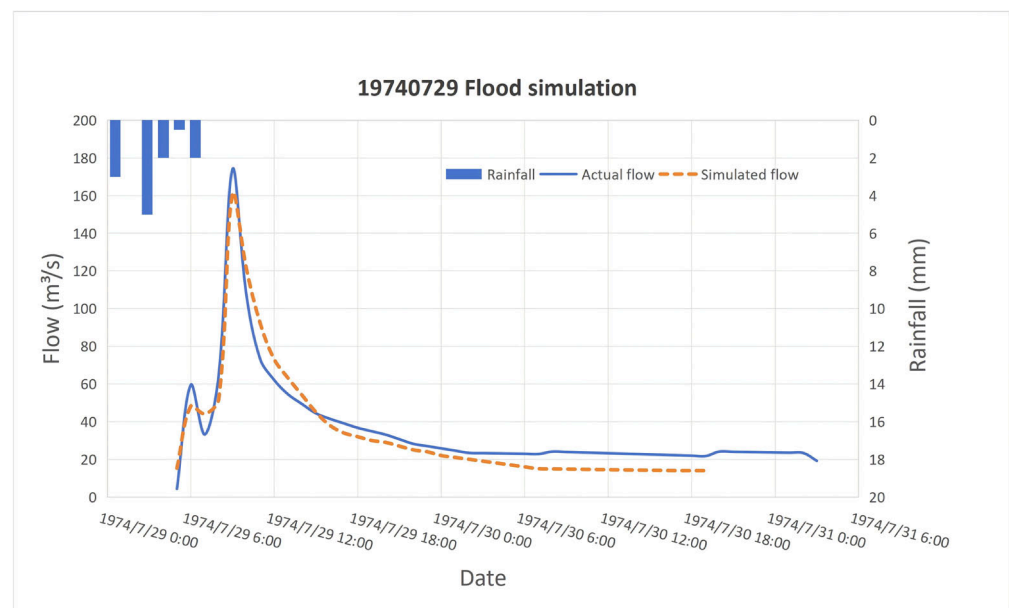


Figure 8. Observed and simulated hydrographs with rainfall hyetograph for the flood event of 29 July 1974.

During the calibration period, the NSE values ranged from 0.69 to 0.90, with a mean value of 0.80. Among the 15 events, 11 (73%) yielded $NSE > 0.75$, and 5 events exceeded 0.85, indicating that the model adequately reproduces the overall flood hydrograph dynamics. The mean absolute R_Q was 7.6%, falling within the generally acceptable range for flood simulation.

The most notable strength of the calibrated model lies in its timing accuracy. The peak flow timing error (E_T) was zero for 6 out of 15 events (40%) and within ± 1 h for 9 events (60%), with a maximum E_T of only 3 h. This demonstrates that the routing parameters (e.g., roughness coefficients and travel times) were well-optimized. However, a clear systematic bias toward peak flow underestimation was identified: 10 out of 15 events (67%) showed negative R_Q values, with five events exhibiting underestimations exceeding 9%. The largest negative deviation occurred for event 19830813 ($R_Q = -17.2\%$), which also produced the lowest NSE (0.69) and a timing delay of 3 h. This event is a clear outlier and may be attributed to unusual spatial rainfall distribution or antecedent soil moisture conditions not captured by the current model setup. Conversely, positive deviations (overestimations) were generally smaller, with a maximum of 11.5% for event 19990806.

The model performed best for the largest flood on record (event 19740729, observed peak = $174 \text{ m}^3/\text{s}$), achieving the highest NSE of 0.90 and an E_T of 0. However, for moderate-to-small flood events (observed peaks $< 100 \text{ m}^3/\text{s}$), the model tended to produce lower NSE values (averaging 0.77) and more frequent underestimations. This suggests that the model parameters may have been over-fitted to extreme flood conditions, and that the runoff generation mechanism (e.g., infiltration excess) might be non-linear across different flow regimes.

The calibrated model was validated using three independent flood events (2013–2023). The validation results show a mean NSE of 0.81, a mean absolute R_Q of 6.1%, and timing errors within 1 h for all events (Table 9). These metrics are consistent with those obtained during calibration (mean NSE = 0.80), confirming that the model is robust, well-generalized, and suitable for operational flood forecasting in the study basin.

Table 9. Model validation performance metrics.

Flood Event Identifier	Measured Peak Flow	Simulated Peak Flow	Nash–Sutcliffe Efficiency (NSE)	Relative Peak Flow Error (R_Q)	Timing Error (E_T)
20130822	62	65	0.83	−4.8%	0
20200829	44.7	41.3	0.79	7.6%	1
20220821	58.6	55.2	0.81	5.8%	0

Among the three validation events, the flood of 22 August 2013 (event 20130822) achieved the best overall performance, with an NSE of 0.83, an R_Q of −4.8%, and a perfect timing error ($E_T = 0$ h). As illustrated in Figure 9, the simulated hydrograph closely tracks the observed flow throughout the entire storm event. The rising limb is well captured, with the model responding promptly to the rainfall pulse that occurred around midnight on 22 August. The peak flow is simulated as $65 \text{ m}^3/\text{s}$, which slightly overestimates the observed peak of $62 \text{ m}^3/\text{s}$ by 4.8%, a minor deviation well within the acceptable tolerance for flood forecasting. The recession limb is also reproduced with reasonable accuracy, though the simulated flow appears to decline slightly more slowly than the observed flow during the falling stage. The rainfall hyetograph shows a concentrated storm pattern with multiple pulses, and the model successfully captures the dual-peak response, demonstrating its capability to simulate complex rainfall–runoff dynamics. The perfect timing match ($E_T = 0$ h) further confirms that the flow routing parameters remain robust when applied to independent events. Overall, this event provides strong evidence that the calibrated model maintains its predictive skill under validation conditions.

