ARTICLE

A Data-Driven Geospatial Framework for Assessing the Persistence Vulnerability of Traditional Villages in Ethnic Minority Areas: Evidence from 1325 Villages in Sichuan, China

Zhiming Mo[1], Yu Tang[2], Xiaobi Jiang[3], Feng Gong[3], Yong Liao[4]*

1. School of Culture and Tourism, Chongqing City Management Vocational University, China
2. School of General Education, Chongqing City Management Vocational University, China
3. School of Culture and Tourism, Chongqing Vocational College of Culture and Arts, China
4. School of Culture and Tourism, Chongqing Youth Vocational and Technical College, China
* Email: liaoyong460@gmail.com

DOI: https://doi.org/10.64376/xm426w17

Abstract: Traditional villages are the spatial carriers of living ethnic minority culture, yet vulnerability assessments rarely ask whether villages can reach cultural resources nearby. This study develops a reproducible, data-driven geospatial assessment pipeline that links village coordinates, open heritage registers, terrain data and statistical yearbooks, and applies it to 1325 traditional villages at or above the provincial level in the 21 cities (prefectures) of Sichuan Province, China. Within the vulnerability scoping diagram (VSD) framework, exposure is split into disturbance exposure (B1) and structural exposure (B2); the latter is measured by village-level nearest distances to heritage protection units, intangible cultural heritage (ICH) items and A-class scenic areas. Twenty-three indicators are synthesised with entropy weighting into a persistence vulnerability index (VI), and kernel density estimation, gravity-centre analysis and K-means clustering validated by silhouette and Davies–Bouldin indices are used to characterise spatial patterns and village types. Mean nearest distances to the three resource types are 14.6, 18.7 and 9.2 km, highest in Liangshan and lowest in Chengdu. VI shows a steep gradient (Ganzi 1.000, Aba 0.829, Liangshan 0.428; all other cities < 0.028), and 92.6% of provincial inequality lies between the five economic zones (Theil index 1.536). Accessibility covaries with VI (r = 0.602) and with the terrain score (r = 0.698), and the association persists in the 18-city basin subsample (r = 0.727); because B2 is a component of VI, no independent effect is inferred. The 21 prefectures are grouped into four exploratory management zones. By framing vulnerability assessment as a data-fusion and unsupervised-learning task, the pipeline provides a transparent, updatable and interpretable decision-support layer for heritage-oriented rural governance.

Keywords: Traditional villages; Persistence vulnerability; Geospatial analytics; Unsupervised learning; Intelligent spatial decision support; Cultural resource accessibility; Ethnic minority areas

1 Introduction

1.1 Problem statement

Traditional villages are social communities woven together by kinship, locality and occupational ties. They are also living carriers of local knowledge, building crafts and ritual traditions [1–3]. As rural revitalization advances, they carry two roles at once: object of protection and resource for development [4–6]. Village persistence is therefore no longer only a question of heritage protection. It has become a governance issue produced jointly by population, industry, public services and cultural reproduction.

Sichuan is one of the provinces with the greatest variety of traditional village types and the densest superposition of ethnic cultures in China. Its 1325 villages at or above the provincial level span four geomorphological units — the basin, the basin-margin mountains, the northwestern Sichuan plateau and the Panxi rift — and are home to Han, Tibetan, Qiang, Yi, Miao, Tujia and other peoples. Northwestern Sichuan and Panxi are both the core section of the Tibetan-Qiang-Yi corridor and the part of the province with the lowest public service accessibility and the most severe population outflow [7–9]. Rich cultural resources combined with weak public services: this spatial coincidence is the entry point for understanding village vulnerability in ethnic minority areas.

Existing studies fall roughly into three strands. The first addresses spatial patterns and landscape genes, revealing how village distribution follows terrain, river basins and cultural areas [10–12]. The second

assesses vulnerability and resilience, mostly measuring disturbance from hazards, tourism and urbanization with VSD, PSR or social-ecological system frameworks [13–18]. The third discusses revitalization paths and governance mechanisms, including differentiated measures, neo-endogenous development and culture-tourism integration [6,19–21].

In parallel, intelligent and data-driven methods have changed how spatial problems of this kind can be approached. GeoAI and spatially explicit machine learning combine geographic data with computational learning to extract patterns from large point and raster data sets [22,23], and spatial decision support systems (SDSS) turn such analyses into repeatable decision aids for planners [24]. In the heritage field, smart-village and digital-twin initiatives have begun to integrate monitoring data into rural development and conservation practice [25,26]. Yet most vulnerability assessments of traditional villages remain one-off indicator exercises: data assembly, weighting, classification and zoning are seldom specified as a transparent workflow that a provincial agency could audit and rerun as new data arrive.

Computational approaches are also entering traditional-village research itself. Deep learning has been used to identify traditional villages of different cultural types from imagery [27], digital-media data have been used to trace the spatiotemporal change and place-making of villages [28], and digital platforms have been used to optimise environmental comfort and landscape visibility in individual villages [29]. From an intelligent-systems perspective, a provincial vulnerability assessment can be formulated as a data-fusion and unsupervised-learning problem: heterogeneous inputs (point registers, terrain rasters, statistical tables and weather series) must be harmonised into a common feature space; indicator weights can be learned from the information content of the data rather than set by experts, as in Shannon-entropy weighting [30]; latent village types can be discovered by clustering and checked with internal validity indices [31–33]; and the outputs must be robust and interpretable enough to support decisions [34]. This study adopts that formulation.

Three gaps follow. First, the spatial accessibility of cultural resources has not yet entered vulnerability measurement. Heritage protection units, ICH items and A-class scenic areas are strongly concentrated towards central cities, while villages remain in peripheral mountains. Existing studies mostly describe resource density or centroid deviation qualitatively and do not convert "whether a village can obtain cultural resources nearby" into a comparable vulnerability dimension. Spatial mismatch theory originates in explanations of the separation between urban jobs and residence [35,36] and has been applied in the heritage field in recent years [37], but its distance metric has not been connected to vulnerability measurement. Second, there is little quantitative evidence on whether cultural enrichment in ethnic minority areas can be converted into resilience. Ethnography emphasizes villager subjectivity, cultural self-consistency and embedded relations [1,38,39], but these mechanisms are presented qualitatively and can hardly support resource allocation decisions at the provincial scale. Third, existing assessments seldom document their processing chain in a form that can be re-run, audited and updated, which limits their use as decision-support tools.

1.2 Research aims and analytical positioning

This study advances one confirmatory descriptive proposition: **P1**, that vulnerability is strongly spatially differentiated at the provincial scale, that the ethnic minority areas of northwestern Sichuan and Panxi are higher than the basin, and that between-group differences dominate total differences.

This study sets no hypothesis on whether cultural accessibility has an independent effect. The reason is given in Section 3.5: the accessibility score is a component of the vulnerability index, so the two are definitionally linked; ScoreB is in turn a discretization of the same set of distances and correlates with B2 at 0.962. With a sample of 21 prefectures, this constructive association cannot be separated from a real effect. Accessibility is therefore used as a descriptive dimension of structural exposure, for mapping, zoning and policy dialogue, not for effect inference.

The study makes three contributions. First, it specifies a modular geospatial assessment pipeline (Fig. 1) that converts village registers, cultural resource lists, terrain data and statistical yearbooks into indicators, a composite index, data-driven village typologies and a zoning matrix, with cluster-validity and specification-

sensitivity checks built in, so that the assessment can be audited and repeated as data are updated. Second, cultural resource accessibility is measured by village-level nearest distance, which gives the dimension an unambiguous adverse direction, while regional size imbalance is kept separate as a direction-neutral covariate; in this way the anthropological propositions of cultural niche and cultural subjectivity are translated into describable spatial indicators and carried through to the zoning. Third, a basin subsample, two distance metrics and five index constructions are used to test whether the mapping results are driven by extreme cases or by analytical choices.

