Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Understanding social vulnerability in multi-ethnic inter-embedded communities: structural inequalities and disaster risk in Southwest China

Lu Gan
Sichuan Agricultural University
Leqi Zhou
Sichuan Agricultural University
Shan Yang
Sichuan Agricultural University
Xia Liao
Sichuan Agricultural University
Jiangjun Wan
2426509131@qq.com

Sichuan Agricultural University

Research Article

Keywords: Multi-ethnic Inter-embedded Communities, Social Vulnerability, DEMATEL, XGBoost-SHAP, Disaster Governance

Posted Date: July 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10100076/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Applied Spatial Analysis and Policy on September 21st, 2026. See the published version at https://doi.org/10.1007/s12061-026-10021-2.

Understanding social vulnerability in multi-ethnic inter-embedded communities: structural inequalities and disaster risk in Southwest China

Abstract:
Multi-ethnic inter-embedded communities consist of diverse ethnic groups interacting within shared spaces. Such emerging new-type communities are formed with interdependent patterns in social relations, resource access, and governance. Accordingly, it is result to these types exhibit greater complexity than single ethnic ones. In particularly, these regard population structure, social networks, and governance coordination—a complexity further intensified in disaster-prone mountainous regions. While existing social vulnerability research predominantly focuses on urban or conventional communities, studies on multi-ethnic inter-embedded communities remain limited. Many such settlements are located in hazard-prone mountainous areas facing combined challenges of natural disasters, population aging, high mobility, and uneven public services. All the above lead to elevate vulnerability. Therefore, identifying vulnerability characteristics and their driving mechanisms is crucial. This study develops a framework to assess social vulnerability by integrating ethnic-related indicators into the Social Vulnerability Index (SoVI). Utilizing Lushan County as a case study, the research employs Decision Making Trial and Evaluation Laboratory (DEMATEL) and eXtreme Gradient Boosting (XGBoost) coupled with Shapley Additive Explanations (SHAP). The results reveal significant spatial disparities across townships, with ethnic composition, age structure, and education identified as key dimensions. In specific, education levels, dependency ratios, and healthcare capacity emerge as primary influencing factors. The study concludes the social vulnerability stems from both structural and service-related factors. It suggests a governance should shift toward structural identification and capacity building in mountainous multi-ethnic regions.

Keywords Multi-ethnic Inter-embedded Communities • Social Vulnerability • DEMATEL • XGBoost-SHAP • Disaster Governance

Acknowledgements This work was supported by XXX

1 Introduction

Multi-ethnic inter-embedded communities refer to a type of community in which different ethnic groups coexist and interact within a specific geographical space. Over time, these communities form a mutually embedded structure. This structure appears in economic activities, daily life, social relations, and grassroots governance. Supported by shared values, collaborative participation and institutional norms, these communities promote ethnic integration and strengthen a strong sense of national unity. (Li 2024;Xiaona et al. 2026). Unlike single-ethnic communities, they are far more complex. Their complexity is reflected not only in spatial integration but also in deep interaction across cultural exchange, resource sharing and collaborative governance. (Tian and Li 2024).

As the above, such complexity is particularly pronounced in mountainous regions characterized by rugged terrain, active fault zones, and frequent geological hazards. At meantime, dispersed settlements, weak infrastructure and limited public services significantly exacerbate the vulnerability of

multi-ethnic communities(Perry and Green 1982; Fothergill et al. 1999; Carter-Pokras et al. 2007; Hung et al. 2021; Rong et al. 2025) .Statistical data show that between 2000 and 2020, more than half of the direct economic losses and deaths or missing persons resulting from natural disasters in China occurred in multi-ethnic regions (National Bureau of Statistics of China, 2023). This indicates that these regions face significantly greater risk pressures under disaster impacts. Also, they may exhibit higher levels of uncertainty and imbalance when encountering external shocks (Adger, 2006; Graus et al., 2025a; Wisner, 2004). Within this context, social vulnerability is often intensified by social stratification and unequal resource distribution. Moreover, it diminishes the adaptive capacity of sensitive groups(Birkmann et al., 2022; S. Chen et al., 2023; Emrich and Cutter, 2011; Goktas and Rabiei-Dastjerdi, 2026; Tate et al., 2021). Consequently, it is imperative to identify the key drivers of vulnerability in multi-ethnic inter-embedded communities. And then, it is necessary to examine their influence to improve disaster prevention and recovery strategies.

Based on such consideration, this study aims to establish a social vulnerability evaluation system applicable to multi-ethnic inter-embedded communities. In disaster-prone contexts, regional disparities in social development and structure are often significant. Further, they are intensified by population aging, declining birth and marriage rates, and diverse ethnic compositions. Thus, vulnerability levels differ greatly among multi-ethnic inter-embedded regions when they are exposed to disasters. Building on the traditional social vulnerability framework, this study develops an index system tailored to multi-ethnic inter-embedded communities. The system incorporates multiple dimensions, including demographic structure, spatial distribution characteristics, socioeconomic conditions, public service provision, emergency resources, and community governance capacity. Rather than treating residents of multi-ethnic regions as a homogeneous group, the study focuses on internal structural differences and their impact on disaster risk. By integrating machine learning methods to identify key influencing factors, the study reveals the vulnerabilities and constraints of ethnically intermingled communities in disaster response. Thereby, it provides a basis for developing targeted disaster prevention, mitigation, and community governance strategies.

Through the above research framework, this study aims to address the following three questions. i What spatial patterns characterize social vulnerability in multi-ethnic inter-embedded communities? ii Which key factors drive the formation of social vulnerability in these communities? iii In the governance of social vulnerability in multi-ethnic communities, which key issues should be prioritized?This study addresses a key gap in existing research. Results are often identified, but mechanisms remain unclear. This limits their use in decision-making. By linking patterns, drivers, and governance priorities, the study improves the applicability of social vulnerability assessment and supports more targeted disaster risk governance.

The remainder of this paper is organized as follows: Section 2 reviews the relevant literature. Section 3 establishes the problem statement, identifying the research gaps and core questions addressed by this study. Section 4 details the materials and methods. Section 5 introduces the case study of Lushan County and describes its data sources. Section 6 presents the results, outlining the spatial distribution patterns of social vulnerability, its key drivers, and the causal relationships among the various dimensions. Section 7 provides an in-depth discussion of the results. Section 8 concludes the paper by summarizing the main findings and outlining directions for future research.

2 Literature Review

2.1 Multi-ethnic Inter-embedded Communities

Due to inherent natural constraints, multi-ethnic regions often experience relatively slow economic development, weak infrastructure, and complex geographical environments. In addition, linguistic diversity, varied cultural practices, and unique housing structures further exacerbate their susceptibility to disasters(Perry and Green 1982; Fothergill et al. 1999; Carter-Pokras et al. 2007; Hung et al. 2021; Rong et al. 2025).Therefore, reducing disaster losses in these regions has become an important policy concern. Relevant policies emphasize multidimensional integration among ethnic groups. This includes spatial, cultural, economic, and governance dimensions. These policies also encourage equal, interactive, and autonomous relationships among different ethnic groups(Feng, 2022).As a result, multi-ethnic inter-embedded communities have become key units in grassroots emergency response.However, they also exhibit complex social structures and uneven resource distribution.This combination often leads to higher social vulnerability.(Chen et al. 2023).Existing studies show that participatory governance can reduce these vulnerabilities.For example, information exchange and collaboration across ethnic groups can enhance trust and cultural integration.(Driel and Verkuyten 2022; Gafari et al. 2024; Elsenburg et al. 2025).Furthermore, effective governance fosters a collective sense of belonging and transforms ethnic resources into drivers of socioeconomic development(Wen et al. 2009; Tang et al. 2020; Ma et al. 2025) .Moreover, the moderating role of community support suggests that ethnic, religious, and other forms of community support can significantly mitigate the negative impact of disasters on mental health(Nguyen-Thi Tra et al. 2026).However, current research mainly focuses on urban areas with simpler social structures, leaving a gap in studies on social vulnerability in multi-ethnic disaster contexts.Consequently, developing targeted assessment frameworks from a social vulnerability perspective is of profound practical importance.

2.2 Social Vulnerability

“Vulnerability” was initially defined as the sensitivity and resilience a system exhibits when faced with various external shocks(Timmerman 1981). Early research on vulnerability was primarily conducted from a natural science perspective, emphasizing disaster exposure and physical losses(Turner et al. 2003) .However, as research progressed, the concept of vulnerability has expanded to include socioeconomic and institutional dimensions.It highlights the influence of sociocultural factors, resource supply, and management practices.(Dwyer et al. 2004; Emrich and Cutter 2011; Bates et al. 2014).Through this evolution, social vulnerability is understood as a multidimensional condition shaped by demographic, socioeconomic, governance, and service factors.(Cardona 2013). This shift suggests that vulnerability assessment should go beyond the physical aspects of disasters.Instead,it should address structural disparities and social inequalities that shape disaster outcomes.(Adger 2006; Blaikie et al. 2014) .

Within this research framework, demographic characteristics and underlying structural inequalities have gradually become focal points for analyzing social vulnerability. People are central to social systems.They shape socioeconomic development, spatial planning, and policy, and act as key drivers of disaster-related social vulnerability.Research shows that low-income households, older adults, and individuals with low education levels, as well as communities with inadequate infrastructure and healthcare, suffer more severe disaster impacts. This vulnerability results from both exposure and structural inequality.It reflects unequal resource distribution and limited institutional support.(Flanagan et al. 2011; Cutter et al. 2012; Li et al. 2025) . Furthermore, against the backdrop of demographic shifts such as population aging, declining birth and marriage rates, and ongoing migration, community structures are undergoing profound transformations.These changes are not

merely demographic adjustments but will further impact family care giving pressures, the stability of social networks, and community resilience and recovery capacity(Graus et al., 2025b; Hilhorst and Bankoff, 2013;Zhou et al., 2014).Consequently, disparities in population size, age structure, education, and family status—alongside their spatial distribution—serve as key determinants in explaining the spatial heterogeneity of social vulnerability.

2.3 Methods for Measuring Social Vulnerability

At the methodological level, the Pressure and Release (PAR) model reveals the structural mechanisms underlying the formation of vulnerability(Blaikie et al. 2014).It highlights the interactions among social structures, institutions, and the environment. Importantly,it also shifts the focus from purely natural vulnerability to a people-centered perspective, emphasizing social adaptability and resilience as core concerns.The SoVI systematically captures differences in demographic characteristics, economic conditions, and infrastructure. It has become a widely used tool for the quantitative assessment of vulnerability(Cutter et al. 2012). In addition, the Socio-Ecological-Technological Systems(SETS) framework explains vulnerability from the perspective of multi-system coupling(Cardona et al. 2012; Chang et al. 2021).

Meanwhile, machine learning has been extensively applied in vulnerability assessment to enhance robustness and minimize subjectivity(Linh and others 2025).However, multi-ethnic inter-embedded communities are structurally complex systems.In such contexts,quantitative methods alone cannot fully capture their causal relationships or directional interactions. Therefore, applying traditional frameworks to assess social vulnerability in these ethnically integrated contexts remains a formidable challenge.