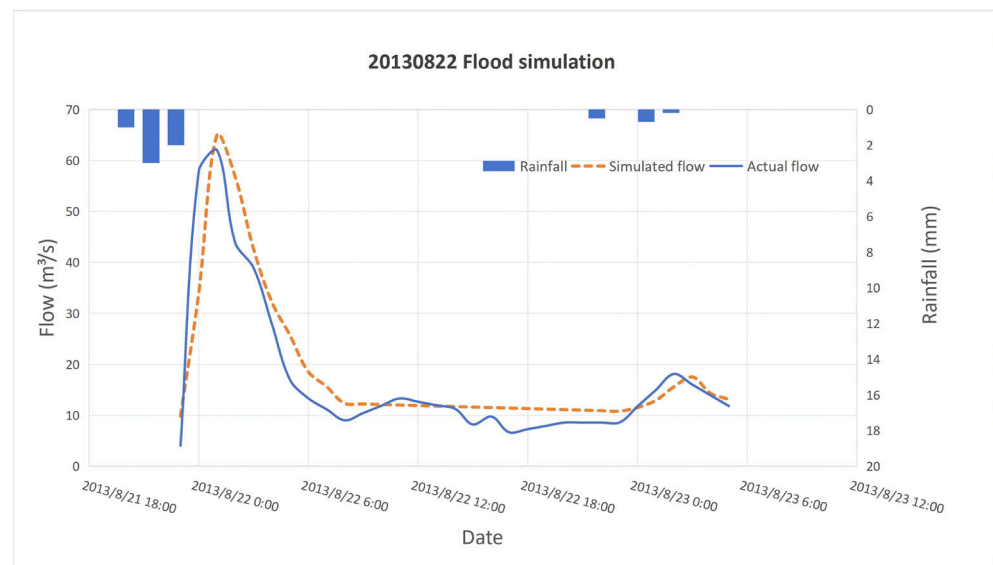


Figure 9. Observed and simulated hydrographs with rainfall hyetograph for the flood event of 22 August 2013.

4.1.2. Validation of Flood-Risk Extent Using Observed Data

Since wall-to-wall observed flood maps are not available for the study area, we validated the reliability of the modeled flood-risk map using field-surveyed high-water marks (HWMs) from two independent flood events. Specifically, HWMs were collected in Lanlong Kou village for the events of 22 August 2013, and 21 August 2022.

The comparison between simulated and observed water levels shows excellent agreement: for the 2013 event, the simulated water level (2474.25 m) deviates from the observed value (2474.19 m) by only 0.06 m; for the 2022 event, the simulated level (2474.20 m) reproduces the observed level (2473.90 m) with a bias of 0.30 m. These discrepancies are well within the acceptable margin of uncertainty for practical flood risk assessment, confirming that our model reliably captures the inundation dynamics at the local scale.

4.2. Hazard, Exposure, and Vulnerability Assessment Results

- **Hazard:** Based on modeling with IFMS (v 4.0.2), the inundation extent and inundation depth of the Xina Chuan River were obtained. Reclassification of the Digital Elevation Model (DEM) data for the watershed enabled the isolation of all areas where elevations fall below the established water level (H), starting from the lowest point. The resulting flood-prone zone is subsequently mapped and rendered as a spatial visualization (Figure 10).
- **Exposure:** This exposure assessment considers road accessibility. The flood risk of an area depends not only on the depth and velocity of inundation but also directly on whether people can evacuate in time during a flood and whether rescue forces can reach the area smoothly [56]. In other words, the vulnerability of the road system directly amplifies flood risk. Based on GIS, the road vector data were processed into road grid data with a 30 m × 30 m resolution (Figure 11).
- **Vulnerability:** The vulnerability assessment for flood disaster risk is based on two key factors: the vulnerability of populations and that of economies. The spatial distribution of population density is a key determinant of population vulnerability. Hence, areas with higher population density tend to suffer greater losses during floods and therefore exhibit higher vulnerability. Economic vulnerability is determined by the distribution of GDP; the higher the GDP density in an area, the greater the loss caused by a flood,

and these regions face heightened risk of severe losses during flood disasters. The geographic patterns of these vulnerabilities are displayed in Figure 12.

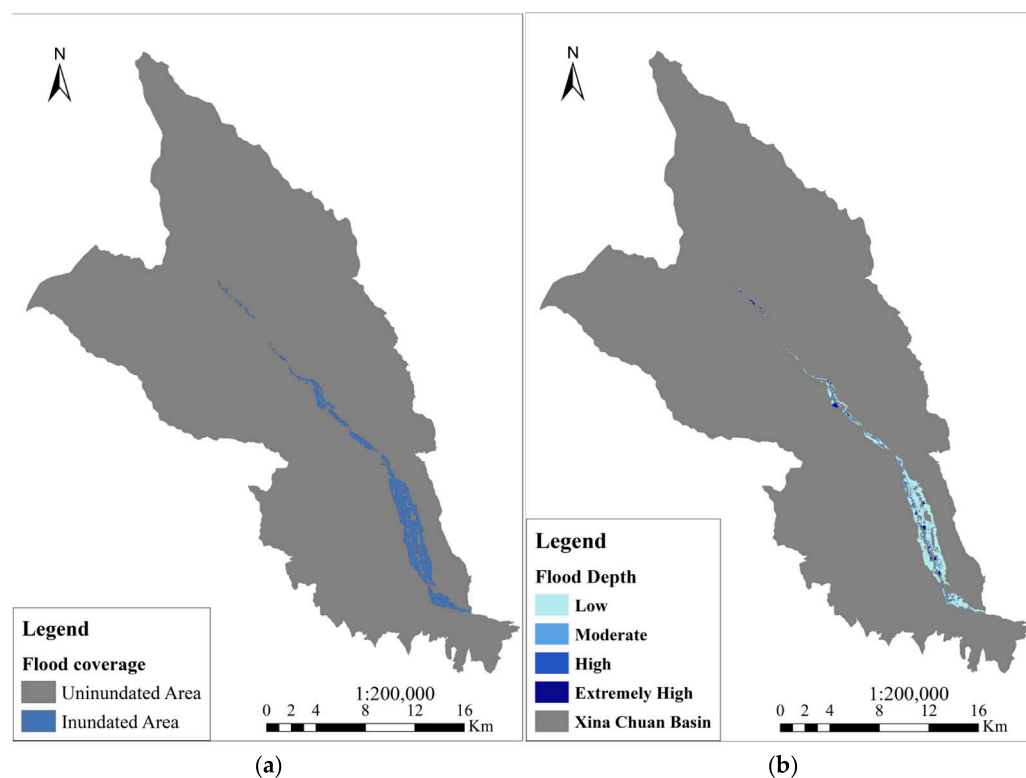


Figure 10. Flood risk and extent under the 100-year return period rainfall scenario in Xina Chuan River basin: (a) inundation area; (b) flood depth.