2 Study area, data and methods

2.1 Analytical workflow

The analysis is organised as a four-layer pipeline (Fig. 1). The data layer assembles the 1325 village points, three cultural resource registers, a digital elevation model, statistical yearbooks for 2021–2025 and daily weather records. The feature-engineering layer computes village-level nearest distances, kernel density surfaces, gravity centres and 23 standardised and direction-aligned indicators. The modelling layer derives entropy weights within each dimension, aggregates them into the composite index (Eq. 1) and builds data-driven village typologies with K-means clustering validated by the silhouette coefficient and the Davies–Bouldin index [31–33]. The decision-support layer decomposes provincial inequality, examines the sensitivity of the results to index construction [34] and maps the prefectures into a two-dimensional zoning matrix linked to measure packages. Each layer takes the outputs of the previous one as inputs, so the whole chain can be re-run when any data source is updated.

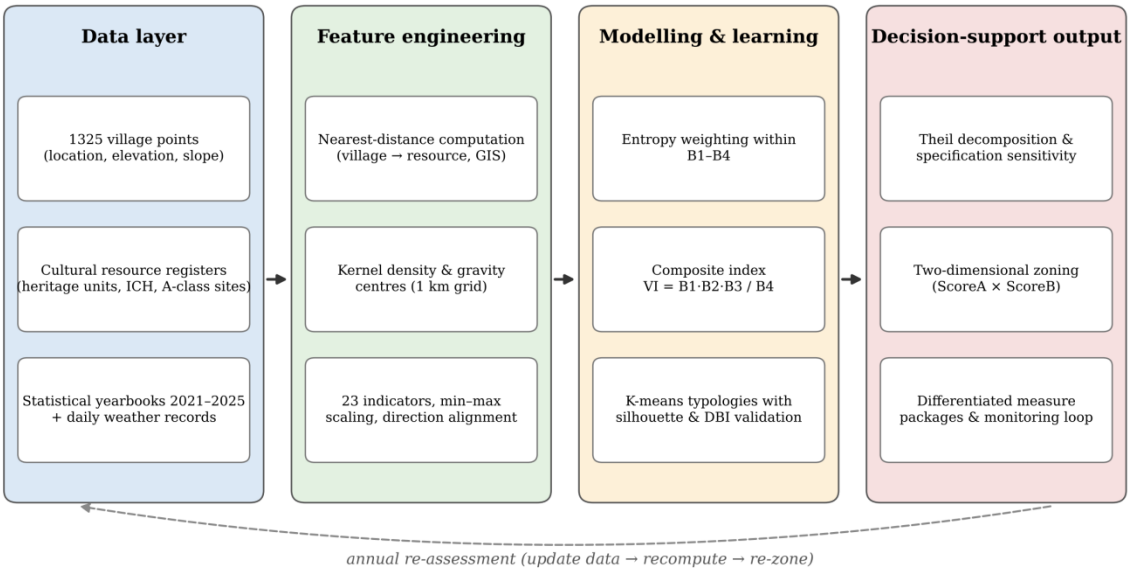


Fig. 1 Data-driven assessment pipeline for the persistence vulnerability of traditional villages

2.2 Study area and data

Sichuan Province lies between 97°21′-108°33′E and 26°03′-34°19′N, with an elevation range exceeding 7000 m. Settlement forms are closely coupled with ethnic cultural types: northwestern Sichuan is dominated by Tibetan and Qiang watchtower houses and fortified villages, Panxi is represented by Yi flat-roofed and shingle-roofed houses, and eastern and southern Sichuan retain relatively intact Han lineage settlements [40].

Villages are most numerous in Ganzi (302), Aba (170) and Bazhong (128), and fewest in Ziyang (6), Panzhihua (9) and Suining (14). The three prefectures of Ganzi, Aba and Liangshan together hold 506 villages, 38.2% of the provincial total, and form the ethnic minority sample on which this study focuses. Sichuan is also a meteorological disaster-prone area, where extreme precipitation and snow disasters continuously threaten the accessibility of high-altitude villages [41,42].

Data fall into five categories. Village points (longitude, latitude, elevation, slope) were taken from the traditional village lists published by the Ministry of Housing and Urban-Rural Development and by Sichuan Province; cultural resource points (heritage protection units, ICH items, A-class scenic areas) were taken from the public lists of the competent cultural and tourism authorities. The nearest distances from villages to the three resource types were computed point by point for all 1325 villages and then averaged arithmetically by prefecture (Supplementary Table S1) [Author note: the village-level distance table and Appendix A cited in the original submission were not provided; please supply them as supplementary material.]. Kernel density and the gravity model were produced in ArcGIS 10.8. Socioeconomic data were taken from the Sichuan Statistical Yearbook 2021-2025 and from prefecture bulletins, and five-year means were used to smooth interannual fluctuation; daily weather data were used to compute the index of adverse-weather travel impact.

Pre-processing had three steps. First, min-max standardization. Second, direction alignment. Of B1-B3, the three items with direction "−" (C13 cultural resource density, C14 village density, C15 share of primary school students) were transformed as $1-x'$, so that a larger value represents a stronger contribution to vulnerability. The seven "−" indicators of B4 were not reversed; their standardized values represent adaptive capacity directly, which corresponds to the position of B4 in the denominator of Eq. (1). Third, the three cultural location indicators (C8-C10) are themselves "the farther the worse", and enter the dimension as standardized values without further sign handling. All computations are documented in the Supplementary Material.

2.3 Conceptual definitions and the two-part structure of exposure

The core concepts are defined as follows. Persistence vulnerability is the degree to which traditional villages fall short of maintaining intact chains of population, livelihood and cultural reproduction under the combined action of external stress and internal fragility; its opposite is not "no development" but "rupture in the course of development". Cultural capital is the cultural and tourism resources that can provide villages with protection resources, market opportunities and identity support, measured here through three carriers: heritage protection units, ICH items and A-class scenic areas. A-class scenic areas are closer to tourism market resources, so the three are collectively termed cultural tourism resources. Cultural location is the spatial accessibility of a village relative to those resources, represented by nearest distance. Cultural niche is the position of a village within the resource-market-institution network in which it is embedded; the greater the distance, the more isolated the niche.

The VSD framework decomposes vulnerability into exposure, sensitivity and adaptive capacity, and is widely used for cross-case comparison because of its "meta-framework" character [43,44]. This study does not add a fourth VSD dimension. Instead, exposure is split into two blocks: disturbance exposure (B1) and structural exposure (B2).

The reason for this split is that the two kinds of pressure operate on different time scales. B1 measures instantaneous, exogenous disturbance such as tourism shocks, population outflow and extreme weather; B2 measures persistent, endogenous spatial conditions, that is, whether a village can obtain cultural resources nearby. Neither belongs to sensitivity or adaptive capacity — the first is the shock itself, the second is the distributional condition of the shock — so both are assigned to exposure. Mixing them in one weighted score would let them dilute each other because of their different mechanisms; reporting them in blocks preserves their separate explanatory power.

To facilitate comparison with earlier work, the numbering B1-B4 and the phrase "four scores" are retained below. The position of B2, however, is that of a structural component of exposure, not a new dimension independent of the three VSD elements.

Once resources are spatially distant from a village, the effect is not a one-off shock. Bearers of transmission struggle to obtain project support, crafts struggle to enter the market, ritual spaces struggle to obtain repair, the chain of cultural reproduction weakens, and the system threshold keeps shifting down [35,45]. This is also why it is listed here as structural exposure rather than merged into disturbance exposure.