Despite its broad application, significant controversy persists regarding assessment frameworks and indicator systems(Cutter and Finch 2008; Ge et al. 2017).Current debates primarily focus on three aspects: first, the selection and weighting of variables across demographic, economic, and cultural dimensions; second, the lack of a unified approach to addressing the systemic inter dependencies between these factors; and third, the difficulty of translating identified drivers into prioritized, actionable governance strategies within complex community settings(Kok and De Coninck 2007; Aroca-Jimenez et al. 2017).

3 Problem Statement

In summary, multi-ethnic inter-embedded community refers to collectives in which people from diverse ethnic backgrounds live within a specific geographical area. They jointly participate in the governance of local affairs or represent the community in external matters. Through these processes, interaction, exchange, and integration among different ethnic groups are continuously reinforced(Liang, 2025).Such communities are more than spatial co-residence.They are structures defined by embedded social relations, resource access, cultural interaction, and grassroots governance.Within this structure, differences in ethnic composition do not exist in isolation. Instead,they are closely interwoven with factors such as age distribution, educational attainment, family structure, access to public services, and governance capacity(liu at el. 2015; Song 2013). Collectively,these factors shape a community's disaster prevention, resilience, and recovery capabilities.Therefore, social vulnerability in multi-ethnic communities is a complex condition shaped by multiple factors.It cannot be explained by a single group or a single facility.

However, despite increasing attention to demographic, economic, and institutional factors, several key gaps remain.First, conventional frameworks often exclude ethnic variables, failing to capture how ethnic composition and its internal interactions shape disaster risk. Second, existing

research primarily prioritizes the measurement of vulnerability outcomes, identifying where vulnerability exists but rarely elucidating why or identifying the underlying drivers.Finally, the lack of systematic identification of key drivers and governance priorities hampers the development of targeted, practice-oriented interventions at the grassroots level.

Consequently, this paper proposes a systematic framework to assess social vulnerability and its driving mechanisms within these unique communities. It aims to reveal spatial patterns, underlying logic, and governance priorities.To achieve this, the study adopts a three-level analytical approach. These include vulnerability identification, mechanism explanation, and governance refinement..The overall research framework (Fig.1) comprises three research phases validated through a case study:

Phase I: Expanding the traditional SoVI framework, this phase integrates ethnic indicators to establish a tailored evaluation system and quantify township-level vulnerability.

Phase II: The DEMATEL method is utilized to analyze causal inter-dependencies and influence directions, distinguishing between causal and effect factors of social vulnerability.

Phase III: Utilizing measurement results and causal insights, the XGBoost-SHAP model identifies critical variables driving vulnerability disparities to distill actionable governance priorities.

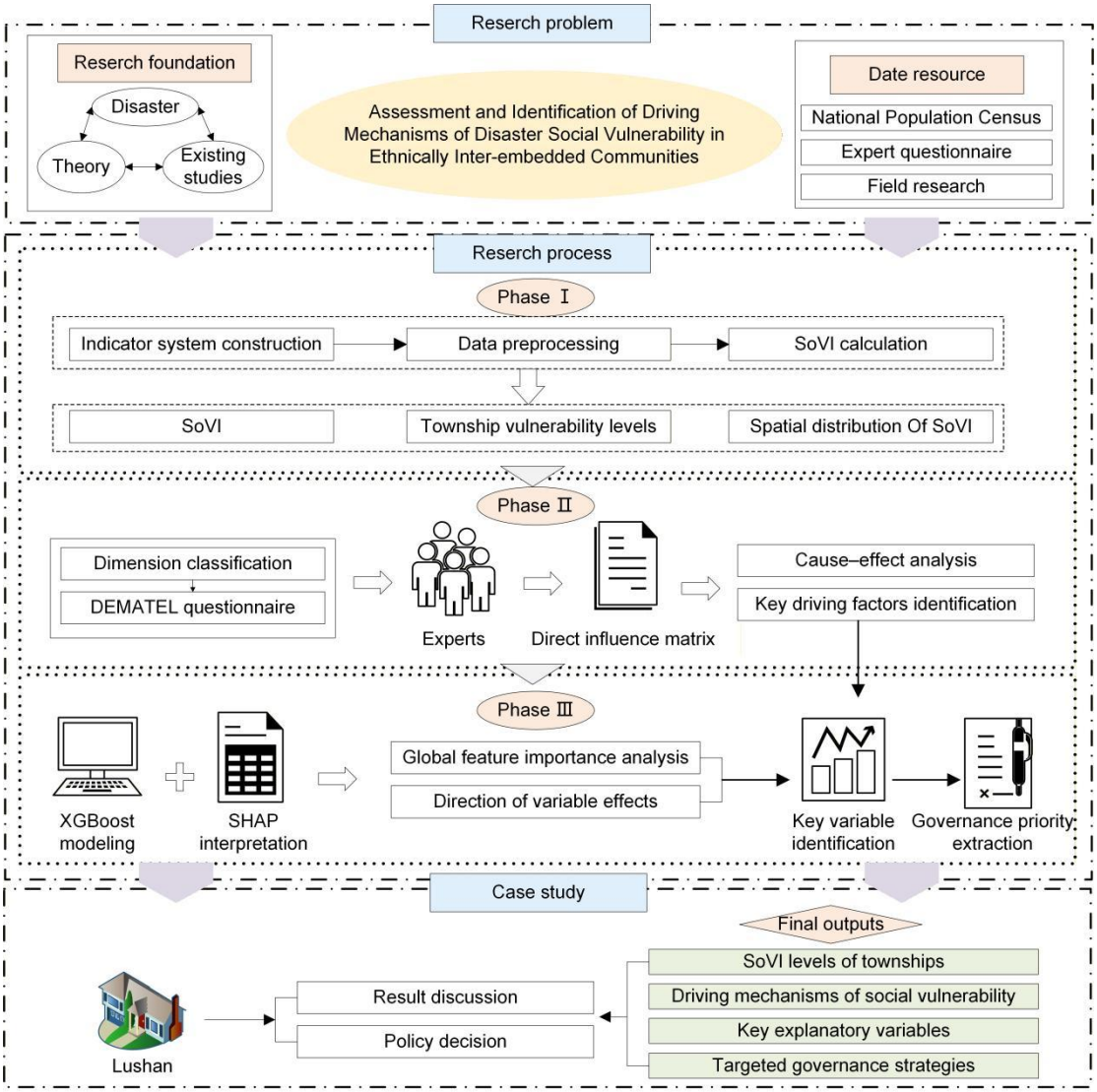


Fig. 1 Research Framework

4 Materials and Methods

189 4.1 SoVI

190 4.1.1 Selection of Social Vulnerability Variables

191 In selecting indicators, this study builds upon the SoVI index framework proposed by
192 Cutter(Cutter et al. 2012) , and fully considered the comprehensiveness, representativeness, scientific
193 validity, and availability of data to select variables for the SoVI(Ma et al. 2020) .First, we selected
194 social vulnerability indicators based on prior studies.These include population, socioeconomic status,
195 and education(Flanagan et al. 2011; Cutter et al. 2012) .We also incorporated indicators relevant to
196 multi-ethnic communities.These include population aging and the share of ethnic minorities(Elsenburg
197 et al. 2025; Rong et al. 2025) Second, to reflect the region’s material foundation for disaster response
198 and recovery, this study selected indicators that reflect the level of community public services and
199 healthcare.These include medical resources and the number of supermarkets.(Cutter et al. 2012) .
200 Finally, to reflect local emergency response capabilities, the number of emergency shelters was selected
201 as a quantitative indicator. Detailed indicator selections are shown in Table 1.

202 **Table1** Table of Social Vulnerability Indicators

Dimension	Indicator	Indicator Description	Indicator Attributes	References
Demographic Characteristics	Proportion of the Elderly	Elderly aged 60 and over in each township	+	(Flanagan et al. 2011; Cutter et al. 2012; Blaikie et al. 2014; Guillard-Gonçalves et al. 2015; Jia et al. 2024; Ouvrard et al. 2025)
	Proportion of the youth population	Children aged 0-14 in each township	+	
	Proportion of females	Proportion of females in the total population of each township	+	
Economic status	Households Receiving Minimum Living Allowance	Households in each township that receive state minimum living allowance subsidies because their household income or housing conditions fall below the local minimum living allowance standard	+	(Holand et al. 2011; Cutter et al. 2012)
Education	Low	Individuals	+	(Holand et al. 2011; Cutter et al.

	educational attainment	aged 15 and older in each township who have not received a high school education or higher		2012; Guillard-Gonçalves et al. 2015; Jia et al. 2024)
	Higher education	Individuals in each township who have received college-level education or higher	-	
Ethnicity	Percentage of Ethnic Minorities	Percentage of the population in each township belonging to ethnic minorities	+	(Bolin and Stanford 1991; Bolin 1994; Peacock et al. 1997; Cutter et al. 2012)
Family Structure	Marriage rate	Degree of community family formation, reflecting social cohesion	-	(Zhou et al. 2014)
	Density of medical institutions	Reflects the accessibility and spatial equity of regional healthcare resources	-	(Cutter et al. 2012; Guillard-Gonçalves et al. 2015; Jia et al. 2024)
Infrastructure	Average Number of Beds per Healthcare Facility	Reflects the accessibility of regional healthcare resources		(Cutter et al. 2012)
	Number of healthcare workers per 10,000 people	Indicates community health protection and	-	(Cutter et al. 2012; Birkmann et al. 2013)

	Number of supermarkets per capita	emergency response capacity Indicates residents' access to basic daily necessities	-	(Cutter et al. 2012)
Emergency response capacity	Number of emergency shelters	Number of emergency shelters in each township	-	(Birkmann et al., 2013)

4.1.2 Calculation of the SoVI

After the indicator system was established, the indicators were standardized to eliminate the impact of differences in units and value ranges on the comprehensive evaluation results. Due to the varying attributes of the indicators, positive and negative standardization methods were applied respectively to convert the original indicator values into dimensionless standard values. Subsequently, Pearson correlation tests were used to remove indicators with extreme multicollinearity.The VIF values were then calculated to ensure acceptable correlation levels.To mitigate the uncertainty arising from subjective weighting, this study employs the Entropy-CRITIC weighting method, as shown in Eq.(1). Weights were calculated using the entropy method, the CRITIC method, and the Entropy-CRITIC method, respectively. Robustness tests and sensitivity analyses were then conducted on the results to verify their consistency, with the robustness test formula shown in Eq.(2) and the sensitivity analysis formula in Eq.(3).Finally, a stable composite SoVI is obtained. A higher value of the SoVI indicates that the region faces greater potential risks under disaster impacts, and that the resilience and recovery capacity of the social system are relatively weak.

$$w_j^{final} = \beta w_j^{entropy} + (1-\beta) w_j^{CRITIC}, 0 \le \beta \le 1 \quad (1)$$

$$\rho_{AB} = 1 - \frac{6\sum_i (r_i^A - r_i^B)^2}{n(n^2-1)} \quad (2)$$

Where r_i^A, r_i^B represents the ranking of sample i under Method A and Method B. $\rho \in [-1,1]$; the closer the value is to 1, the more consistent the ranking.

$$r_{AB} = \frac{\sum_i (S_i^A - \bar{S}^A)(S_i^B - \bar{S}^B)}{\sqrt{\sum_i (S_i^A - \bar{S}^A)^2} \sqrt{\sum_i (S_i^B - \bar{S}^B)^2}} \quad (3)$$

Where S_i^A, S_i^B represent the vulnerability scores of the i-th township under Method A and Method

B, respectively; $\overline{S}^{A}, \overline{S}^{B}$ denote the means of each method; $r_{AB} \in [-1,1]$.