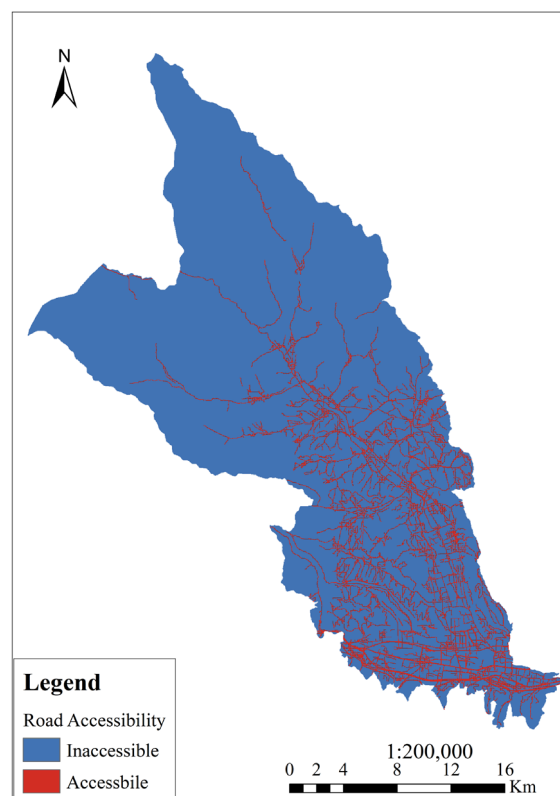


Figure 11. Spatial distribution of road accessibility exposure in the study area.

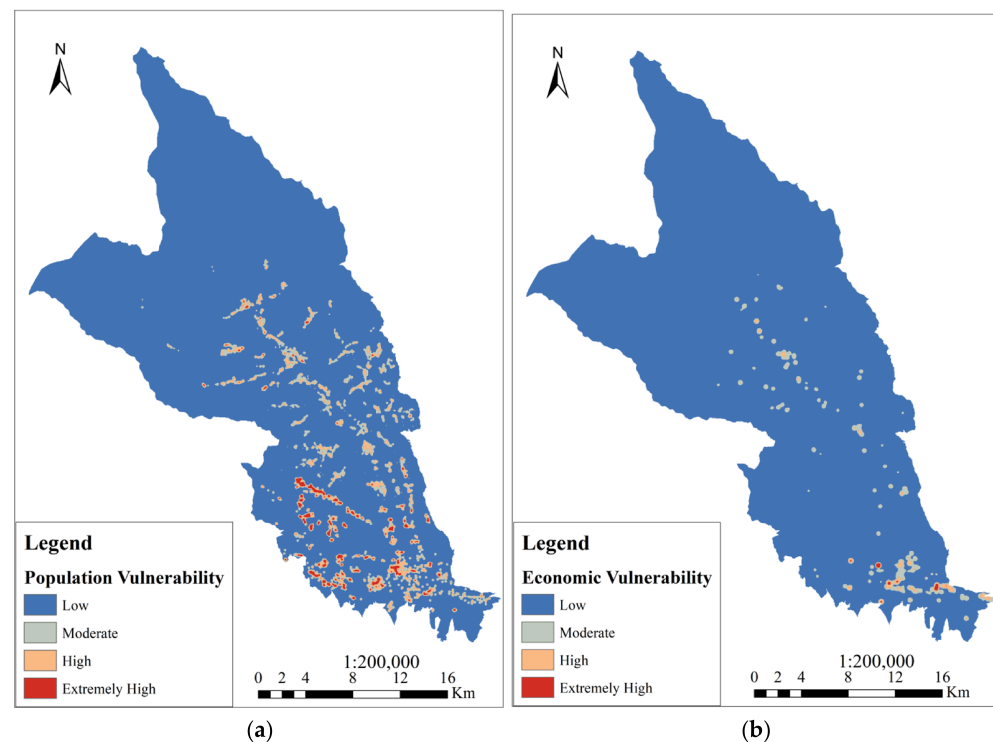


Figure 12. Spatial distribution of vulnerability: (a) population vulnerability; (b) economic vulnerability.

4.3. Assessment Results of Flood Risk Using the Delphi–BN Framework

The Delphi–BN framework used to characterize flood risk is shown in Figure 13. The Delphi–BN method served to construct the Conditional Probability Tables (CPTs) for quantifying and mapping the flood risk of Xina Chuan River in the Huangshui River Basin, Qinghai Province. The CPTs of flood risk reveal the following: the lowest risk level accounts for 78.31% of the total area. This suggests that, under the 100-year rainfall scenario, most of the Xina Chuan River region faces relatively low flood hazard. The moderate risk level accounts for 10.12%, followed by the high risk level at 6.34%, and the extremely high risk level at 5.23%. Upon comprehensive evaluation of the integrated components of hazard, vulnerability, and exposure, the overall flood risk node demonstrates a high probability of occupying the low-risk level, signifying a comparatively low-risk condition. However, proximity to mountainous regions is generally associated with higher flood risk. Several high-risk points have been detected in city center areas of higher socioeconomic vulnerability, a number of which are principally distributed along riverbanks (Figure 13). Compared to other areas, the southern urban area of Xina Chuan River exhibits markedly higher flood risk.

The flood risk results derived from the Delphi–Bayesian Network (Delphi–BN) framework comprehensively consider flood hazard, exposure, and vulnerability, thereby identifying the areas in need of rescue in a more targeted manner. Based on GIS, the very-high-risk zones are identified as the areas most in need of rescue, i.e., the first batch of evacuation areas; high-risk zones are designated as the second batch of rescue areas; and medium-risk zones are designated as the third batch of rescue areas (Figure 13). This study relies on the Delphi–Bayesian Network (Delphi–BN) method to construct a systematic and dynamic flood risk assessment framework. Focusing on the core elements of flood hazard formation, this framework integrates expert experience with spatial big data, breaking through the limitations of traditional assessment methods to achieve a refined quantitative assessment of regional flood risk. This provides a core basis for subsequent emergency decision-making tasks such as emergency rescue, personnel evacuation, and resource dispatch.