The functions of the four score blocks are as follows. Disturbance exposure (B1) is the input signal of external stress. Structural exposure (B2) is the spatial condition of accessibility to cultural resources. Sensitivity (B3) is the internal fragility of the system, covering the natural base, the cultural niche and the generational structure. Adaptive capacity (B4) is the buffering mechanism. The composite vulnerability index is synthesized as follows:

$$VI_{\text{raw}}=\frac{B_1\times B_2\times B_3}{B_4} \tag{1}$$

Eq. (1) is multiplicative, so vulnerability tends to zero if any dimension approaches zero, that is, "no exposure, no vulnerability". This accords with the theoretical meaning of vulnerability as a conditional state [46,47], but it also produces boundary effects and compression at the low end (see Section 4.3). $VI_{(raw)}$ is min-max normalized and denoted VI.

2.4 Indicator system and entropy weighting

The indicator system has three tiers and 23 indicators (Table 1). B1 to B4 contain 7, 3, 6 and 7 indicators respectively. The share, information entropy and weight in the entropy weighting method are:

$$p_{ij}=\frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}},\qquad e_j=-k\sum_{i=1}^{n} p_{ij}\ln p_{ij},\qquad w_j=\frac{1-e_j}{\sum_{j=1}^{m}(1-e_j)} \tag{2}$$

Here $k=1/\ln n$ and $n=21$. Entropy weighting assigns weights objectively and avoids subjective bias. Its weights reflect the amount of information rather than mechanistic importance, so an indicator with little dispersion may receive a low weight even if it matters theoretically [48]. To make dimensions comparable, entropy weights are computed separately within each of B1-B4, and the weighted sum within a dimension serves as the dimension score. [30]

Table 1 Indicator system for assessing the persistence vulnerability of traditional villages (2021-2025 annual means)

Criterion layer	Code	Indicator and calculation	Direction	Data source	Theoretical meaning
Exposure B1	C1	Total tourism revenue/gross regional product × 100%	+	Statistical yearbook	Economic dependence on tourism; a high share brings over-commercialization and loss of authenticity
	C2	Total tourist arrivals/year-end permanent resident population	+	Statistical yearbook	Intensity of external visitors, reflecting carrying pressure on facilities and daily life
	C3	(Registered population − permanent resident population)/registered population × 100%	+	Statistical yearbook	Intensity of population outflow, corresponding to hollowing out and loss of bearers of transmission
	C4	Commercial housing floor space sold/year-end permanent resident population	+	Statistical yearbook	Encroachment of urban construction land on village space
	C5	Rural minimum living allowance recipients/year-end permanent resident population × 100%	+	Statistical yearbook	Proxy for poverty depth; a high share implies a weak economic buffer
	C6	Highway passenger turnover/year-end permanent resident population	+	Statistical yearbook	Intensity of population mobility; excessive mobility tends to dismantle community networks
	C7	(Annual snow days + annual rain days)/365 × 100%	+	Weather stations	Blocking of accessibility by adverse weather, a seasonal natural constraint

Criterion layer	Code	Indicator and calculation	Direction	Data source	Theoretical meaning
Cultural location B2	C8	Prefecture mean distance from villages to the nearest heritage protection unit/km	+	Point proximity analysis	Cost of obtaining protection resources; the farther the distance, the harder to obtain repair and project support
	C9	Prefecture mean distance from villages to the nearest ICH item/km	+	Point proximity analysis	Accessibility of living cultural resources; the farther the distance, the looser the transmission tie
	C10	Prefecture mean distance from villages to the nearest A-class scenic area/km	+	Point proximity analysis	Degree of spatial separation from the tourism market, affecting visitor intake and benefit sharing
Sensitivity B3	C11	Arithmetic mean elevation of villages within the prefecture/m	+	DEM extraction	High elevation means ecological fragility and difficult infrastructure, hence a high baseline of fragility
	C12	Mean slope of villages within the prefecture/(°)	+	DEM extraction	Steep slopes raise hazard risk and constrain construction and transport
	C13	(ICH + heritage protection units + scenic areas)/administrative area (units/100 km²)	−	Lists + GIS	The higher the density, the richer the regional cultural-ecological network
	C14	Traditional villages/administrative area (villages/100 km²)	−	Lists + GIS	The higher the density, the tighter the collaborative protection network among villages
	C15	Primary school students/year-end permanent resident population × 100%	−	Statistical yearbook	Size of the young generation; the lower the share, the higher the risk of broken transmission
	C16	Urban permanent resident population/year-end permanent resident population × 100%	+	Statistical yearbook	The higher the urbanization, the greater the pressure of population and land absorption
Adaptive capacity B4	C17	Per capita disposable income of rural residents/yuan	−	Statistical yearbook	Household economic strength, the fundamental guarantee for resisting shocks and investing in protection
	C18	Gross output value of agriculture, forestry, animal husbandry and fishery/administrative area (100 million yuan/100 km²)	−	Statistical yearbook	Regional agricultural output capacity, the basic buffer for maintaining livelihoods
	C19	Local general public budget expenditure/100 million yuan	−	Statistical yearbook	Fiscal basis for public services and for investment in heritage protection
	C20	Balance of deposits in financial institutions/year-end permanent resident population (yuan/person)	−	Statistical yearbook	Household stock wealth, an economic cushion accumulated over the long term
	C21	Internet broadband subscribers/total households × 100%	−	Statistical yearbook	Capacity for information access and external connection
	C22	Beds in medical and health institutions/year-end permanent resident population ×10000	−	Statistical yearbook	Level of basic public services, a core component of the social safety net
	C23	Highway mileage/administrative area (km/km²)	−	Statistical yearbook + GIS	Transport accessibility, affecting material flows, tourist entry and emergency response

Note: A direction of "+" means that a larger value indicates higher vulnerability, and "−" means that a larger value indicates lower vulnerability. C8-C10 are village-level nearest distances, not centroid distances.

2.5 Spatial analysis: kernel density, gravity centres and regional size imbalance

Kernel density estimation is used to characterize the continuous distribution intensity of features:

$$\hat{f}(x)=\frac{1}{n\mathrm{h}^2}\sum_{i=1}^{n}K\left(\frac{x-x_i}{\mathrm{h}}\right) \tag{3}$$

Here h is the bandwidth, $K(\cdot)$ is the kernel function, the bandwidth is set by Silverman's rule of thumb, and the output is a 1 km grid. The gravity model is used to identify the distribution centre of features:

$$\overline{X}=\frac{\sum_i w_i x_i}{\sum_i w_i}, \qquad \overline{Y}=\frac{\sum_i w_i y_i}{\sum_i w_i} \tag{4}$$

Feature counts are used as weights: prefecture centroids are computed first, and the provincial centroid is then synthesized from them. On the basis of centroid deviation, a regional mismatch index is constructed:

$$SMI_k=D_k\times G_k, \qquad D_k=\mathrm{sgn}(\Delta x_k)\ \frac{d_k}{d_{\max}}, \qquad G_k=1-\frac{|s_k^v-s_k^r|}{\max\{s_k^v,s_k^r\}} \tag{5}$$

Here d_k is the Euclidean distance between the village centroid and the resource centroid of prefecture k, and Δx_k is the longitudinal difference, which takes $+1$ when the village centroid lies west of the resource centroid and -1 when it lies east; s^v_k and s^r_k are the shares of villages and of resources in the provincial total; G_k ranges from 0 to 1 and is the size-match score.

Two points need emphasis. First, SMI_k is a Euclidean distance metric based on centroids. It reflects only distributional deviation at the regional scale and is not equivalent to the travel distance from villages to resources; detours in mountain road networks and terrain resistance make actual costs higher than straight-line distances. Second, the sign of SMI_k indicates direction rather than intensity: the same value may come either from "many villages, few resources" or from "many resources, few villages", and the policy implications of the two are opposite. This study therefore does not use SMI_k as a constituent dimension of vulnerability; it serves only the regional-scale characterization in Figs. 6–8. The cultural location dimension that enters the index is instead carried by village-level nearest distances (C8-C10 in Table 1).