4.2 Analysis of Distribution Patterns and Impact Mechanisms of Social Vulnerability

4.2.1 Analysis of Distribution Differences

To better tailor interventions to regions with different levels of vulnerability, cluster analysis was performed using the K-means algorithm. Among numerous clustering algorithms, the K-means algorithm is widely favored for its simplicity and efficiency(McQueen 1967) . Using social vulnerability indices and related indicators as inputs, the K-means algorithm classifies townships through unsupervised learning.By comparing within-cluster variation and between-cluster differences under different cluster numbers, the optimal number of clusters is determined. The study townships are then classified into several social vulnerability levels.Using the ArcGIS geographic information system, thematic maps were created based on the SoVI results to illustrate the geographic distribution of social vulnerability.

Indicators reflecting structural characteristics—including population density, proportion of the elderly, and proportion of ethnic minorities—were selected. The social vulnerability clustering results were linked with these indicator data at the administrative unit level to ensure that each township possesses complete social vulnerability categories and characteristic fields. ArcGIS was used to visually display the distribution differences of each indicator, and box plots were employed to describe the degree of data dispersion and distribution patterns.Finally, based on the clustering results and box plots, we examine the differences in social vulnerability across different townships.

4.2.2 Analysis of the Mechanisms of Influence Based on DEMATEL and XGBoost-SHAP

To examine relationships among factors, this study uses DEMATEL and XGBoost.These methods form a framework that combines expert knowledge with data-driven analysis.It analyzes social vulnerability from two perspectives.These include causal structure identification and impact intensity estimation.It focuses on social structure, infrastructure, and governance factors.

(1) Causal Relationship Identification of Influencing Factors Based on DEMATEL

The DEMATEL method was originally developed to analyze causal structures among factors in complex decision-making problems(Gabus and Fontela 1973) . In vulnerability research, influencing factors are often not independent of one another but interact through multi-level, multi-path mechanisms(Jia et al. 2024) .The DEMATEL method analyzes causal relationships among factors using expert knowledge without requiring large datasets(Huang et al. 2007; Tzeng et al. 2007) .Therefore, it is effective for explaining complex system mechanisms.Applying this method reveals causal relationships at the system level.It also supports identifying key drivers and setting governance priorities.The main steps are as follows:

Based on the established framework of influencing factors, an expert questionnaire was designed. Nine experts were invited from disaster risk management, social governance, and regional development.They assessed mutual influence among factors using a 5-point scale (0–4). This process ultimately yielded the direct influence matrix Eq.(4):

$$A = [a_{ij}] \tag{4}$$

Where a_{ij} represents the degree of direct influence of factori on factorj .

After standardizing the matrix, the comprehensive influence matrix is calculated using Eq. (5):

$$T = B(1-B)^{-1} \quad (5)$$

Where T represents the total direct and indirect influence between indicators, reflecting the direct and indirect influence relationships of each factor within the system.

To characterize the role of different factors in the formation of social vulnerability, Eq. (6) was used to further calculate the influence (R_i) and susceptibility (C_i) :

$$R_i = \sum_{j=1}^{n} t_{ij}, C_i = \sum_{j=1}^{n} t_{ji} \quad (6)$$

(2) Quantitative Analysis Method for Influencing Factors Based on XGBoost-SHAP

Building on the identified causal structure, this study further applies the XGBoost-SHAP model. This model provides a data-driven analysis of factor influence on social vulnerability.XGBoost is an ensemble learning algorithm based on the gradient boosting framework(Chen 2016) . This method constructs high-performance predictive models by integrating multiple weak learners and offers significant advantages in handling nonlinear relationships, complex interactions among variables, and high-dimensional data(Nachappa et al. 2020; Nguyen et al. 2021) .

In this study, XGBoost is used to quantify the importance of influencing factors on the SoVI. This complements the expert-based results from the DEMATEL method.Thus, both expert knowledge and data-driven insights are integrated.During model construction, the SoVI or its change indicators are used as the target variable. Social structural characteristics and related factors are used as input variables.These variables are used to build an XGBoost regression model.Through model training and parameter tuning, the model learns the mapping relationships between various influencing factors and social vulnerability. Once model training is complete, feature importance and interpretability methods are used to quantify the contributions of influencing factors. These results reveal their potential impacts on social vulnerability from a data-driven perspective.

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (7)$$

Where $l(y_i, \hat{y}_i)$ serves as the loss function, and $\Omega(f_k)$ represents the model complexity penalty term.

Although XGBoost possesses strong predictive capabilities, its model structure is relatively complex and exhibits certain “black-box” characteristics(Nguyen et al. 2021) .To enhance model interpretability, this study introduces the SHAP method to interpret the output results of the XGBoost model(Lundberg and Lee 2017) . This method is a model interpretation approach based on the Shapley value in game theory(Li et al. 2024; Wei et al. 2025) (Li et al., 2024b; Wei et al., 2025). By fairly allocating the contribution of each driving factor to the model’s prediction results, it provides a

quantitative attribution analysis tool for social vulnerability research. Its basic expression is:

$$f(x) = \phi_0 + \sum_{j=1}^{M} \phi_j x_i \quad (8)$$

Where ϕ_0 denotes the extreme prediction value, and ϕ_j represents the SHAP contribution value of the th feature (j).

In this study, mean squared error (MSE) and the coefficient of determination (R^2) were selected as evaluation metrics for the regression task. (The dataset was split into an 80% training set and a 20% test set to ensure the robustness of the model evaluation.)

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \quad (9)$$

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \quad (10)$$

Where n denotes the number of samples, and $\hat{y}_i$ represents the predicted value for the th$i$ sample, while y_i represents its true value. A lower MSE corresponds to a higher R^2, and a higher R^2 indicates superior performance of the social vulnerability assessment framework.

(3) Analytical Approach for the Integration of DEMATEL and XGBoost-SHAP

These two methods analyze the mechanisms influencing social vulnerability from the perspectives of expert cognition and data-driven approaches, respectively, and are complementary in their analytical focus. DEMATEL places greater emphasis on revealing the causal structural relationships among influencing factors, while XGBoost quantifies the relative impact of different factors on the level of social vulnerability. The SHAP method provides interpretations for the output results of the XGBoost model.

By comparing DEMATEL drivers with the high-importance features from XGBoost-SHAP model, we construct a comprehensive framework to explain mechanisms of social vulnerability.On the one hand, DEMATEL results are used to identify key nodes for governance interventions. On the other hand, the XGBoost-SHAP model provides quantitative evidence to inform governance priorities and resource allocation. This integrated approach enables a more systematic and comprehensive analysis of the mechanisms underlying social vulnerability without relying on a single analytical perspective. It helps overcome the limitations of traditional correlation analysis in revealing causal relationships and avoids the pitfalls of relying solely on expert judgment while neglecting data-driven insights.

5 Case Study

5.1 Study Area

This study selects Lushan County in Ya＇an City as the primary case study (Figure 2).Geographically,Lushan is located on the western edge of the mountainous region surrounding the Sichuan Basin, in the northeastern part of Ya'an City (29° 30′ -30° 00′ N, 102° 45′ -103° 10′ E). Situated in the transition zone between the Sichuan Basin and the Qinghai-Tibet Plateau, and within the southern segment of the Longmen Shan Fault Zone, it is one of the most seismically active regions in China.The county covers a total area of approximately 1,191km^2.It is dominated by mountainous and hilly terrain.The topography is higher in the northwest and lower in the southeast. As a result,elevation varies significantly across the county. With Weita Mountain in the center serving as a boundary, the northern and eastern parts belong to the medium-to-high mountainous region. They are characterized by deep valleys and steep slopes, with an average elevation above 2,000 meters. These areas account for more than 90% of the county's total area.The southern region consists of low-mountain river valleys with relatively open terrain and an average elevation below 1,000 meters, exhibiting a typical high-mountain canyon landscape. The county's overall terrain is fragmented with steep slopes, creating a natural environment prone to geological hazards.

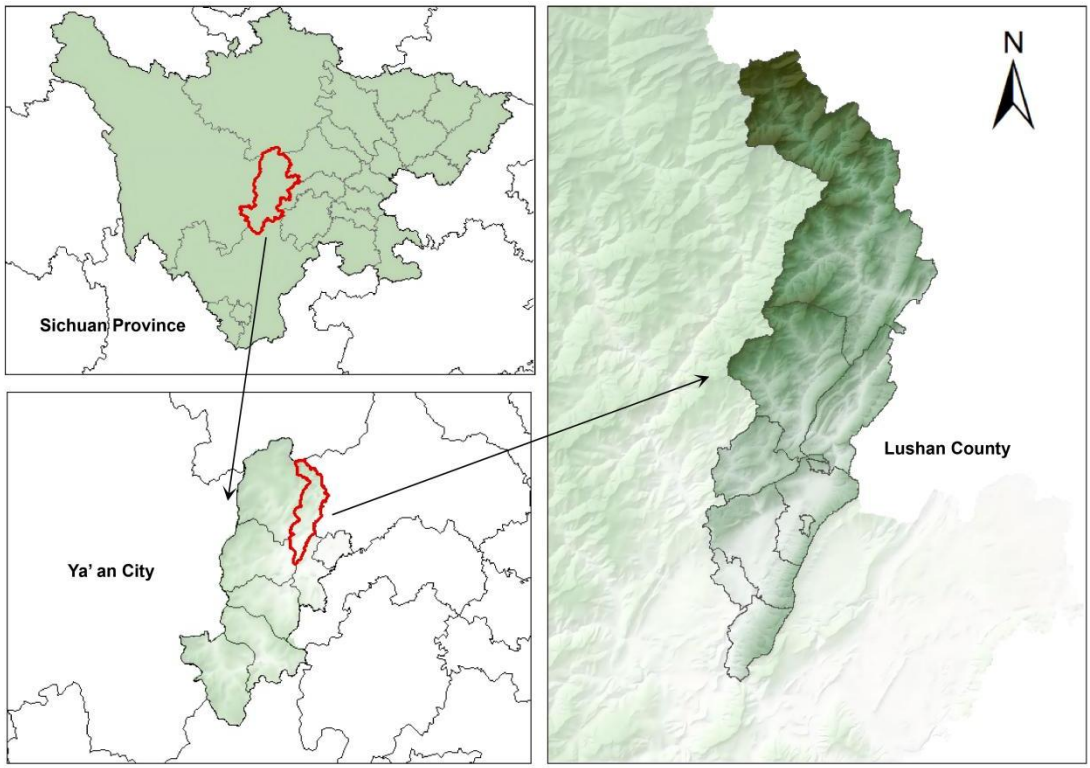


Fig. 2 Study Area

The region has been repeatedly tested by severe seismic events, most notably the 7.0-magnitude earthquake on April 20, 2013, which caused extensive casualties and infrastructure damage while triggering secondary disasters like landslides. Subsequent aftershocks in 2022 and 2025 have posed persistent threats to regional safety and community resilience.In terms of society and culture, Lushan County comprises eight townships with a population of approximately 100,000. In addition to the Han ethnic group, it is home to multiple ethnic groups, including the Tibetan, Qiang, and Hui, exhibiting distinct characteristics of multi-ethnic intermingling. Economically, it is a mountainous agricultural county with a weak development base and a clear urban–rural divide. Despite recent recovery driven by tourism and specialty agriculture, population aging and labor outflow remain significant

challenges.According to data from the Seventh National Population Census, the proportion of the local population aged 65 and older reached 18.29%, higher than the national average.