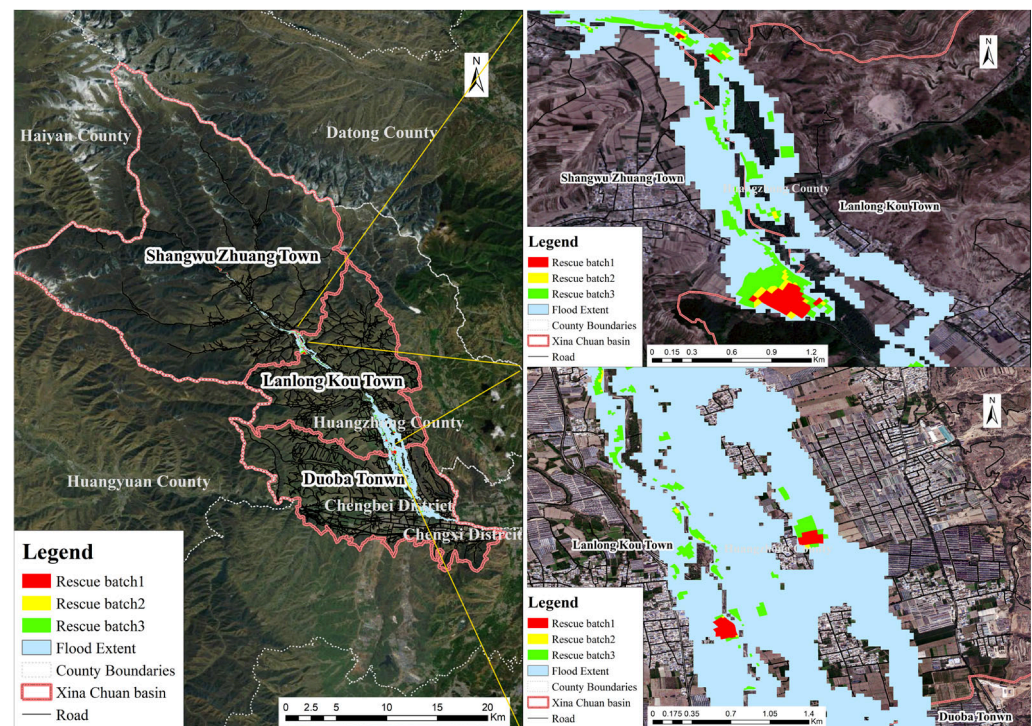


Figure 13. Flood risk assessment results based on the Delphi–Bayesian Network (Delphi–BN) in Xina Chuan Basin, Qinghai. The remote sensing images are sourced from the National Geographic Information Public Service Platform of China, Tianditu (Map World). For more information, visit <http://lbs.tianditu.gov.cn/server/MapService.html> (accessed on 17 August 2026).

To bridge the gap between hazard assessment and emergency operations, we aggregated flood-exposed elements into four risk tiers aligned with the planned rescue sequencing. Table 10 summarizes the key decision-relevant indicators for each category, including affected population, GDP, road length, and village counts.

Table 10. Decision-relevant metrics by risk category.

Indicator	Unit	First Rescue Batch (Extremely High)	Second Rescue Batch (High)	Third Rescue Batch (Moderate)	Low Risk	Total
Total flooded area	km ²	1.84	2.44	3.80	4.17	12.25
Flooded residential land	10,000 m ²	18.67	22.41	22.41	11.21	74.70
Flooded residential buildings	units	80	96	96	49	321
Flooded farmland	ha	98.61	141.75	135.70	240.23	616.29
Flooded roads	km	12.26	11.73	11.63	1.03	36.65
Affected population	persons	6750	6100	4200	450	17,500
Affected GDP	10,000 CNY	4600	4250	3200	1089.71	13,139.71
Direct economic loss	10,000 CNY	689.60	769.60	636.10	558.72	2654.03
Affected villages	units	8	10	10	5	33

Note(s): The “low-risk” category includes flooded tiles with minimal expected impact. These areas are not assigned to a specific rescue batch, as they do not require immediate response, but they remain within the flooded zone and are monitored for potential deterioration.

The tiered pattern reveals a clear inverse relationship between hazard priority and spatial extent. The first rescue batch (extremely high risk) covers only 1.84 km², the smallest among all categories, but it concentrates the largest affected population (6750 persons) and the highest GDP exposure (4600×10^4 CNY). This confirms that emergency response should prioritize population density and economic value over areal coverage, as the first

batch exhibits a population density of 3668 persons/km², compared to 2500 persons/km² for the second batch and 108 persons/km² for the low-risk zone.

The second and third rescue batches present a contrasting exposure profile. Although their population and GDP figures are moderately lower (6100 and 4200 persons; 4250 and 3200×10^4 CNY, respectively), they involve substantially longer flooded road sections (11.73 km and 11.63 km) and affect ten villages each, which are the highest village counts among all batches. This indicates that rescue operations in these tiers may face greater logistical constraints due to disrupted transportation networks, and that coordination across multiple rural settlements will be critical.

For the low-risk category, the dominant exposure shifts toward agricultural assets. This zone contains the largest flooded farmland (240.23 ha) and accounts for the second-highest direct economic loss (558.72×10^4 CNY), primarily driven by crop damage rather than infrastructure or population losses. Notably, only 450 people and 5 villages are affected in this category, supporting its exclusion from the immediate rescue schedule. Nevertheless, the sheer farmland area (representing 39% of the total flooded farmland) warrants continued monitoring, as prolonged inundation could amplify secondary losses or trigger cascading effects on downstream food supply chains.

Cumulatively, the flood event affects 33 villages and 17,500 people across the entire study area, with total direct economic losses reaching 2654.03×10^4 CNY. These decision-relevant metrics not only justify the proposed three-batch rescue strategy but also provide operational targets, including prioritizing road clearance in the second and third batches and deploying agricultural drainage teams to the low-risk zone, that can be directly incorporated into emergency action plans.

4.4. Sensitivity Analysis

According to the sensitivity analysis, the relative importance of individual factors varies considerably in the context of flood risk prediction. Mutual information measures how much information a node provides to the target node (e.g., flood disaster risk), with larger values reflecting a greater influence of that node's uncertainty on the target. Proportion normalizes the mutual information into a percentage, allowing an intuitive comparison of the relative importance of each node. Flood depth (30%) and flood extent (42%) are the most sensitive nodes, as indicated by their mutual information values. Slight variations in them can cause significant changes in the assessment of flood risk. This finding corroborates the basic theory of flood risk, in which water depth and inundated area directly determine the extent of physical damage. The low mutual information for population vulnerability (0.8%) and economic vulnerability (0.5%) is attributed to the small urban area of the Xina Chuan Basin and the extensive, uninhabited upstream regions. Moreover, the area is not economically developed; most of the Xina Chuan Basin is mountainous with sluggish economic development, resulting in a lower impact of floods on these factors. The mutual information of road accessibility is also minimal at 0.4%, reflecting the basin's high road construction and design standards, which ensure that major interruptions hindering rescue access rarely happen during floods. These findings serve as a valuable reference for decision makers, enabling them to rank risk mitigation strategies and allocate resources efficiently for flood risk assessment in the Xina Chuan Basin, Qinghai Province.