Regional size imbalance is listed separately as a covariate. The size imbalance index is defined as

$$\mathrm{RAI}=\frac{1}{3}\sum_{r=1}^{3}(1-G_r) \tag{6}$$

Here G_r is the size-match score of the three resource types: heritage protection units, ICH items and A-class scenic areas. RAI is direction-neutral. It measures only the degree of imbalance between the regional resource share and the village share, presumes neither benefit nor harm, and enters the regression as an independent explanatory variable.

2.6 Village typology

K-means clustering ($k=3$) is run separately on two key variables, minimizing the within-cluster sum of squares by iteration: [31]

$$\arg\min_S \sum_{j=1}^{k}\sum_{x\in S_j}\left\|x-\mu_j\right\|^2 \tag{7}$$

The number of clusters is determined by the elbow method (Figs. 9–10), and quality is assessed with the silhouette coefficient and the Davies-Bouldin index (DBI): [32,33]

$$s(i)=\frac{b(i)-a(i)}{\max\{a(i),b(i)\}}, \qquad s\in[-1,1] \tag{8}$$

Here $a(i)$ is the mean distance between sample i and samples in the same cluster, and $b(i)$ is the mean distance to samples in the nearest other cluster. Scheme A takes village elevation as input and distinguishes a

low-elevation valley type, a mid-elevation hill type and a high-elevation mountain type; Scheme B takes the distance from a village to the nearest cultural resource as input and distinguishes a core-radiation zone, an adjacent-linkage zone and a peripheral-isolated zone.

The SSE curve of Scheme B shows no clear break between k=3 and k=4, and the original elbow plot is labelled k=4. This study adopts k=3 because it corresponds to the three-tier accessibility design of "core radiation-adjacent linkage-peripheral isolation", and because a silhouette coefficient of 0.584 and a DBI of 0.534 indicate that the three clusters are separable. The alternative k = 4 solution is reported in Supplementary Table S2. [Author note: the k = 4 result announced in the original text is not reported anywhere; please add it.]

2.7 Prefecture aggregation and the descriptive regression

To raise village-level type information to the prefecture level, scores of 3, 2 and 1 are assigned to high, medium and low stress respectively:

$$\text{ScoreA}=\frac{3N_H+2N_M+1N_L}{N_{\text{total}}},\qquad \text{ScoreB}=\frac{3N_E+2N_N+1N_C}{N_{\text{total}}} \tag{9}$$

Here N_H,N_M,N_L are the numbers of high-elevation mountain, mid-elevation hill and low-elevation valley villages, and N_E,N_N,N_C are the numbers of peripheral-isolated, adjacent-linkage and core-radiation villages. Cluster separability is described by analysis of variance, with effect size reported:

$$F=\frac{\frac{\text{SS}_{\text{between}}}{(k-1)}}{\frac{\text{SS}_{\text{within}}}{(n-k)}},\qquad \eta^2=\frac{\text{SS}_{\text{between}}}{\text{SS}_{\text{total}}} \tag{10}$$

This test uses the clustering input variable as the grouping basis. It is a descriptive indicator of clustering quality, is circular, and serves neither as proof of validity nor as causal inference. The driver structure is estimated by multiple regression:

$$VI=\beta_0+\beta_1\cdot\text{ScoreA}+\beta_2\cdot\text{ScoreB}+\varepsilon \tag{11}$$

OLS is used, and standardized coefficients, R^2, adjusted R^2 and VIF are reported. Given that the sample contains only 21 observations, leverage diagnostics are added, using the diagonal elements of the hat matrix and Cook's distance to identify influential points:

$$\text{h}_{ii}=x_i^{\text{T}}\left(X^{\text{T}}X\right)^{-1}x_i,\qquad D_i=\frac{e_i^2}{p\cdot\text{MSE}}\cdot\frac{\text{h}_{ii}}{(1-\text{h}_{ii})^2} \tag{12}$$

The sources of provincial variance are decomposed with the Theil index:

$$T=\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\bar{y}}\ln\frac{y_i}{\bar{y}},\qquad T=T_{\text{b}}+T_{\text{w}} \tag{13}$$

Here y_i is the VI of prefecture i, *bar(y)* is the provincial mean, g indexes the five economic planning zones, n_g and *bar(y)*$_g$ are the number of units and the mean within group g, and T_g is the within-group Theil index. T_b/T is the share of between-group variance.

Five robustness checks are then carried out: re-estimation after removing high-leverage observations; additive aggregation, VI_add = [B1 + B2 + B3 + (1 − B4)]/4; geometric-mean aggregation; the structural ratio with B2 removed, VI = *B1·B3/B4; and the structural ratio with B3 removed, VI = B1·B2/B4* [34]. Rank consistency is compared with Spearman rank correlation, and the significance level is α = 0.05. Computations were performed in SPSSPRO (V1.0.11) and in custom scripts; spatial analysis and mapping were performed in ArcGIS 10.8. [Author note: please add a Monte Carlo uncertainty analysis of weights, normalisation and aggregation, and report rank intervals for each prefecture.]

3 Results

3.1 Spatial pattern and regional separation

Traditional villages are dense in the east and sparse in the west (Fig. 2) [Author note: the kernel-density legends in Figs. 2–5 (0.001–0.034) do not match the densities quoted here (> 0.8, < 0.2, > 1.0); please

harmonise units or values.]. The Chengdu Plain forms a high-value core with kernel density above 0.8, while Bazhong and Dazhou in northeastern Sichuan form a second-order cluster and large parts of Ganzi and Aba in northwestern Sichuan fall below 0.2. This pattern is consistent with the basin-plateau terrain gradient, and with a locational tradition in which villages long depended on farming radius, water sources and building materials [10].

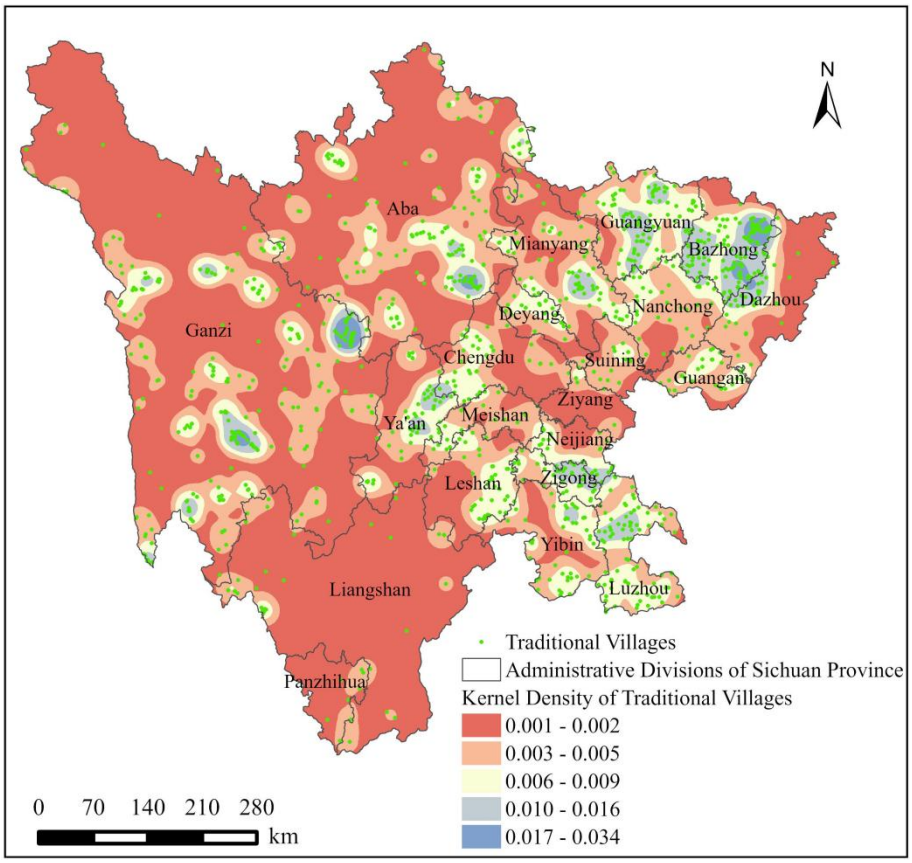


Fig. 2 Kernel density of traditional villages in Sichuan

Cultural resources follow a different logic. Heritage protection units cluster along the Chengdu-Deyang-Mianyang line and the Leshan-Yibin river corridor (Fig. 3). ICH items centre on the Chengdu Plain, where density exceeds 1.0, with a secondary centre in southern Sichuan (Fig. 4); Ganzi and Aba hold 23.7% of the provincial ICH items but only 9.2% of the heritage protection units, a combination of "rich living heritage, scarce material heritage" [37,49]. A-class scenic areas concentrate in the Chengdu metropolitan area and the Leshan-Emei corridor, with only isolated sites such as Jiuzhaigou and Huanglong in northwestern Sichuan (Fig. 5).