In summary, Lushan County shows high disaster risk, complex terrain, ethnic diversity, and socioeconomic vulnerability.Therefore, it provides a representative case for studying social vulnerability in multi-ethnic inter-embedded contexts.

5.2 Data Sources

This study integrates data from multiple sources. First, population data are primarily drawn from the Seventh National Population Census (2020), including variables such as the youth population, elderly population, gender structure, educational attainment, and ethnic composition. These indicators are used to describe demographic structure and human capital characteristics. Second, socioeconomic and public service data were collected from the Lushan County Statistical Yearbook and local government reports, covering indicators such as marriage rates, the proportion of households receiving minimum living allowances, the number of supermarkets per capita, and other socioeconomic characteristics. Third, healthcare and emergency service data were sourced from official government channels and obtained through field surveys, including the number of medical institutions, the number of beds per institution, the number of healthcare workers per 10,000 people, and the number of emergency shelters. These indicators were used to reflect service capacity and emergency response capabilities. Furthermore, the expert assessment data used in the DEMATEL analysis were collected through structured questionnaires. Experts were invited to evaluate the interrelationships among key dimensions of social vulnerability, which served as the basis for constructing the direct influence matrix.

All data were processed at the township level and underwent necessary preprocessing steps, including data cleaning, standardization, and indicator calculation, to ensure comparability among different variables.

6 Results

6.1 Development of the SoVI

Building upon traditional social vulnerability indices, we incorporated indicators tailored to China's national context—such as the proportion of ethnic minorities and governance capacity—to construct a SoVI system applicable to the perspective of ethnically interembedded communities in China. Using three weighting methods — the Entropy method, the CRITIC method, and the Entropy-CRITIC method (β =0.5)—we calculated the social vulnerability indices for each township in Lushan County. The results are shown in Table 2.

Table 2 SoVI of Townships in Lushan County

Township SoVI	Entropy Weighting Method	CRITIC	Entropy-CRITIC	z-score
Luyang Subdistrict	0.353423444	0.437680758	0.452301582	-2.169460016
Feixianguan Town	0.724098949	0.639625718	0.684745629	0.88615998
Shuangshi Town	0.663107204	0.546981909	0.580131414	0.383381693
Taiping Town	0.680285665	0.649070806	0.616746872	0.524990314
Dachuan Town	0.551207776	0.428824161	0.461108402	-0.5390481

Siyan Town	0.710197536	0.597214689	0.660708329	0.771565314
Longmen Town	0.587892386	0.433399775	0.512606956	-0.236642826
Baosheng Township	0.662582171	0.515713139	0.572200125	0.379053641

To verify the reliability and consistency of the results obtained from these three social vulnerability indices, this study employed Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (ρ) to conduct robustness tests on the composite scores derived from different weighting methods.The Pearson correlation coefficient measures the linear consistency among the scores of different schemes, while the Spearman rank correlation coefficient reflects the consistency in the ranking order of different schemes. If both coefficients are close to 1, it indicates that the results of each scheme maintain a high degree of consistency in both numerical values and ranking, thereby demonstrating that the model possesses strong robustness. In specific, the calculation results are shown in Table 3:

Table3 Comparison of Weighting Methods

Comparison of Methods	Spearman ρ	Explanation	Pearson r	Interpretation
Entropy vs. CRITIC	0.8571	More robust	0.9277	More robust
Entropy vs. Hybrid (β=0.5)	0.9988	Extremely robust, with nearly identical rankings	0.9689	Virtually identical
CRITIC vs. Hybrid	0.9994	Extremely robust, with rankings almost identical	0.8707	The results of the two methods are highly consistent, and the model is robust

The results show that the Spearman's ρ values for all three sets of schemes exceed 0.85, indicating that the rankings of social vulnerability remain highly consistent across different weighting methods. In particular, the ρ value between the entropy method and the comprehensive method reaches 0.9988, indicating near-perfect agreement.High correlations among entropy, CRITIC, and the comprehensive method indicate effective information integration.The hybrid model produces stable weighting results.Pearson coefficients above 0.87 confirm strong numerical consistency and similar trends.

Overall, the social vulnerability indices under different weighting schemes show strong consistency in spatial distribution, numerical characteristics, and ranking patterns. This indicates that the indicator system and the comprehensive weighting model used in this study are robust and reliable.

To objectively classify the social vulnerability levels of each township in Lushan County, this study conducted a clustering analysis using the K-means method based on the standardized SoVI composite scores. First, the Elbow Rule was applied to determine the optimal number of clusters (Figure 3).The sum of squared errors within clusters showed a flattening trend after k=3, forming a distinct inflection point.This result indicates that k=3 is the critical point for data partitioning efficiency.Beyond this point, the marginal benefit of increasing the number of clusters is limited. Therefore, the number of clusters is set to k=3, corresponding to three social vulnerability levels: Level I(low vulnerability),Level II(moderate vulnerability),and Level III(high vulnerability). The

vulnerability levels of each township are shown in Table 4:

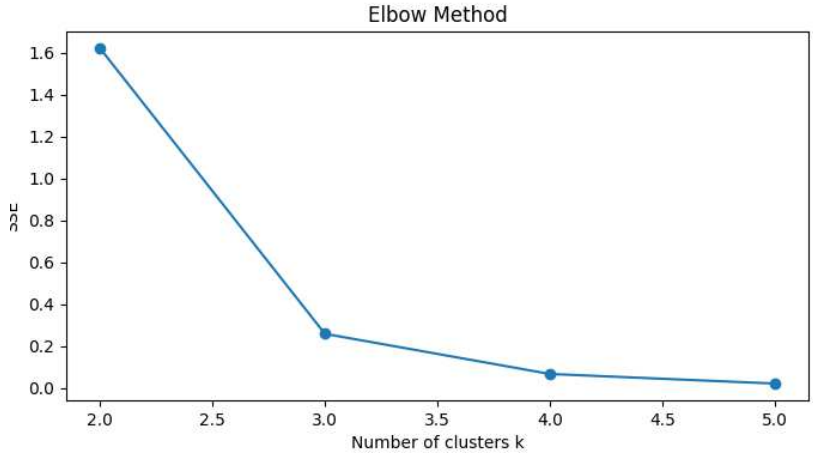


Fig. 3 Elbow Plot of Optimal Number of Clusters

Table 4 Social Vulnerability Levels of Each Township

Township	SoVI	Cluster	Vulnerability_Level	Vulnerability_Label
Luyang Subdistrict	-2.1695	1	1	Low
Feixianguan Town	0.8862	2	3	High
Shuangshi Town	0.3834	2	3	High
Taiping Town	0.5250	2	3	High
Dachuan Town	-0.5390	0	2	Middle
Siyan Town	0.7716	2	3	High
Longmen Town	-0.2366	0	2	Middle
Baosheng Township	0.3791	2	3	High

Based on clustering results, social vulnerability shows clear spatial differentiation across townships in Lushan County (Figure 4). Vulnerability generally increases from north to south and from central areas toward mountainous regions.Level III townships are mainly concentrated in southern and remote mountainous areas. These regions are characterized by complex terrain, scattered public services, and higher aging levels, resulting in greater vulnerability. Level II are mainly distributed in northern and transitional areas. Although basic infrastructure is available, limitations in healthcare and educational accessibility, together with labor outmigration and aging populations, still constrain disaster response and recovery capacity. Level I townships are concentrated near the county center, with Luyang Subdistrict as a typical example. These areas have stronger socioeconomic conditions, better public services, and more balanced population structures, resulting in lower vulnerability levels. Overall, social vulnerability is not uniformly distributed but exhibits a certain spatial coupling relationship with topographical conditions, population structure, and access to public services.

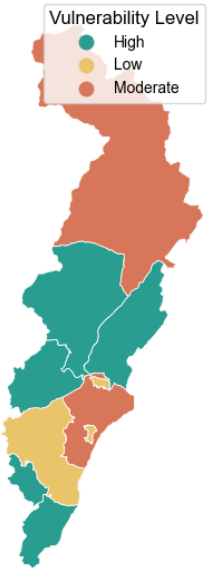


Fig. 4 Distribution Map of Social Vulnerability Levels

To examine the relationship between spatial patterns and demographic characteristics, this study conducted an overlay analysis in Lushan County.It combined vulnerability clustering results with key demographic indicators.These include child share, aging index, population density, and ethnic minority proportion.Using visualization tools in ArcGIS and Python, spatial distribution maps of demographic characteristics(Figure 5)was generated for each township in Lushan County.

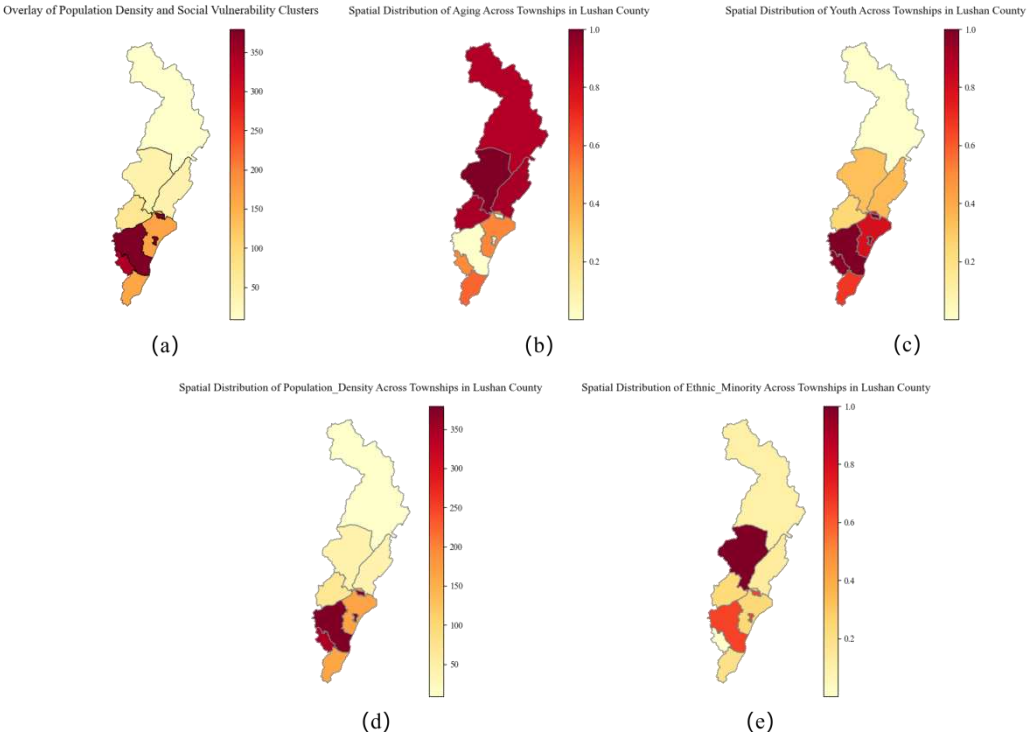


Fig.5 Spatial distribution maps of demographic characteristics

In terms of spatial distribution characteristics, the various demographic indicators in Lushan County exhibit significant regional variations. Areas with a higher proportion of youth are primarily concentrated in the southern townships, located near the county center, indicating strong demographic vitality and potential for development.However, a high dependency ratio places additional pressure on families and communities in disaster scenarios.Areas with higher levels of aging are primarily

distributed across northern and central townships.Comparison shows spatial overlap between areas with high aging rates and high vulnerability zones.This suggests that elderly dependency may reduce community capacity for self-rescue and mutual aid.Population density is highest in the southern regions and areas surrounding towns. These areas exhibit high population concentration and active economic activity, but they may also face social issues such as housing shortages and infrastructure pressures.The ethnic minority population is primarily concentrated in certain townships in the central and southern regions. However, due to the small differences in population proportions across townships, this factor has limited explanatory power regarding current vulnerability disparities at the county level.