5. Discussion

5.1. Model Results Analysis

A systematic validity analysis was conducted to further verify the reliability and superiority of the developed GIS-based Delphi-BN flood disaster risk assessment framework. Successfully applied to the Xina Chuan River basin in Qinghai Province, China,

this framework enables probabilistic quantification and spatial visualization of flood risk, addressing the limitation of traditional methods that fail to integrate probabilistic analysis with spatial representation. This model exhibits several notable features that enhance flood risk assessment, which are elaborated as follows.

First, it can effectively handle the nonlinear coupling relationships among assessment indicators, accurately capturing risk characteristics under complex scenarios in small and medium rivers. For example, in densely populated urban built-up areas, even when inundation depth is relatively low, the concentration of population may still result in high-risk zones due to the potential for casualties and property losses.

Second, through the Delphi–BN coupling, the model integrates multidisciplinary expert knowledge while using BNs CPTs to quantify various uncertainties (such as the randomness of meteorological precipitation, errors in topographic data, and uncertainties in human activities), reducing the interference of uncertain factors on assessment results and overcoming the strong subjectivity of traditional models.

Third, using refined flood risk assessment and GIS spatial analysis, the study area is divided into four zones: very-high-risk, high-risk, medium-risk, and low-risk. A tiered emergency rescue strategy is then developed based on zone-specific risk characteristics, damage levels, and needs. Accordingly, the extremely high-risk, high-risk, and medium-risk zones are assigned as the first, second, and third batches for rescue priority, respectively.

5.2. Rationality of Indicator Selection

The selection of the assessment indicator system in this study is scientifically sound and reasonable, strictly following the intrinsic mechanism of flood risk formation. The selected indicators comprehensively encompass the three core dimensions of flood risk, i.e., hazard, exposure, and vulnerability, thereby constituting a complete and logically consistent indicator framework. Specifically, inundation depth and inundation extent, as core hazard indicators, directly reflect the intensity and impact range of flood disasters and are key to characterizing flood hazard potential. Population density and GDP distribution, as indicators of elements at risk and socio-economic attributes, accurately reflect the regional disaster-bearing capacity and vulnerability level, quantifying the potential scale of casualties and property losses caused by floods and capturing the influence of socio-economic development on flood risk. Road accessibility, as an important exposure indicator, not only reflects the vulnerability of the transportation system during floods but is also closely related to emergency rescue efficiency and post-disaster recovery progress. Its inclusion makes the risk assessment more relevant to actual risk management needs, providing targeted support for decisions such as emergency rescue route planning and post-flood traffic restoration, thereby further enhancing the practical value of the assessment model. Future work should enhance the model by expanding the indicator system to include building type (differences in flood resistance) and flood control measures (protection standards and distribution of reservoirs and embankments). This will lead to greater accuracy and comprehensiveness in flood risk assessment, thereby strengthening technical support for watershed flood risk management.

5.3. Limitations and Improvement Directions

It must be acknowledged that the flood risk assessment framework developed in this study has two main limitations. First, the calculation of the 100-year design flood is highly dependent on the accuracy of basic data, including soil texture and land use type. Although the basic data used in this study have undergone standardized preprocessing, due to limitations in data acquisition methods and monitoring technology, small errors still exist in data such as soil moisture content and dynamic land use changes in some

areas. Such errors propagate through the model, compromising the calculation accuracy of both the design flood and the subsequent flood risk assessment results. Second, although road accessibility is a core exposure indicator currently graded according to the Xina Chuan Basin's traffic conditions and flood characteristics, the assessment remains overly focused on transportation. To further refine exposure indicators and enhance the comprehensiveness of flood risk assessment, future work should introduce additional factors, including agricultural indicators (e.g., farmland distribution, crop types, irrigation facilities) and industrial indicators (e.g., factory locations, industrial park distribution, and critical production equipment).

In response to these limitations, and considering the current practical needs and technological development trends in flood risk management, this paper proposes clear directions for future research. First, real-time hydro-meteorological monitoring data (e.g., precipitation, water level, discharge) can replace historical statistics, enabling a dynamic model that achieves real-time flood risk updating and early warning, thereby improving responsiveness to extreme flood events. Second, algorithms such as random forest and neural networks can be explored to optimize CPT learning in the Delphi-BN coupled model. By integrating Delphi expert scoring with multi-source data, subjective bias can be further reduced, CPT quantitative accuracy improved, and the model's predictive performance optimized. Third, multi-source data fusion techniques can integrate remote sensing imagery, UAV monitoring, ground measurements, and other data types, compensating for the limitations of any single data source regarding coverage, accuracy, and timeliness. This enhances the comprehensiveness and accuracy of basic data and provides more reliable data support for model optimization [57]. Through these improvements, the model structure can be further refined, assessment accuracy enhanced, and the sustainable development of flood risk assessment technology for small and medium rivers in semi-arid regions promoted.

6. Conclusions

This study takes a typical small to medium-sized river basin in the Xina Chuan River basin semi-arid region, Qinghai Province, as the research object. A GIS-based Delphi-BN model for flood disaster risk assessment in small to medium-sized rivers was constructed. Based on a 100-year flood scenario, a refined flood risk assessment was carried out, classifying flood risk into four levels: low, medium, high, and extremely high. Using the risk level results, a stratified and phased flood emergency rescue and personnel transfer plan was established. Areas of extremely high risk were designated as the primary transfer core zones for the first phase, high-risk areas as the second-phase rescue zones, and medium-risk areas as the third-phase rescue zones, providing clear operational guidance for flood emergency response in the basin.

The study systematically reveals the spatial differentiation patterns and underlying driving mechanisms of flood risk in the Xina Chuan River basin. The spatial distribution of flood risk is highly consistent with topographical features and human activity patterns. High-risk areas are mainly concentrated in river valley plains and urban built-up areas, where the terrain is flat and elevation is low, making it easy for floodwater to accumulate and difficult to drain. Combined with dense populations and concentrated infrastructure, regional disaster vulnerability is significantly increased, amplifying flood-induced losses. Medium- to high-risk areas are mostly distributed in the transitional zones between river valley plains and mountainous hilly regions, where flood risk exhibits significant transitional characteristics due to the combined effects of topographic relief, spatial variation in precipitation, and human activities. Low-risk areas are concentrated in high-altitude mountainous and hilly regions of the basin, where steeper slopes lead to rapid flood con-

centration and flow, making sustained inundation disasters unlikely, resulting in an overall low level of risk.