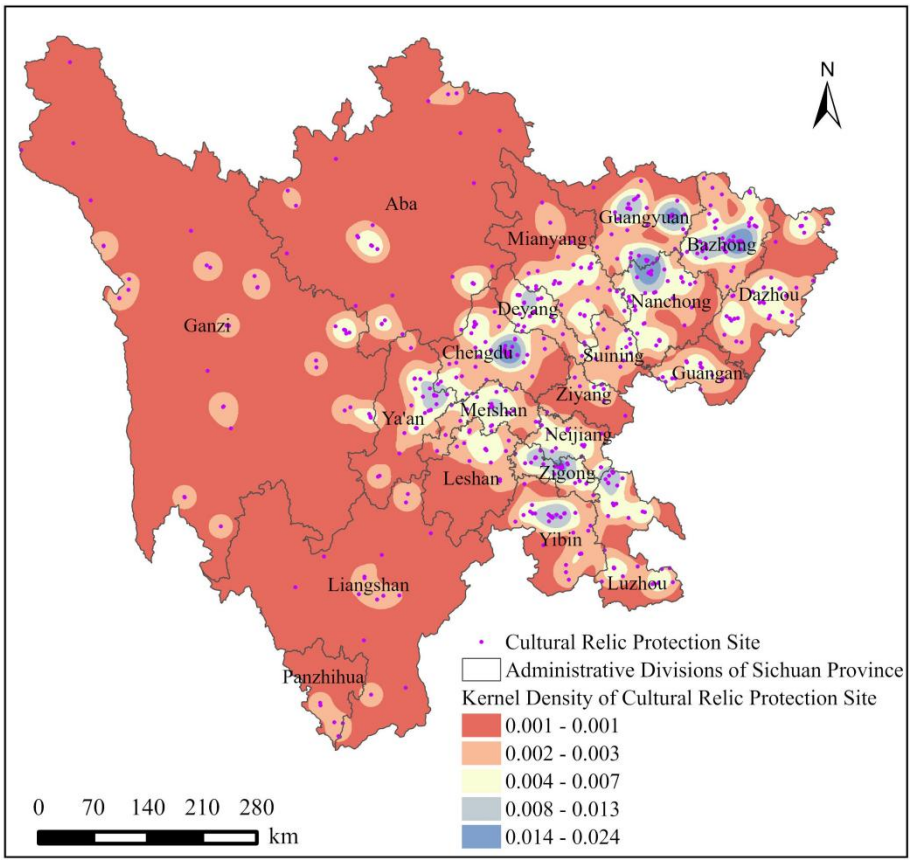


Fig. 3 Kernel density of heritage protection units

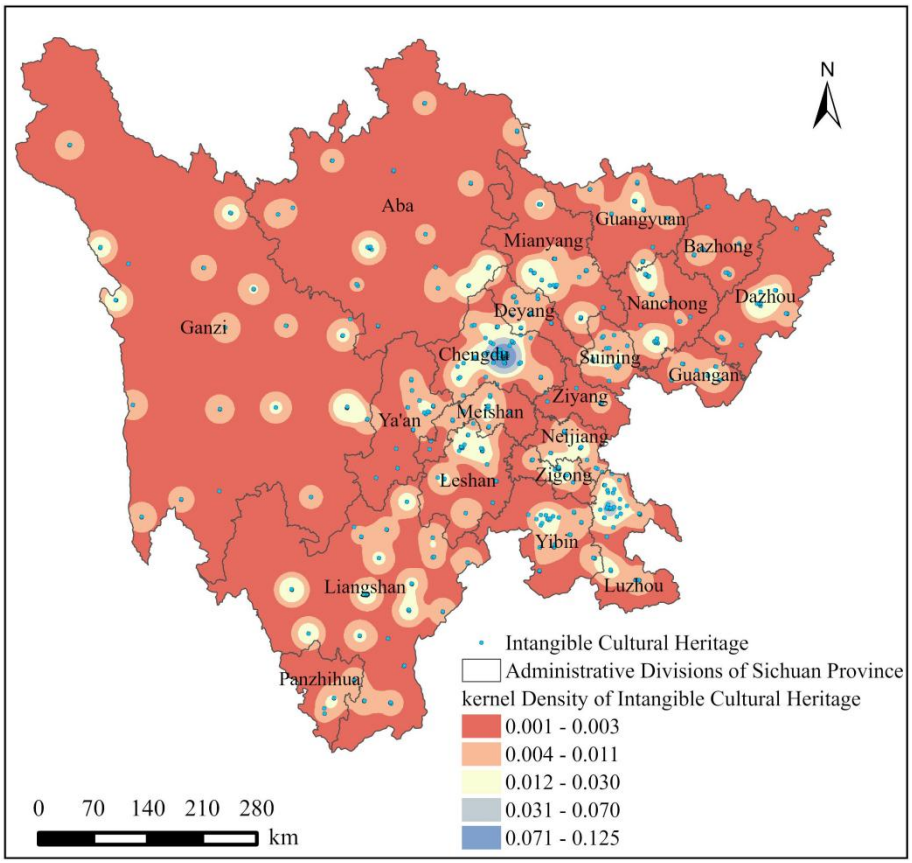


Fig. 4 Kernel density of intangible cultural heritage items

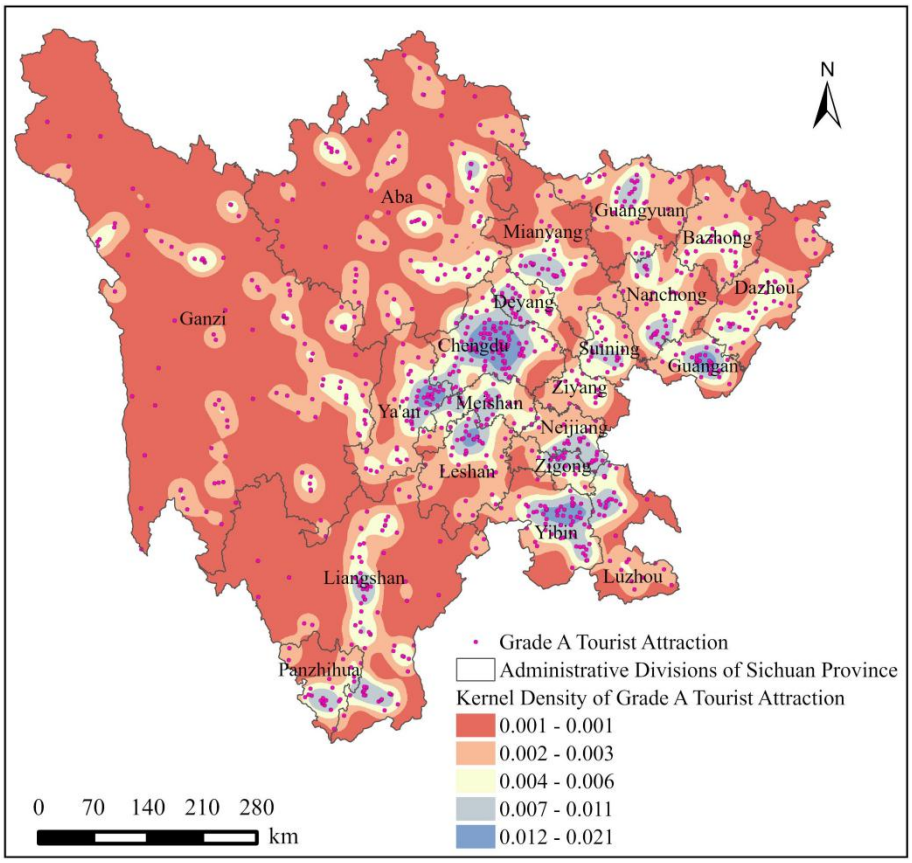


Fig. 5 Kernel density of A-class scenic areas

The gravity model gives the degree of separation at the regional scale (Figs. 6–8). The village centroid lies to the west, 178.6 km from the heritage protection unit centroid, with a size-match score of only 0.32. The ICH centroid lies about 64 km west of the heritage centroid and 114.2 km from the village centroid. The scenic area centroid nearly coincides with the heritage centroid but lies 196.5 km from the village centroid, with a size-match score of only 0.27. All of these are centroid separations. They reflect distributional deviation at the regional scale and are not equal to the travel distance from villages to resources.

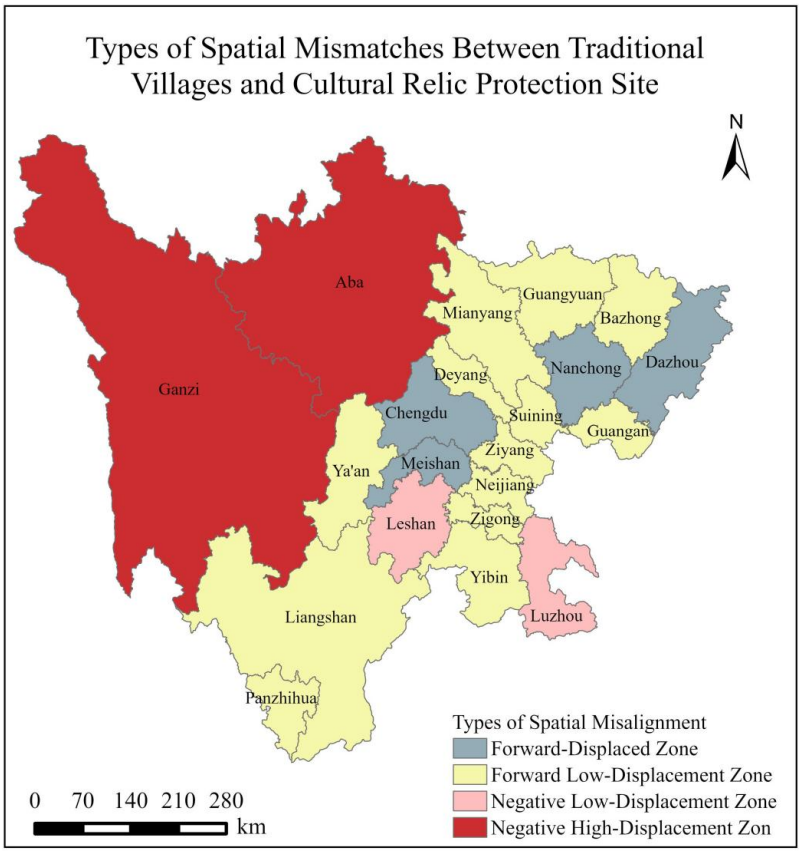


Fig. 6 Spatial mismatch between villages and heritage protection units

Note: All maps used in this study are derived from standard map No. GS(2024)0650 of the Standard Map Service System; the base maps were not modified.