To further identify differences in population characteristics across different types of social vulnerability, this study conducted a box-and-whisker plot analysis of key demographic indicators for each vulnerability category (Figure 6).

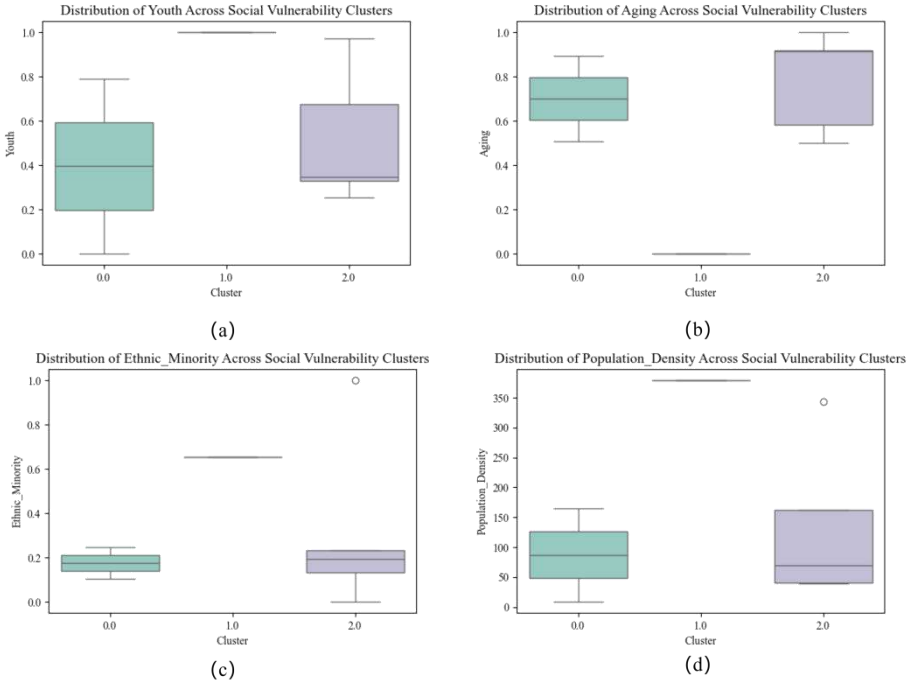


Fig.6 Box-and-whisker plot

The proportion of adolescents in high-vulnerability areas is generally higher and exhibits significant fluctuations.These areas often face a heavier dependency burden. Population aging is generally higher in high-vulnerability areas.In contrast, low-vulnerability areas show weaker clustering patterns.This indicates a positive relationship between aging and vulnerability.Population density fluctuates most in moderately vulnerable areas. In contrast, highly vulnerable areas do not necessarily have the highest density. This suggests that vulnerability is influenced more by demographic structure and resource allocation than by population size alone.Overall, the spatial variations in social vulnerability in Lushan County are closely linked to the structure of the dependency burden.

6.2 Causal Analysis and Impact Intensity Analysis

6.2.1 Confirming Causality Using DEMATEL

To facilitate expert scoring using the DEMATEL method, a direct influence matrix among influencing factors was established. After normalization, a normalized direct influence matrix was obtained. Through iterative calculations, a comprehensive influence matrix was derived, along with measures of influence, susceptibility, centrality, and causality. The 13 social vulnerability indicators were consolidated into 8 dimensions based on their theoretical concepts (Table 5).

Table 5 Dimensions of Social Vulnerability

Number	Dimension
Y1	Socioeconomic Status
Y2	Education Level
Y3	Public and Medical Services
Y4	Emergency and Governance Capacity
Y5	Household Structure
Y6	Age Structure
Y7	Ethnic Composition
Y8	Gender Structure

（1）Constructing the Direct Influence Matrix

First, a questionnaire survey was conducted to interview nine experts with extensive practical and theoretical experience in the field of disaster prevention. Using a “0-4 scale method”，they rated the strength of the relationship between each pair of influencing factors. The scores from all participating experts were then averaged to obtain the direct influence matrix (Table 6).For the direct influence matrix A, the sum of each row was calculated, and the matrix data was uniformly divided by the maximum row sum to obtain the normalized matrix B (Table 7).

Table 6 Direct Influence Matrix

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
Y1	0.000	0.111	0.000	2.111	2.667	2.889	2.000	3.333
Y2	0.125	0.000	0.000	1.778	1.556	1.333	1.333	2.111
Y3	0.250	0.375	0.000	2.222	2.111	2.444	2.333	2.556
Y4	1.375	0.875	1.000	0.000	2.556	3.333	3.000	3.111
Y5	1.500	1.500	1.000	2.250	0.000	2.667	1.444	2.000
Y6	1.875	1.000	1.250	3.500	3.125	0.000	3.444	3.333
Y7	2.000	1.000	1.375	1.750	1.875	2.000	0.000	2.667
Y8	1.375	0.875	0.500	1.250	1.000	2.250	2.500	0.000

Table 7 Normalized Direct Influence Matrix

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
Y1	0.000	0.006	0.000	0.120	0.152	0.165	0.114	0.190
Y2	0.007	0.000	0.000	0.101	0.089	0.076	0.076	0.120
Y3	0.014	0.021	0.000	0.127	0.120	0.139	0.133	0.146
Y4	0.078	0.050	0.057	0.000	0.146	0.190	0.171	0.177
Y5	0.086	0.086	0.057	0.128	0.000	0.152	0.082	0.114
Y6	0.107	0.057	0.071	0.200	0.178	0.000	0.197	0.190
Y7	0.114	0.057	0.078	0.100	0.107	0.114	0.000	0.152
Y8	0.078	0.050	0.029	0.071	0.057	0.128	0.143	0.000

（2） Comprehensive Impact Matrix

By calculating the normalized direct influence matrix B using Eq.(5), the composite influence matrix T can be obtained. The values in matrix T represent the composite influence of one factor on another.See Table 8 for the specific distribution of these values.

Table 8 Comprehensive Influence Matrix

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
Y1	0.227	0.154	0.139	0.427	0.452	0.508	0.461	0.566
Y2	0.148	0.092	0.086	0.289	0.276	0.296	0.293	0.356
Y3	0.225	0.158	0.130	0.412	0.404	0.463	0.455	0.503
Y4	0.325	0.211	0.207	0.363	0.488	0.573	0.554	0.609
Y5	0.283	0.212	0.177	0.414	0.299	0.472	0.411	0.479
Y6	0.378	0.236	0.237	0.573	0.556	0.464	0.620	0.673
Y7	0.309	0.187	0.196	0.390	0.395	0.445	0.336	0.512
Y8	0.242	0.155	0.131	0.310	0.298	0.387	0.397	0.307

（3） Influence, Affectedness, Centrality, Causality

D, R, D+R, and D-R are calculated using Eq.(6), as shown in Table 9.D+R is used to measure the overall strength of an influencing factor's effect and the extent to which a dimension participates in interactions within the entire system.The larger the value, the more central the dimension is. Meanwhile, D-R is used to reveal the causal characteristics of factors.A positive D-R corresponds to causal-category factors and a negative value corresponds to effect-category factors.The difference in signs serves only to distinguish between these types.

Table 9 D, R, D+R, and D-R Values

Dimension	D (Influence Out)	R (Influence In)	D+R (Prominence)	D-R (Relation)	Type
Y1	3.870	3.878	7.748	-0.008	Effect
Y2	3.597	3.279	6.876	0.318	Cause
Y3	2.817	3.854	6.671	-1.037	Effect
Y4	2.211	4.452	6.663	-2.241	Effect
Y5	2.900	3.335	6.236	-0.435	Effect
Y6	3.236	2.156	5.392	1.080	Cause
Y7	3.025	1.312	4.338	1.713	Cause
Y8	2.025	1.416	3.440	0.609	Cause

Based on DEMATEL's centrality (D+R) and causality (D-R) metrics, the social vulnerability system of multi-ethnic communities in Lushan County exhibits a distinct hierarchical structure characterized by "social structure as the driver—capacity as the carrier" (Figure 7).Among these, socioeconomic status and educational attainment are the most central structural variables in the entire social vulnerability system, indicating that human capital and the economic foundation play a fundamental role in the formation of vulnerability at the township level. At the same time, although public services and emergency governance capacity belong to the capability level, they also possess high structural importance within the system, indicating that they play a key role in bearing social risks.

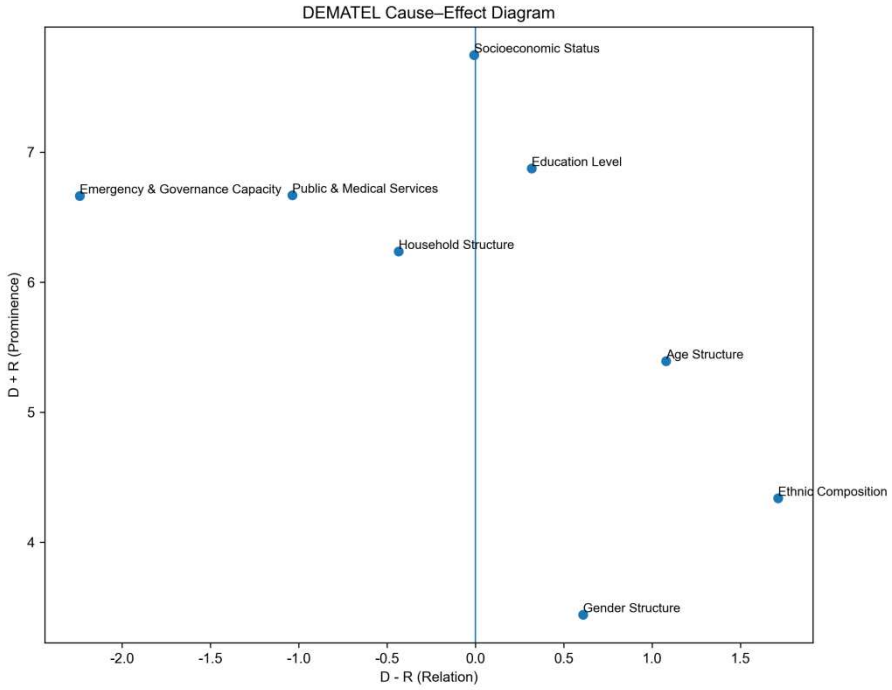


Fig. 7 Causal Structure Diagram

Causal analysis shows that demographic and social structural factors, such as ethnic composition, age structure, and gender structure, act as primary driving variables. In contrast, emergency resources, governance capacity, public services, and medical resources are mainly outcome variables.This indicates that a community's disaster response capacity does not develop in isolation but is shaped by the layered transmission of long-term demographic structures and social conditions. Therefore, improving disaster governance requires more than increasing emergency resources.It depends on strengthening demographic structures and social development conditions.Moreover, governance capacity is not an independent variable.It results from interactions among demographic structure, education, and socioeconomic conditions.