Factor sensitivity analysis and importance assessment indicate that inundation depth and inundation extent are the core factors determining the flood risk level in the Xina Chuan River basin, directly influencing the destructive intensity and scale of flood disasters. Population density and GDP spatial distribution are key factors regulating regional disaster vulnerability and profoundly affecting the extent of flood disaster losses. Road accessibility determines the feasibility and efficiency of flood emergency rescue and disaster response operations. The interaction and coupling of these factors ultimately shape the differentiated spatial pattern of flood risk in the basin.

Compared with traditional single-method flood risk assessment approaches, the Delphi–BN assessment framework developed in this study focuses on the core elements of flood disaster risk formation, integrates expert experience with spatial big data, and effectively overcomes the limitations of traditional methods, such as strong subjectivity, insufficient precision, and poor quantification. It achieves a refined and quantitative assessment of flood risk in small to medium-sized river basins in semi-arid regions.

The research outcomes not only clarify the evolution mechanisms and distribution characteristics of regional flood risk but also provide scientific and reliable technical support and decision-making basis for flood emergency response, phased personnel transfer, optimization of emergency resource allocation, and other disaster prevention and mitigation efforts.

Nevertheless, several limitations should be acknowledged. The performance of the Delphi–BN model is contingent upon the quality and representativeness of the expert panel, as well as the spatial resolution of the input data. The consensus criteria ($IQR \leq 1.0$ and $C_V \leq 0.25$) were established to ensure robustness, but alternative thresholds could yield different results. Furthermore, the model was developed and tested in a single semi-arid basin; its transferability to other regions, particularly those with different climatic, topographic, or socioeconomic conditions, requires local recalibration and adaptation. Future work will focus on incorporating real-time data assimilation and dynamic updating capabilities to enhance the model's applicability in rapidly changing flood scenarios.

Author Contributions: The study's conceptualization, original manuscript preparation, funding procurement, project oversight, methodological design, and formal analysis were led by B.Z., J.Q. focused on software implementation, validation, experimental investigation, data management, figure and visualization production, as well as manuscript revision. Z.W. contributed to methodological work, software coding, formal analysis, investigation, and funding. L.Z. handled validation, data curation, revision, and editing of the manuscript, and visualization. L.W. was involved in investigation, visualization, and software-related tasks. W.M. performed formal analysis, validation, and the final review and editing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Achievement Transformation Fund, China Institute of Water Resources and Hydropower Research, grant number ZS1003A012021.

Data Availability Statement: The data presented in this study are not publicly available due to restrictions imposed by the data provider. These data were obtained from the Qinghai Provincial Department of Water Resources and are available from the corresponding author upon reasonable request, subject to the permission of the department.

Acknowledgments: The authors express sincere thanks to the Qinghai Provincial Department of Water Resources for their support, as well as to the data providers. The remote sensing imagery utilized in this study was sourced from Tianditu (Map World). For more information, visit <http://lbs.tianditu.gov.cn/server/MapService.html> (accessed on 17 August 2026).

Conflicts of Interest: Bingyu Zhang, Jing Qin, Zhen Wang, Lingyun Zhao and Lu Wang are employed by Beijing China Institute of Water Resources and Hydropower Research Engineering Group Co., Ltd. The remaining author, Wencai Ma, declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BNs	Bayesian networks
Delphi-BNs	Integrating the Delphi method and Bayesian networks
GIS	Geographic Information System
IFMS	Integrated Flood Modeling System