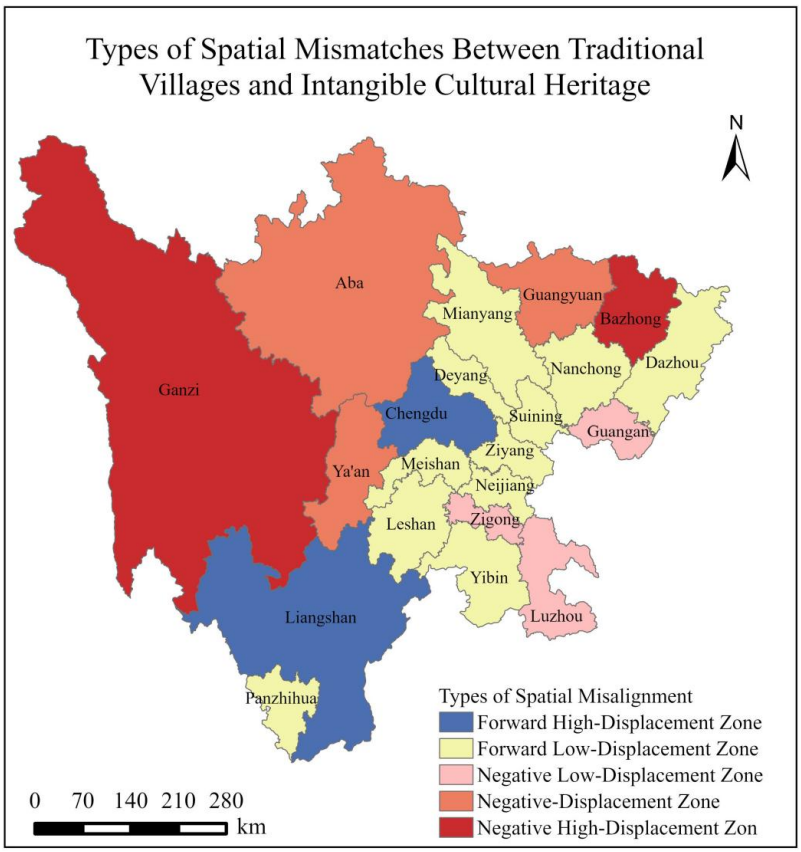


Fig. 7 Spatial mismatch between villages and intangible cultural heritage

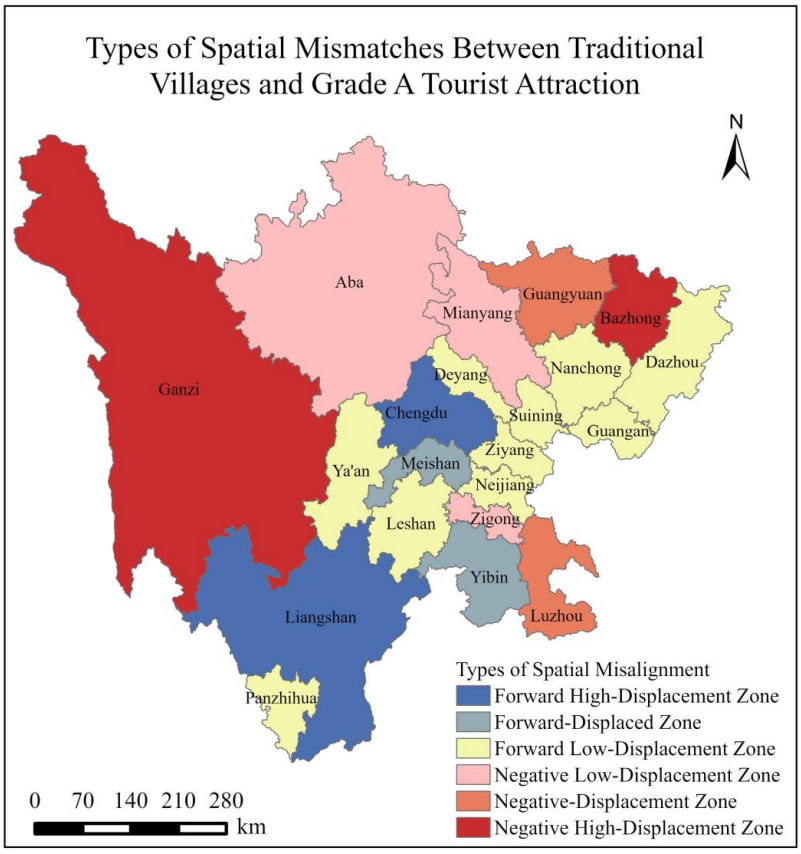


Fig. 8 Spatial mismatch between villages and A-class scenic areas

The village scale is different. The mean distances from the 1325 villages to the nearest heritage protection unit, ICH item and A-class scenic area are 14.6, 18.7 and 9.2 km. Liangshan has the highest values on all three (43.5, 40.4 and 19.3 km), followed by Ganzi (28.8, 27.9 and 12.6 km) and Aba (25.3, 23.6 and 7.2 km), while Chengdu has the lowest (8.0, 13.1 and 6.1 km). The two kinds of metric are not synchronized: the centroid deviation index of Ganzi is only 0.067, yet its mean village distance reaches 28.8 km. Centroid distance is masked by resource concentration inside a prefecture, whereas village-level distance reflects actual accessibility.

Regional size imbalance takes another form. *RAI* has a mean of 0.952 and a range of 0.772-0.993, and most prefectures approach the upper limit: whether resources are rich (Chengdu) or scarce (Ganzi), the resource share and the village share differ markedly, only in opposite directions. This is precisely why size imbalance cannot be treated directly as stress — it carries no direction.

3.2 Indicator weights

Entropy weights are uneven within the four dimensions (Table 2). Within exposure, C1 (tourism revenue share, 31.96%) and C6 (highway passenger turnover, 25.35%) together account for 57.3%, indicating that economic dependence on tourism and population mobility are the main sources of external pressure. The three cultural location weights are close to each other (C8 44.02%, C9 26.44%, C10 29.55%), and the distance to heritage protection units has the highest weight, reflecting the greater importance of the cost of obtaining protection resources.

Sensitivity weights concentrate on natural conditions. C11 (mean elevation) rises to 45.68% and C12 (slope) takes 20.39%, 66.1% together. Adaptive capacity is dominated by accumulated wealth and fiscal capacity: C20 (deposits per capita, 37.91%) and C19 (fiscal expenditure, 20.30%) rank first, while C22 (hospital beds) takes only 5.42%. In the provincial comparison, differences in adaptive capacity come mainly from household wealth and government finance, not from per capita medical resources.

The weight structure of B4 is highly sensitive to extreme values. Deposits per capita in Chengdu are about 59 times those in Ganzi, and entropy weighting concentrates 37.91% of the weight there (see Section 4.3).

Table 2 Entropy weights of the indicators

Criterion layer	Code	Information entropy e_j	Utility value d_j	Weight/%
Exposure B1	C1	0.792	0.208	31.96
	C2	0.941	0.059	9.08
	C3	0.976	0.024	3.69
	C4	0.940	0.060	9.18
	C5	0.936	0.064	9.80
	C6	0.835	0.165	25.35
	C7	0.929	0.071	10.94
Cultural location B2	C8	0.833	0.167	44.02
	C9	0.900	0.100	26.44
	C10	0.888	0.112	29.55
Sensitivity B3	C11	0.775	0.225	45.68
	C12	0.899	0.101	20.39
	C13	0.950	0.050	10.11
	C14	0.949	0.051	10.39
	C15	0.977	0.023	4.71
	C16	0.957	0.043	8.72