A comprehensive analysis reveals that the social vulnerability system in Lushan County's multi-ethnic communities exhibits a distinct hierarchical structure. Demographic and social structural factors serve primarily as driving forces, while public services and emergency governance capabilities manifest as outcome variables.Education and socioeconomic conditions play a crucial mediating role between the two. This indicates that a community's disaster response capacity is fundamentally the result of long-term social structural evolution, rather than a matter of resource allocation.

6.2.2 Identifying the Strength of Influencing Factors Using the XGBoost-SHAP Model

To identify the key drivers of variations in the SoVI across townships in Lushan County, this study constructed an XGBoost-based regression model at the township level (n=8) and employed the SHAP method to conduct interpretability analysis of the model results (Figure 8). Additionally, given the small sample size, Leave-One-Out Cross-Validation (LOOCV) was used to test the model's robustness(Table 10).

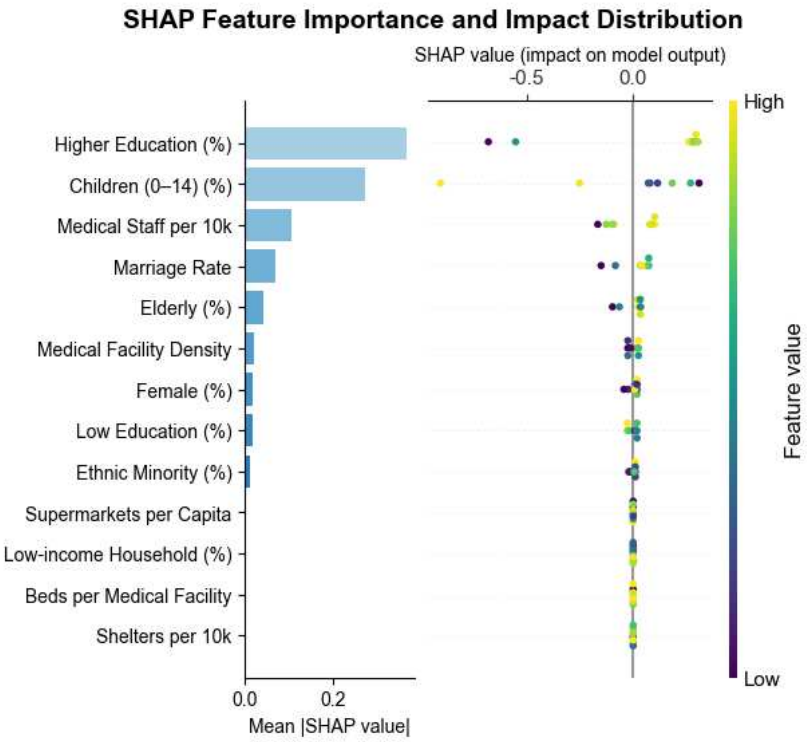


Fig. 8 SHAP Feature Importance and Impact Distribution

Table 10 LOOCV Prediction Table

township	SoVI_true	SoVI_pred_LOOCV
Luyang Subdistrict	-2.169460016	0.770005822
Feixianguan Town	0.88615998	0.011581494
Shuangshi Town	0.383381693	0.18842046
Taiping Town	0.524990314	-1.075784087
Dachuan Town	-0.5390481	0.350578994
Siyan Town	0.771565314	0.708783388
Longmen Town	-0.236642826	0.817340434
Baosheng Township	0.379053641	0.465547889

(1) Model Fitting Results and Robustness Tests

LOOCV results show that the simplified model achieves lower prediction errors than the initial model (MAE = 0.8146; RMSE = 1.1218), indicating improved generalization under small-sample conditions. Overall, the model captures relative vulnerability differences across townships, although some bias remains in extreme samples. Given the limited sample size and multiple variables, the model performance is acceptable for township-level social vulnerability analysis. Therefore, this study mainly applies XGBoost-SHAP to identify key drivers and their influence directions rather than for high-precision prediction.

(2) SHAP Global Importance Analysis: Identification of Key Drivers

Global SHAP importance rankings reveal clear hierarchical differences among variables. The share of highly educated individuals has the highest importance, highlighting the dominant role of human capital in explaining vulnerability differences. The percentage of children and healthcare workers per 10,000 people also show strong explanatory power, indicating the importance of dependency structure and healthcare capacity. In contrast, variables such as medical institution density,

female proportion, low-education share, and ethnic minority proportion contribute relatively little under current sample conditions.

(3) SHAP Directional Analysis: Parsing the Mechanisms of Variable Effects

The SHAP beeswarm plot further reveals the directional and nonlinear effects of variables on social vulnerability (Figure 8). Higher education levels are associated with higher SoVI values in this sample, suggesting that education does not always function as a protective factor. Its effects may depend on spatial and demographic conditions. The proportion of children also shows a significant positive effect, reflecting greater caregiving and resource pressures during disasters.

In contrast, healthcare workforce allocation shows a clear vulnerability-buffering effect, indicating that stronger healthcare capacity improves emergency response and recovery. The marriage rate negatively affects vulnerability, suggesting that stable family structures support community resilience. Meanwhile, the elderly share shows a relatively limited marginal effect.Several infrastructure-related indicators, including supermarkets per capita, hospital beds, and shelters, show low SHAP importance. This may result from limited inter-township variation and overlapping effects with demographic and healthcare variables. Overall, vulnerability differences across townships are mainly driven by human capital, dependency structure, and healthcare capacity rather than infrastructure quantity (Figure 9).

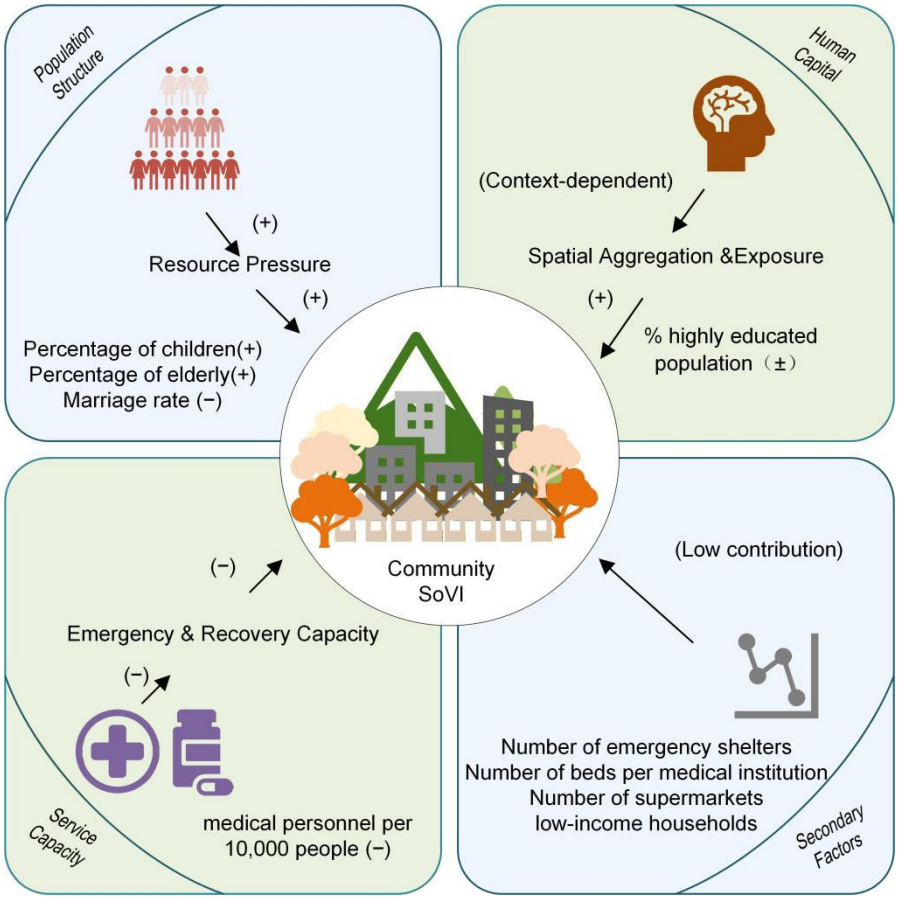


Fig.9 The mechanism diagram corresponding

7 Discussion

The frequent occurrence of extreme natural disasters is caused by a combination of natural and human factors. By examining interactions among factors influencing social vulnerability, this study highlights the role of multi-ethnic coexistence, social ties, and diverse cultural and institutional needs.

It further analyzes social, economic, and governance conditions to identify key issues, providing a basis for more effective disaster risk management and decision-making.

More specifically, this study uses county-level data through a hybrid DEMATEL and XGBoost-SHAP approach. This approach reveals both formation mechanisms and governance pathways.Unlike conventional studies relying on a single methodology, this research addresses two fundamental questions:First,which variables operate at the causal dimension within the system's structure. Second, which variables best explain variations in SoVI across townships. Consequently,this helps identify vulnerable areas and populations more clearly. At the same time,it also supports the design of targeted interventions based on local conditions, thereby strengthening disaster response capacity.Moreover,research from a community perspective in multi-ethnic regions remains limited.Therefore,this study provides evidence to support disaster risk management policies in these areas.

7.1 Presentation of Key Findings

By synthesizing the results of SoVI, DEMATEL causal identification, and XGBoost-SHAP interpretation,several fundamental patterns emerge:

(1) **Spatial Heterogeneity of Vulnerability:** The SoVI results reveal that social vulnerability is not homogeneously distributed but varies significantly at the township level. Therefore, vulnerability is not determined solely by hazard exposure.Instead, it depends on internal social structures.These include population composition, social resources, and service distribution.While the disaster context imposes a uniform external shock, the divergent internal social architectures of these townships determine their unique vulnerability profiles.

(2) **Dominance of Demographic and Social Structural Determinants:** Both causal analysis and SHAP interpretations indicate that human capital, dependency burdens, and public service capacity are the primary drivers of vulnerability. In multi-ethnic inter-embedded communities, vulnerability is less a product of immediate resource scarcity and more a result of deep-seated structural factors.These factors shape socioeconomic status, healthcare capacity, and governance effectiveness by influencing education, labor division, family structure, risk communication, and public services.In particular, dependency structures, reflected in the shares of children and older adults, are critical.Importantly,they show that social vulnerability is often formed under normal conditions through accumulated caregiving and resource pressures, rather than during disasters alone.

(3) **Systemic Coupling of Multi-dimensional Factors:** Structural and service-related factors are not mutually exclusive; they jointly shape vulnerability through systemic coupling. While DEMATEL identifies the causal direction of these dimensions, XGBoost-SHAP elucidates their explanatory power regarding spatial variations. In Lushan County, vulnerability arises from the convergence of ethnic composition, age structure, and educational attainment, which collectively constrain governance capacity and socioeconomic resilience. Ultimately, social vulnerability in these communities represents the concentrated manifestation of pre-existing structural disparities when faced with external shocks.

7.2 Application of the Findings to Regions Similar in Southwest China

The findings from Lushan County are not isolated but represent a broader phenomenon across the multi-ethnic regions of Southwest China. The mountainous areas of Southwest Sichuan exhibit a high degree of consistency in their socio-spatial characteristics. For example, Aba, Garzê, and Liangshan show comparable conditions.These regions have diverse ethnic compositions and complex terrain.They also face limited public services and high disaster exposure.