References

- Asaridis, P.; Molinari, D.; Di Maio, F.; Ballio, F.; Zio, E. A probabilistic modeling and simulation framework for power grid flood risk assessment. *Int. J. Disaster Risk Reduct.* **2025**, *120*, 105353. [\[CrossRef\]](#)
- Scussolini, P.; Luu, L.N.; Philip, S.; Berghuijs, W.R.; Eilander, D.; Aerts, J.C.J.H.; Kew, S.F.; van Oldenborgh, G.J.; Toonen, W.H.J.; Volkholz, J.; et al. Challenges in the attribution of river flood events. *Wiley Interdiscip. Rev. Clim. Change* **2023**, *15*, e874. [\[CrossRef\]](#)
- Chen, F.; Xie, T.; Yang, Y.; Chen, S.; Chen, F.; Huang, W.; Chen, J. Discussion of the “warming and wetting” trend and its future variation in the drylands of Northwest China under global warming. *Sci. China Earth Sci.* **2023**, *66*, 1241–1257. [\[CrossRef\]](#)
- Chen, Y.; Wang, H.; Wang, Z.; Hui, Z. Characteristics of extreme climatic hydrological events in the arid region of northwestern China. *Arid Land Geogr.* **2017**, *40*, 1–9. [\[CrossRef\]](#)
- Gain, A.; Giupponi, C.; Renaud, F. Climate Change Adaptation and Vulnerability Assessment of Water Resources Systems in Developing Countries: A Generalized Framework and a Feasibility Study in Bangladesh. *Water* **2012**, *4*, 345–366. [\[CrossRef\]](#)
- Harrald, J.R. Agility and Discipline: Critical Success Factors for Disaster Response. *Ann. Am. Acad. Polit. Soc. Sci.* **2006**, *604*, 256–272. [\[CrossRef\]](#)
- Son, J.; Aziz, Z.; Peña-Mora, F. Supporting disaster response and recovery through improved situation awareness. *Struct. Surv.* **2008**, *26*, 411–425. [\[CrossRef\]](#)
- Zwirgmaier, V.; Garschagen, M. Linking urban structure types and Bayesian network modelling for an integrated flood risk assessment in data-scarce mega-cities. *Urban Clim.* **2024**, *56*, 102034. [\[CrossRef\]](#)
- Reinstaller, S.; Krebs, G.; Pichler, M.; Muschalla, D. Identification of High-Impact Uncertainty Sources for Urban Flood Models in Hillside Peri-Urban Catchments. *Water* **2022**, *14*, 1973. [\[CrossRef\]](#)
- Nkwunonwo, U.C.; Whitworth, M.; Bailly, B. A review of the current status of flood modelling for urban flood risk management in the developing countries. *Sci. Afr.* **2020**, *7*, e00269. [\[CrossRef\]](#)
- Daupras, F.; Antoine, J.M.; Becerra, S.; Peltier, A. Analysis of the robustness of the French flood warning system: A study based on the 2009 flood of the Garonne River. *Nat. Hazards* **2015**, *75*, 215–241. [\[CrossRef\]](#)
- Balbi, S.; Villa, F.; Mojtahed, V.; Hegetschweiler, K.T.; Giupponi, C. A spatial Bayesian network model to assess the benefits of early warning for urban flood risk to people. *Nat. Hazards Earth Syst. Sci.* **2016**, *16*, 1323–1337. [\[CrossRef\]](#)
- Joo, H.; Choi, C.; Kim, J.; Kim, D.; Kim, S.; Kim, H.S. A Bayesian Network-Based Integrated for Flood Risk Assessment (InFRA). *Sustainability* **2019**, *11*, 3733. [\[CrossRef\]](#)
- Wang, Y.; Zhang, Q.; Lin, K.; Liu, Z.; Liang, Y.; Liu, Y.; Li, C. A novel framework for urban flood risk assessment: Multiple perspectives and causal analysis. *Water Res.* **2024**, *256*, 121591. [\[CrossRef\]](#) [\[PubMed\]](#)
- Du, Z.; Zhong, Q.; Dong, H. Rapid evaluation method of landslide dam stability based on Bayesian network. *Adv. Sci. Technol. Water Resour.* **2023**, *43*, 37–45. (In Chinese) [\[CrossRef\]](#)
- Guo, Z.; Chen, R. Safety monitoring model of gravity dam zoning based on spatial clustering and Bayesian model averaging. *Adv. Sci. Technol. Water Resour.* **2024**, *44*, 41–47. (In Chinese) [\[CrossRef\]](#)
- Wang, H.; Liu, B.; Li, D.; Pan, W. A human factors reliability analysis method for levee safety using Bayesian networks. *Water Resour. Hydropower Eng.* **2024**, *55*, 275–282. (In Chinese) [\[CrossRef\]](#)
- Wu, Z.; Shen, Y.; Wang, H.; Wu, M. Assessing urban flood disaster risk using Bayesian network model and GIS applications. *Geomat. Nat. Hazards Risk* **2019**, *10*, 2163–2184. [\[CrossRef\]](#)
- Abebe, Y.; Kabir, G.; Tesfamariam, S. Assessing urban areas vulnerability to pluvial flooding using GIS applications and Bayesian Belief Network model. *J. Clean. Prod.* **2018**, *174*, 1629–1641. [\[CrossRef\]](#)
- Feng, D.; Shi, X.; Renaud, F.G. Assessing hurricane-induced flood risks in urban areas using Bayesian networks and GIS applications. *Nat. Hazards* **2026**, *122*, 102. [\[CrossRef\]](#)