Adaptive capacity B4	C17	0.914	0.086	8.94
	C18	0.902	0.098	10.27
	C19	0.805	0.195	20.30
	C20	0.637	0.363	37.91
	C21	0.924	0.076	7.91
	C22	0.948	0.052	5.42
	C23	0.911	0.089	9.25

Note: Weights are normalized within each dimension separately, so B1 to B4 each sum to 100%.

3.3 Vulnerability index and regional differentiation

Synthesis by Eq. (1) and normalization give the persistence vulnerability index of the 21 cities (prefectures) (Table 3). The distribution shows a steep gradient: 1.000 in Ganzi, 0.829 in Aba and 0.428 in Liangshan, while the other 18 cities are all below 0.028, and 15 of them are below 0.015.

The three prefectures share two features: very low adaptive capacity and very high sensitivity. B4 is 0.034, 0.040 and 0.093 respectively, only 16%-44% of the provincial mean of 0.212; B3 is 0.777, 0.818 and 0.661, all among the highest in the province. The cultural location shortfall is most prominent in Panxi: B2 is 1.000 in Liangshan and 0.314 in Panzhihua, whereas in Ganzi and Aba the shortfall appears mainly in B4 and B3.

Differences within the three prefectures are clearly smaller than those between the three prefectures and the basin: Ganzi is 1.21 times Aba, Aba is 1.94 times Liangshan, while Liangshan is 15.6 times Guangyuan, the fourth-ranked prefecture. The three prefectures form a relatively coherent high-risk group whose boundary is clearer than its internal gradient.

Table 3 Persistence vulnerability index and dimension scores of traditional villages in the cities (prefectures) of Sichuan

Economic zone	City (prefecture)	B1	B2	B3	B4	VI
Chengdu Plain	Chengdu	0.3615	0.0551	0.2865	0.8173	0.0005
	Deyang	0.1755	0.0263	0.2840	0.3050	0.0000

Economic zone	City (prefecture)	B1	B2	B3	B4	VI
	Mianyang	0.2184	0.3483	0.3426	0.3182	0.0147
	Suining	0.2224	0.0881	0.2922	0.1952	0.0048
	Leshan	0.3382	0.2485	0.3420	0.2279	0.0231
	Meishan	0.2637	0.0937	0.3401	0.1338	0.0111
	Ya'an	0.2466	0.0770	0.6425	0.1241	0.0179
	Ziyang	0.2213	0.1726	0.2489	0.2132	0.0076
Southern Sichuan	Zigong	0.2448	0.0403	0.1718	0.2235	0.0006
	Luzhou	0.2434	0.3217	0.2877	0.1884	0.0219
	Neijiang	0.2382	0.0756	0.2589	0.1541	0.0049
	Yibin	0.2465	0.2764	0.2274	0.2964	0.0091
Northeastern Sichuan	Guangyuan	0.2664	0.2757	0.3559	0.1750	0.0275
	Nanchong	0.3332	0.2141	0.1880	0.3227	0.0071
	Guang'an	0.2320	0.1430	0.2718	0.1292	0.0124
	Dazhou	0.2155	0.3319	0.3666	0.1928	0.0250
	Bazhong	0.3172	0.2036	0.3064	0.1345	0.0271
Panxi	Panzhihua	0.0937	0.3143	0.6071	0.1379	0.0238
	Liangshan	0.3166	1.0000	0.6613	0.0927	0.4280
Northwestern Sichuan	Aba	0.5691	0.3777	0.8183	0.0402	0.8291
	Ganzi	0.4030	0.5726	0.7774	0.0340	1.0000

Note: $VI_{(raw)}=B_1\times B_2\times B_3/B_4$; VI is the min-max normalized value of $VI_{(raw)}$.

Aggregated by the five economic zones, vulnerability is ordered along the terrain gradient (Table 4): 0.9146 in northwestern Sichuan, 0.2259 in Panxi, 0.0198 in northeastern Sichuan, 0.0100 in the Chengdu Plain and 0.0091 in southern Sichuan. Dispersion within zones is equally sharp: Liangshan is 18.0 times Panzhihua in Panxi, and Ganzi is 1.21 times Aba in northwestern Sichuan.