Official data from Aba Tibetan and Qiang Autonomous Prefecture confirms its status as a multi-ethnic hub primarily comprising Tibetan, Qiang, Hui, and Han populations.Due to its complex and variable environment, Aba faces high disaster risks.China's 14th Five-Year Plan for Geological Disaster Prevention identifies 4,062 hazard sites in its key zones.Therefore, it is one of the most affected regions in China.Similarly, Ganzi Tibetan Autonomous Prefecture features a landscape of high mountains and deep valleys similar to that of Lushan. It along with a multi-ethnic population comprising Tibetans, Han, Yi, and Qiang, among others. Recent seismic activity and extreme rainfall have further increased the frequency and scale of disasters in this region. Furthermore, Liangshan Prefecture hosts the largest Yi population in China.Its 2023 disaster reduction plan identifies it as one of the highest-risk regions in the province.Thus,these shared demographic and environmental commonalities provide a robust empirical basis for generalizing the Lushan County findings to other ethnically regions across Southwest China.

Against this backdrop, the finding presented in this paper—that social vulnerability is primarily driven by demographic and social structures—demonstrates strong regional applicability.Taking Lushan County as an example, ethnic composition, age structure, educational attainment, and the quality of healthcare services are key factors in explaining variations in the SoVI across townships.Similar logic holds true in other ethnic regions of Southwest China. In multi-ethnic communities in Aba, Ganzi, and Liangshan, several factors are critical.These include age structure, education, healthcare, and governance capacity.They play a greater role in resilience and recovery than infrastructure quantity.Essentially, for mountainous ethnic regions of similar scale and typology, vulnerability is a function of whether structural factors within the community have reached a state of equilibrium.

From the perspective of ethnic relations, Lushan County exemplifies the challenges of interethnic coexistence, interaction, negotiation, and resource sharing within a single geographical space. Compared to single ethnic settings, the governance of multi-ethnic inter-embedded communities is complicated by divergent cultural practices, linguistic barriers, and social identities. The critical concern is whether these cultural differences translate into inequitable risk-bearing capacities during disasters.This observation holds significant implications for other multi-ethnic inter-embedded communities as well.

7.3 Practical Value for Different Stakeholders

Based on the findings of this study, different departments and organizations can develop more targeted action priorities in accordance with their respective responsibilities. **For local governments**, The SoVI classification should be used to institutionalize tiered township management. By prioritizing resource allocation based on social vulnerability rather than just geographical size, governments can ensure that limited fiscal resources are directed toward the highest-risk areas.Research shows that the key factors influencing social vulnerability extend beyond the quantity of shelters and infrastructure. They also include dependent populations, such as children and the elderly, as well as the capacity of primary healthcare services. Therefore, **emergency management departments** should pay greater attention to the actual needs of vulnerable groups and enhance service capabilities when formulating emergency response plans and organizing routine drills. At the same time, ethnic affairs work should not be limited to cultural exchange and social integration but should be further integrated into disaster governance.**The National Ethnic Affairs Commission** should strengthen cross-ethnic communication, information sharing, and collaborative participation to address vulnerabilities rooted in structural disparities.**Natural resources and planning authorities** must not base spatial planning and public

service allocation solely on natural disaster risk levels and the principle of balanced facility distribution.They should also factor in variations in social vulnerability across different townships to enhance the social relevance and risk-specificity of planning decisions. For social forces and community organizations,enhancing grassroots resilience should focus on social capital over large-scale investment. Cultivating mutual aid networks and conducting routine, culturally sensitive risk education are more effective for long-term recovery than infrastructure alone.Ultimately, disaster governance in multi-ethnic contexts must integrate structural identification, capacity building, and targeted intervention into a synergistic, holistic strategy.

7.4 Policy Recommendations

Based on the empirical evidence, this study proposes four strategic recommendations. These measures should be co-designed and implemented in partnership with local communities to ensure site-specific calibration and enhanced operational effectiveness.

（1）**Institutionalize a tiered governance framework.** Given that social vulnerability is non-homogeneous, a monolithic "one-size-fits-all" approach to disaster mitigation is inadequate. Instead, a tiered governance system should be established, calibrating interventions based on the specific risk profiles of high-, medium-, and low-vulnerability townships. This ensures that finite fiscal and administrative resources are directed precisely toward the populations and regions where the need for intervention is most acute.

（2）**Establish a longitudinal social vulnerability monitoring mechanism.** A social vulnerability monitoring mechanism covering the township level should be established by leveraging county-level statistics, grassroots governance, and emergency management systems. These indicators, including aging, child share, migration, ethnic composition, education, and healthcare capacity, should be regularly monitored and stored in standardized databases. Regular assessments of trends in social vulnerability across townships should be conducted to ensure that highly vulnerable areas and key vulnerable groups can be identified in a timely manner prior to disasters.

（3）**Prioritize operational optimization over physical expansion.** Future strategies should move beyond the mere pursuit of facility expansion, focusing instead on the efficiency of public resource utilization. This involves strengthening the synergistic coordination between shelters, medical institutions, and community governance networks. Furthermore, investment should be channeled into human capital—enhancing the risk identification and emergency response capabilities of grassroots medical staff, emergency personnel, and community workers.

（4） **Advance inclusive governance through inter-ethnic embeddedness.** In multi-ethnic contexts, resilience is rooted in social capital as much as administrative resource allocation. Policy design must prioritize inter-ethnic communication and mutual assistance networks. This includes fostering community-based platforms for risk awareness and linguistic-cultural adaptation of disaster warnings. By reducing communication barriers and ensuring equitable access to institutional services, governance can mitigate the structural inequalities that often exacerbate disaster impacts in ethnically diverse regions.

8 Conclusions

As critical spatial units for cohabitation and joint governance, multi-ethnic inter-embedded communities face increasingly complex social vulnerability risks amidst frequent mountainous disasters. However, research on these communities remains limited.Therefore, this study develops a

comprehensive framework to assess social vulnerability and its drivers.

The results reveal significant spatial heterogeneity in social vulnerability across townships.Vulnerability varies significantly rather than being evenly distributed.Moreover, key drivers are mainly rooted in demographic and structural factors.These include age structure, ethnic composition, and education level. In specific, human capital disparities and healthcare service capacity explain the observed variations in vulnerability.Therefore,social vulnerability is not determined by a single dimension. Instead, it results from interactions among multiple dimensions.These include structural factors and service-related conditions.

Although this paper offers valuable insights into the assessment of social vulnerability and the analysis of driving mechanisms in multi-ethnic inter-embedded communities, there is still room for further exploration. Future research can be advanced in three areas. First, incorporating multi-temporal data would enable analysis of the dynamic evolution of social vulnerability and its long-term drivers. Second, indicators reflecting ethnic inter-embeddedness should be included to better capture its core characteristics. Third, mixed methods such as surveys, interviews, and social network analysis can enhance validation of causal mechanisms and key variables. Expanding research across temporal, spatial, and social dimensions will provide a more comprehensive understanding of vulnerability formation and support more effective disaster resilience strategies.

References

Adger WN (2006) Vulnerability. Glob Environ Change 16(3):268–281. https://doi.org/10.1016/j.gloenvcha.2006.02.006

Aroca-Jimenez E, Bodoque JM, Garcia JA, Diez-Herrero A (2017) Construction of an integrated social vulnerability index in urban areas prone to flash flooding. Nat Hazards Earth Syst Sci 17(9):1541–1557. https://doi.org/10.5194/nhess-17-1541-2017

Bates S, Angeon V, Ainouche A (2014) The pentagon of vulnerability and resilience: A methodological proposal in development economics by using graph theory. Econ Model 42:445–453. https://doi.org/10.1016/j.econmod.2014.07.027

Birkmann J, Cardona OD, Carreño ML, Barbat AH, Pelling M, Schneiderbauer S, Kienberger S, Keiler M, Alexander D, Zeil P (2013) Framing vulnerability, risk and societal responses: the MOVE framework. Nat Hazards 67(2):193–211. https://doi.org/10.1007/s11069-013-0558-5

Birkmann J, Jamshed A, McMillan JM, Feldmeyer D, Totin E, Solecki W, Ibrahim ZZ, Roberts D, Kerr RB, Poertner H-O, Pelling M, Djalante R, Garschagen M, Leal Filho W, Guha-Sapir D, Alegría A (2022) Understanding human vulnerability to climate change: A global perspective on index validation for adaptation planning. Sci Total Environ 803:150065. https://doi.org/10.1016/j.scitotenv.2021.150065

Blaikie P, Cannon T, Davis I, Wisner B (2014) At risk: natural hazards, people's vulnerability and disasters. Routledge

Bolin R, Stanford L (1991) Shelter, housing and recovery: A comparison of US disasters. Disasters 15(1):24–34. https://doi.org/10.1111/j.1467-7717.1991.tb00424.x

Bolin RC (1994) Household and community recovery after earthquakes. Institute of Behavioral Science University of Colorado

Cardona OD (2013) The need for rethinking the concepts of vulnerability and risk from a holistic perspective: a necessary review and criticism for effective risk management. In: Mapping

vulnerability. Routledge, pp 37–51

Cardona OD, Van Aalst MK, Birkmann J, Fordham M, Mc Gregor G, Rosa P, Pulwarty RS, Schipper ELF, Sinh BT, Décamps H, others (2012) Determinants of risk: exposure and vulnerability. In: Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change. Cambridge University Press, pp 65–108

Carter-Pokras O, Zambrana RE, Mora SE, Aaby KA (2007) Emergency preparedness: Knowledge and perceptions of Latin American immigrants. J Health Care Poor Underserved 18(2):465–481. https://doi.org/10.1353/hpu.2007.0026

Chang H, Pallathadka A, Sauer J, Grimm NB, Zimmerman R, Cheng C, Iwaniec DM, Kim Y, Lloyd R, McPhearson T, Rosenzweig B, Troxler T, Welty C, Brenner R, Herreros-Cantis P (2021) Assessment of urban flood vulnerability using the social-ecological-technological systems framework in six US cities. Sustain Cities Soc 68:102786. https://doi.org/10.1016/j.scs.2021.102786

Chen S, Zou Q, Wang B, Zhou W, Yang T, Jiang H, Zhou B, Yao H (2023) Disaster risk management of debris flow based on time-series contribution mechanism (CRMCD): Nonnegligible ecological vulnerable multi-ethnic communities. Ecol Indic 157:111266. https://doi.org/10.1016/j.ecolind.2023.111266

Chen T (2016) XGBoost: A Scalable Tree Boosting System. Cornell Univ. https://doi.org/10.1145/2939672.2939785

Cutter SL, Boruff BJ, Shirley WL (2012) Social vulnerability to environmental hazards. In: Hazards vulnerability and environmental justice. Routledge, pp 115–132

Cutter SL, Finch C (2008) Temporal and spatial changes in social vulnerability to natural hazards. Proc Natl Acad Sci 105(7):2301–2306. https://doi.org/10.1073/pnas.0710375105

Driel E, Verkuyten M (2022) Gossip, diversity and community cohesion: The case of multi-ethnic Riace. Ethn Racial Stud 45(16):591–613. https://doi.org/10.1080/01419870.2022.2114806

Dwyer A, Zoppou C, Nielsen O, Day S, Roberts S (2004) Quantifying social vulnerability: a methodology for identifying those at risk to natural hazards. Geoscience Australia Canberra