21. Harris, R.; Furlan, E.; Pham, H.V.; Torresan, S.; Mysiak, J.; Critto, A. A Bayesian network approach for multi-sectoral flood damage assessment and multi-scenario analysis. *Clim. Risk Manag.* **2022**, *35*, 100410. [\[CrossRef\]](#)
22. Lu, Y.; Zhai, G.; Zhou, S. an integrated bayesian networks and geographic information system (BNs-GIS) approach for flood disaster risk assessment_a case study of Yinchuan China. *Ecol. Indic.* **2024**, *166*, 112322. [\[CrossRef\]](#)
23. Liu, J.; Xu, X.; Zou, C.; Lin, N.; Zhang, K.; Shan, N.; Zhang, H.; Liu, R. A Bayesian network-GIS probabilistic model for addressing human disturbance risk to ecological conservation redline areas. *J. Environ. Manag.* **2023**, *344*, 118400. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Gräfe, P.; Blanco-Vogt, A.; Behr, F.-J. Bayesian Networks and GIS-based analysis for flood risk assessment in agriculture during heavy rainfall events. *Int. J. Disaster Stud. Clim. Resil.* **2025**, *1*, 8–21. [\[CrossRef\]](#)
25. Merz, B.; Thieken, A.; Gocht, M. Flood Risk Mapping at the Local Scale: Concepts and Challenges. In *Flood Risk Management in Europe Innovation in Policy and Practice*; Springer: Dordrecht, The Netherlands, 2007; pp. 231–251. [\[CrossRef\]](#)
26. Fernández, D.S.; Lutz, M.A. Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Eng. Geol.* **2010**, *111*, 90–98. [\[CrossRef\]](#)
27. Ma, J.; Yu, H.; Zhang, D.; Zhang, H.; Wu, B.; Mu, J. Application of flood analysis software in flood risk map compilation. *China Water Resour.* **2017**, *5*, 17–20.
28. Yu, H.; Ma, J.; Zhang, D.; Mu, J. Application of IFMS Urban software in urban flood risk map compilation. *China Flood Drought Manag.* **2018**, *28*, 13–17. [\[CrossRef\]](#)
29. Li, N. Application of GIS technology in flood risk map compilation. *China Water Resour.* **2005**, *17*, 17–19.
30. Zhou, G.; Ning, C.; Liu, G. Research on a Dynamic Flood Risk Mapping System—A Case Study of Binjiang District, Hangzhou. *Water Resour. Inform.* **2019**, *20*, 80–84. [\[CrossRef\]](#)
31. Zhong, H.; Ma, T.; Wang, X.; Ji, H. Research on dynamic flood risk mapping system based on open-source GIS and database. *Water Resour. Inform.* **2022**, *6*, 33–37. [\[CrossRef\]](#)
32. Jiang, C.B.; Zhou, Q.; Shen, Y.X. Research progress on coupled hydrological and hydrodynamic models for flood forecasting in mountainous watersheds. *J. Hydraul. Eng.* **2021**, *52*, 1137–1150. [\[CrossRef\]](#)
33. Liu, T.; Hu, B.; Yu, L.; Qiou, C. Research on dynamic flood risk mapping system based on numerical model. *Water Resour. Inform.* **2020**, *4*, 68–72. [\[CrossRef\]](#)
34. Liu, Y.; Zhang, W.; Cui, X.; Meng, B.; Niu, S.; Liang, D. Dynamic real-time flood risk analysis and management system. *J. Jilin Univ. (Earth Sci. Ed.)* **2021**, *51*, 1160–1171. [\[CrossRef\]](#)
35. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*; Morgan Kaufmann: San Francisco, CA, USA, 2014; pp. 7–141. [\[CrossRef\]](#)
36. Zwirgmaier, V.; Reimuth, A.; Garschagen, M. How suitable are current approaches to simulate flood risk under future urbanization trends? *Environ. Res. Lett.* **2024**, *19*, 073003. [\[CrossRef\]](#)
37. Uusitalo, L. Advantages and challenges of Bayesian networks in environmental modelling. *Ecol. Model.* **2007**, *203*, 312–318. [\[CrossRef\]](#)
38. Aguilera, P.A.; Fernández, A.; Fernández, R.; Rumí, R.; Salmerón, A. Bayesian networks in environmental modelling. *Environ. Model. Softw.* **2011**, *26*, 1376–1388. [\[CrossRef\]](#)
39. Sucar, L.E. *Probabilistic Graphical Models*; Springer: Cham, Switzerland, 2021; pp. 111–151.
40. Li, G.; Wu, X.; Han, J.; Li, B.; Huang, Y.; Wang, Y. Flood risk assessment by using an interpretative structural modeling based Bayesian network approach (ISM-BN): An urban-level analysis of Shenzhen, China. *J. Environ. Manag.* **2023**, *329*, 117040. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Kanés, R.; Ramirez Marengo, M.C.; Abdel-Moati, H.; Cranefield, J.; Véchet, L. Developing a framework for dynamic risk assessment using Bayesian networks and reliability data. *J. Loss Prev. Process Ind.* **2017**, *50*, 142–153. [\[CrossRef\]](#)
42. Patiño, E.I.A.; Ariza, M.P.S.; Nogal, M.; Sousa, H.S.; Matos, J.C. Flood risk assessment for road infrastructures using Bayesian networks: Case study of Santarem—Portugal. In *Proceedings of the 19th International Probabilistic Workshop, Stellenbosch, South Africa, 8–9 September 2022*. [\[CrossRef\]](#)
43. Huang, S.; Wang, H.; Xu, Y.; She, J.; Huang, J. Key Disaster-Causing Factors Chains on Urban Flood Risk Based on Bayesian Network. *Land* **2021**, *10*, 210. [\[CrossRef\]](#)
44. Li, M.; Hong, M.; Zhang, R. Improved Bayesian Network-Based Risk Model and Its Application in Disaster Risk Assessment. *Int. J. Disaster Risk Sci.* **2018**, *9*, 237–248. [\[CrossRef\]](#)
45. Zhang, W.; Cui, J.; Yao, W.; Guida, M.; Yeboah, F.K.; Zhou, X.; Li, Y.; Zhang, L.; Liu, G. Bayesian network modeling of flood cascade and climate risks in the Pearl River Delta. *npj Nat. Hazards* **2025**, *2*, 68. [\[CrossRef\]](#)
46. Shahi, K.R.; Sairam, N.; Schoppa, L.; Sang, L.T.; Tan, D.L.H.; Kreibich, H. BN-FLEMOΔ: A Bayesian-network-based flood loss estimation model for adaptation planning in Ho Chi Minh City, Vietnam. *Nat. Hazards Earth Syst. Sci.* **2025**, *25*, 2845–2861. [\[CrossRef\]](#)
47. Schneider, M.; Halekotte, L.; Mentges, A.; Fiedrich, F. Dependent Infrastructure Service Disruption Mapping (DISruptionMap): A method to assess cascading service disruptions in disaster scenarios. *Sci. Rep.* **2025**, *15*, 5736. [\[CrossRef\]](#) [\[PubMed\]](#)

48. Xu, Y.; Wu, J. SMOTE-BN-FLA: Enhanced Bayesian network for rainfall-induced flood loss estimation and mechanism decoding in data-scarce regions. *J. Hydrol.* **2026**, *667*, 134928. [[CrossRef](#)]
49. Gnatzy, T.; Warth, J.; Gracht, H.V.D.; Darkow, I.L. Validating an innovative real-time Delphi approach—A methodological comparison between real-time and conventional Delphi studies. *Technol. Forecast. Soc. Change* **2011**, *78*, 1681–1694. [[CrossRef](#)]
50. Rowe, G.; Wright, G. The Delphi technique: Past, present, and future prospects—Introduction to the special issue. *Technol. Forecast. Soc. Change* **2011**, *78*, 1487–1490. [[CrossRef](#)]
51. Chan, H.; Darwiche, A. Sensitivity analysis in Bayesian networks: From single to multiple parameters. In Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence, Banff, AB, Canada, 7–11 July 2004. [[CrossRef](#)]
52. Kjrulff, U.B.; Madsen, A.L. *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*; Springer: New York, NY, USA, 2008; pp. 69–109. [[CrossRef](#)]
53. Gu, X.; Jiang, H.; Ai, Q.; Xu, C. An Extraction Method of Waterlogged Areas and Analysis of Road Network Traffic Capacity Based on Multi-source Remote Sensing Data Fusion. *J. Transp. Inf. Saf.* **2025**, *43*, 42–53. Available online: <http://www.jtxa.net/en/article/doi/10.3963/j.jssn.1674-4861.2025.06.005> (accessed on 17 August 2026).
54. Song, Y.Q.; Guo, L.; Wang, C.H.; Zhu, J.S.; Li, Z.W. Urban road waterlogging multi-level assessment integrated flood models and road network models. *Transp. Res. Part D Transp. Environ.* **2024**, *133*, 104305. [[CrossRef](#)]
55. Qiao, Y.; Wang, Y.; Ma, G.T. Assessing flood risk to urban road users based on rainfall scenario simulations. *Transp. Res. Part D Transp. Environ.* **2023**, *123*, 103919. [[CrossRef](#)]
56. Cutter, S.L.; Boruff, B.J.; Shirley, W.L. Social Vulnerability to Environmental Hazards. *Soc. Sci. Q.* **2003**, *84*, 242–261. [[CrossRef](#)]
57. Saber, M.; Boulmaiz, T.; Guermoui, M.; Abdrabo, K.I.; Kantoush, S.; Sumi, T.; Boutaghane, H.; Hori, T.; Binh, D.V.; Nguyen, B.Q. Enhancing flood risk assessment through integration of ensemble learning approaches and physical-based hydrological modeling. *Geomat. Nat. Hazards Risk* **2023**, *14*, 2203798. [[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.