Elsenburg LK, Nicolaou M, Galenkamp H, Lakerveld J, Stronks K (2025) The clustering of disadvantage in different life dimensions across ethnic groups: A network analysis of indicators of precariousness in the HELIUS study. Soc Sci Med 375:117970. https://doi.org/10.1016/j.socscimed.2025.117970

Emrich CT, Cutter SL (2011) Social vulnerability to climate-sensitive hazards in the southern United States. Weather Clim Soc 3(3):193–208. https://doi.org/10.1175/2011WCAS1092.1

Feng L (2022) Seven departments jointly implement the plan for embedded development among ethnic groups. Xinhua Net. https://www.news.cn/politics/2022-09/07/c_1128984738.htm.Accessed 19 October 2025

Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B (2011) A social vulnerability index for disaster management. J Homel Secur Emerg Manag 8(1). https://doi.org/10.2202/1547-7355.1792

Fothergill A, Maestas EG, Darlington JD (1999) Race, ethnicity and disasters in the United States: A review of the literature. Disasters 23(2):156–173. https://doi.org/10.1111/1467-7717.00111

Gabus A, Fontela E (1973) Perceptions of the world problematique: Communication procedure, communicating with those bearing collective responsibility

Gafari O, Bahrami-Hessari M, Norton J, Parmar R, Hudson M, Ndegwa L, Agyapong-Badu S, Asante KP, Alwan NA, McDonough S, others (2024) Building trust and increasing inclusion in public health research: co-produced strategies for engaging UK ethnic minority communities in research. Public Health 233:90–99. https://doi.org/10.1016/j.puhe.2024.05.007

Ge Y, Dou W, Liu N (2017) Planning resilient and sustainable cities: Identifying and targeting social vulnerability to climate change. Sustainability 9(8):1394. https://doi.org/10.3390/su9081394

Goktas P, Rabiei-Dastjerdi H (2026) Artificial intelligence-driven approaches for assessing social vulnerability to natural hazards: a comprehensive bibliometric review. Nat Hazards 122(7):322. https://doi.org/10.1007/s11069-025-07808-w

Graus S, Ferreira TM, Vasconcelos G, Ortega J, Sarmento J (2025a) A novel methodology for assessing social vulnerability in depopulating rural communities: Application to Montesinho Natural Park, Portugal. Int J Disaster Risk Reduct :105815. https://doi.org/10.1016/j.ijdrr.2025.105815

Graus S, Ferreira TM, Vasconcelos G, Ortega J, Sarmento J (2025b) A novel methodology for assessing social vulnerability in depopulating rural communities: Application to montesinho natural park, portugal. Int J Disaster Risk Reduct :105815. https://doi.org/10.1016/j.ijdrr.2025.105815

Guillard-Gonçalves C, Cutter SL, Emrich CT, Zêzere JL (2015) Application of Social Vulnerability Index (SoVI) and delineation of natural risk zones in Greater Lisbon, Portugal. J Risk Res 18(5):651–674. https://doi.org/10.1080/13669877.2014.910689

Hilhorst D, Bankoff G (2013) Introduction: mapping vulnerability. In: Mapping vulnerability. Routledge, pp 1–9

Holand IS, Lujala P, Rød JK (2011) Social vulnerability assessment for Norway: A quantitative approach. Nor Geogr Tidsskr-Nor J Geogr 65(1):1–17. https://doi.org/10.1080/00291951.2010.550167

Huang C-Y, Shyu JZ, Tzeng G-H (2007) Reconfiguring the innovation policy portfolios for Taiwan's SIP Mall industry. Technovation 27(12):744–765. https://doi.org/10.1016/j.technovation.2007.04.002

Hung KK, MacDermot MK, Chan EY, Liu S, Huang Z, Wong CS, Walline JH, Graham CA (2021) CCOUC Ethnic Minority Health Project: a case study for Health EDRM initiatives to improve disaster preparedness in a rural Chinese population. Int J Environ Res Public Health 18(10):5322. https://doi.org/10.3390/ijerph18105322

Jia Y, Jia S, Pang R (2024) Measurement and impact analysis of social vulnerability to geological disasters in Guizhou Province, China. Sci Rep 14(1):25098. https://doi.org/10.1038/s41598-024-76705-2

Li J (2024) Social work innovation and progress of national unity in multi-ethnic areas: from the perspective of inter-embedded community building. Int J Front Sociol 9(6):41–49. https://doi.org/10.25236/IJFS.2024.060907

Kok M, De Coninck H (2007) Widening the scope of policies to address climate change: directions for mainstreaming. Environ Sci Policy 10(7–8):587–599. https://doi.org/10.1016/j.envsci.2007.07.003

Li L, Xia R, Dou M, Zhang K, Chen Y, Jia R, Li X, Dou J, Li X, Hu Q, others (2024) Integrated machine learning reveals aquatic biological integrity patterns in semi-arid watersheds. J Environ Manage 359:121054. https://doi.org/10.1016/j.jenvman.2024.121054

Li Y, Cheng W, Zhang Y, Stanny M, Du G (2025) Mapping urban-rural divide of disaster vulnerability and implications for community resilience building in China. Habitat Int 165:103580.

https://doi.org/10.1016/j.habitatint.2025.103580
Liang LH (2025) The generation of governance community: governance innovation in multi-ethnic inter-embedded communities. J Hubei Minzu Univ Philos Soc Sci 43(6):111–120. https://doi.org/10.13501/j.cnki.42-1328/c.20251029.001
Linh NT, others (2025) Enhancing the reliability of social vulnerability assessment to natural hazards through combined machine learning methods: A case study in Vietnam. Nat Hazards Res. https://doi.org/10.1016/j.nhres.2025.09.004
Liu ZL, Liao L, Niu CL (2015) The impact of community social capital on residential satisfaction: an empirical analysis based on a survey of low- and middle-income communities in Beijing. Hum Geogr 30(3):21–27. https://doi.org/10.13959/j.issn.1003-2398.2015.03.004
Lundberg SM, Lee S-I (2017) A unified approach to interpreting model predictions. Adv Neura l Inf Process Syst 30. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.Accessed 16 May 2026
Ma L, Liu S, Fang F, Che X, Chen M (2020) Evaluation of urban-rural difference and integration based on quality of life. Sustain Cities Soc 54:101877. https://doi.org/10.1016/j.scs.2019.101877
Ma W, Jiang H, Bao Y, Liu G (2025) The sense of community for the Chinese nation among different ethnic groups: The role of perceived ethnic group continuity and ethnic group identity. Int J Intercult Relat 108:102222. https://doi.org/10.1016/j.ijintrel.2025.102222
McQueen JB (1967) Some methods of classification and analysis of multivariate observations. In: Proc. of 5th Berkeley Symposium on Math. Stat. and Prob. pp 281–297
Nachappa TG, Piralilou ST, Gholamnia K, Ghorbanzadeh O, Rahmati O, Blaschke T (2020) Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory. J Hydrol 590:125275. https://doi.org/10.1016/j.jhydrol.2020.125275
National Bureau of Statistics of China (2023) China statistical yearbook 2023. China Statistics Press, Beijing. https://www.stats.gov.cn/sj/ndsj/2023/indexch.htm.Accessed 20 October 2025
Nguyen DH, Le XH, Heo J-Y, Bae D-H (2021) Development of an extreme gradient boosting model integrated with evolutionary algorithms for hourly water level prediction. Ieee Access 9:125853–125867. https://doi.org/10.1109/ACCESS.2021.3111287.
Nguyen-Thi Tra, M., Trinh, T.-A., & Nguyen-Ngoc, D. (2026). The effects of home-country natural disasters on refugee mental health. Population and Environment, 48(2), 11. https://doi.org/10.1007/s11111-026-00525-x
Ouvrard C, Jiatsa IK, Gbessemehlan A, Simon Y, Letenneur L, Dartigues J-F, Andrew MK, Amieva H, Pérès K (2025) Social vulnerability increases the risk of death differently in men and women: longitudinal analysis over 15 years in the Paquid Study. Arch Gerontol Geriatr :106079. https://doi.org/10.1016/j.archger.2025.106079
Peacock W, Morrow BH, Gladwin H (1997) Hurricane Andrew and the reshaping of Miami: Ethnicity, gender, and the socio-political ecology of disasters. Gainsville FL Univ Press Fla
Perry RW, Green MR (1982) The role of ethnicity in the emergency decision-making process. Sociol Inq 52(4):306–334. https://doi.org/10.1111/j.1475-682X.1982.tb01257.
Rong X, Zhai B, Sun R, Li R, Jing H, Meriggi M, Semprebon G (2025) A full-chain assessment and prediction of social-ecological system vulnerability and its driving factors in multi-ethnic regions: Evidence from southwestern Gansu, China. Ecol Indic 178:114105. https://doi.org/10.1016/j.ecolind.2025.114105

Song WH (2013) The dilemma of residents' participation cognition in urban community cultural construction and its resolution. Adm Forum 20(4):89–92. https://doi.org/10.16637/j.cnki.23-1360/d.2013.04.008

Tang N, Cheng L, Cai C (2020) Making collective policy entrepreneurship work: the case of China's post-disaster reconstruction. J Asian Public Policy 13(1):60–78. https://doi.org/10.1080/17516234.2019.1701388

Tate E, Rahman MA, Emrich CT, Sampson CC (2021) Flood exposure and social vulnerability in the United States. Nat Hazards 106(1):435–457. https://doi.org/10.1007/s11069-020-04470-2

Tian Z, Li Q (2024) Exploring the ethnic cultural integration path of immigrant communities based on ethnic inter-embedding. Economics 18(1):20220090. https://doi.org/10.1515/econ-2022-0090

Timmerman P (1981) Vulnerability, resilience and the collapse of society: a review of models and possible climatic applications

Turner BL, Kasperson RE, Matson PA, McCarthy JJ, Corell RW, Christensen L, Eckley N, Kasperson JX, Luers A, Martello ML, others (2003) A framework for vulnerability analysis in sustainability science. Proc Natl Acad Sci 100(14):8074–8079. https://doi.org/10.1073/pnas.1231335100

Tzeng G-H, Chiang C-H, Li C-W (2007) Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL. Expert Syst Appl 32(4):1028–1044. https://doi.org/10.1016/j.eswa.2006.02.004

Wei J, Li Y, Jia L, Liu B, Jiang Y (2025) The Impact of Spatiotemporal Effect and Relevant Factors on the Urban Thermal Environment Through the XGBoost-SHAP Model. Land 14(2):394. https://doi.org/10.3390/land14020394

Wen M, Lauderdale DS, Kandula NR (2009) Ethnic neighborhoods in multi-ethnic America, 1990–2000: Resurgent ethnicity in the ethnoburbs? Soc Forces 88(1):425–460. https://doi.org/10.1353/sof.0.0244

Wisner B, Wisner B (2004) At risk: natural hazards, people's vulnerability and disasters. Psychology Press

Xiaona Z, Busababodhin P, Suraphee S (2026) Exploring the drivers of satisfaction in multi-ethnic inter-embedded urban communities:Towards sustainable urban development. Sustain Futur 11:101738. https://doi.org/10.1016/j.sftr.2026.101738

Zhou Y, Li N, Wu W, Wu J (2014) Assessment of provincial social vulnerability to natural disasters in China. Nat Hazards 71(3):2165–2186. https://doi.org/10.1007/s11069-013-1003-5