The Theil decomposition gives a provincial index of $T=1.536$. The between-group component is $T_b=1.423$, or 92.6%; the within-group component is $T_w=0.113$, or 7.4%. The Gini coefficient is 0.816, the coefficient of variation is 2.363 and the mean is 0.1189. P1 is supported: provincial differences are almost entirely explained by the structural ordering among the five economic zones, and the residual within-group differences fall mainly between prefectures inside the ethnic minority areas.

Table 4 Vulnerability structure and village size of the five economic planning zones

Economic zone	Prefectures	Villages	B1	B2	B3	B4	Mean VI
Chengdu Plain	8	268	0.256	0.139	0.347	0.292	0.0100
Southern Sichuan	4	201	0.243	0.178	0.236	0.216	0.0091
Northeastern Sichuan	5	341	0.273	0.234	0.298	0.191	0.0198
Panxi	2	43	0.205	0.657	0.634	0.115	0.2259
Northwestern Sichuan ecological demonstration zone	2	472	0.486	0.475	0.798	0.037	0.9146
Province	21	1325	0.275	0.250	0.385	0.212	0.1189

Note: Theil index $T=1.536$ (between groups 1.423, 92.6%; within groups 0.113, 7.4%); Gini coefficient 0.816; coefficient of variation 2.363.

3.4 Village types and clustering quality

Scheme A clusters villages by elevation, with a silhouette coefficient of 0.705 and a DBI of 0.467, indicating good quality (Table 5). The three clusters are the low-elevation valley type (777 villages, 58.642%, 637.32 ± 299.14 m), the high-elevation mountain type (353 villages, 26.642%, 3668.87 ± 382.45 m) and the mid-elevation hill type (195 villages, 14.717%, 2345.65 ± 433.80 m). Separation between clusters is significant ($F=9770.647$, $p<0.001$), with $\eta^2=0.937$.

Scheme B clusters villages by distance to the nearest cultural resource, with a silhouette coefficient of 0.584 and a DBI of 0.534, indicating moderate quality. The three clusters are the core-radiation zone (818 villages, 61.736%, 9.50 ± 4.04 km), the adjacent-linkage zone (373 villages, 28.151%, 22.77 ± 4.53 km) and the peripheral-isolated zone (134 villages, 10.113%, 43.09 ± 8.86 km). Separation between clusters is again significant ($F=3146.638$, $p<0.001$), with $\eta^2=0.826$.

Peripheral-isolated villages account for only 10.1%, yet they concentrate in Ganzi (85), Aba (18) and Liangshan (15); the three prefectures together hold 88.1% of this type. Villages with the least access to cultural resources converge geographically towards the ethnic minority areas.

Table 5 Results of the two K-means clusterings

Scheme	Cluster	Count	Share/%	Centre ± SD	Assigned type
A (elevation, m)	1	195	14.717	2345.65 ± 433.80	Mid-elevation hill
	2	353	26.642	3668.87 ± 382.45	High-elevation mountain
	3	777	58.642	637.32 ± 299.14	Low-elevation valley
B (distance, km)	1	373	28.151	22.77 ± 4.53	Adjacent-linkage zone
	2	818	61.736	9.50 ± 4.04	Core-radiation zone
	3	134	10.113	43.09 ± 8.86	Peripheral-isolated zone

Note: Scheme A has a silhouette coefficient of 0.705, DBI 0.467, $F=9770.647$ ($p<0.001$) and $\eta^2=0.937$; Scheme B has a silhouette coefficient of 0.584, DBI 0.534, $F=3146.638$ ($p<0.001$) and $\eta^2=0.826$. The elbow curves are shown in Figs. 9–10. The analysis of variance uses the clustering input variable as the grouping basis and serves only as a descriptive indicator of clustering quality, not as proof of validity.

The silhouette coefficient of Scheme B is clearly lower than that of Scheme A, which indicates that a single distance dimension discriminates among villages less well than elevation. Accessibility to cultural resources is affected jointly by resource type, spatial layout and administrative level, and summarizing cultural marginality with a single distance threshold loses part of the information.

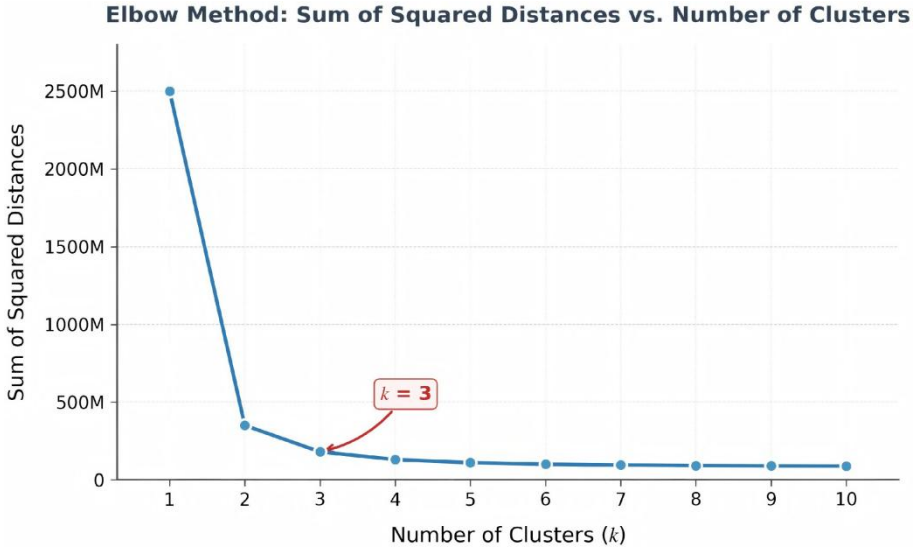


Fig. 9 Elbow curve for elevation clustering

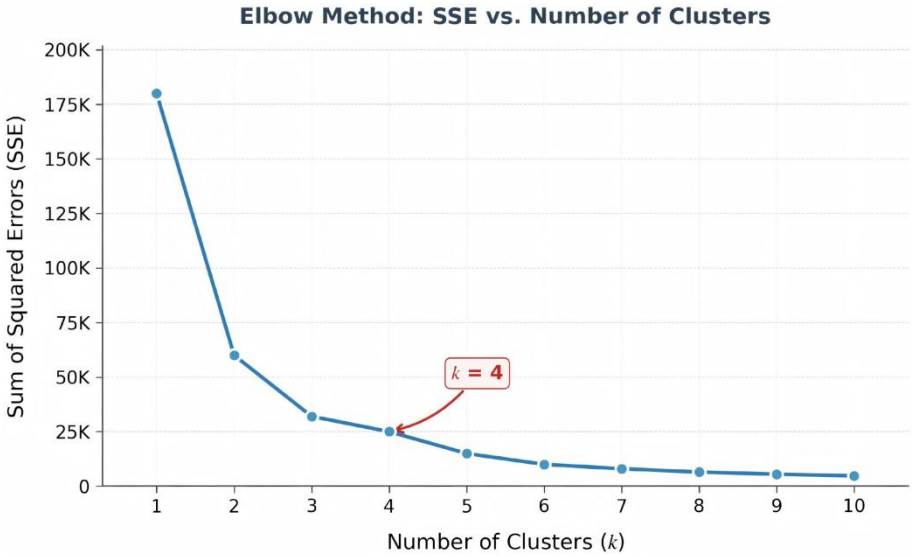


Fig. 10 Elbow curve for nearest-resource-distance clustering

3.5 Association structure and a non-circular test

The sources of the two aggregated scores come first. Scoring village types by Eq. (9) gives the scheme score of each prefecture (Table 6 and Table 7): ScoreA is highest in Ganzi (2.808), Aba (2.577) and Liangshan (2.265); ScoreB is highest in Liangshan (2.324), Ganzi (1.911) and Panzhihua (1.667). ScoreA has a step-like distribution: 10 prefectures exceed 1.000 because they contain mid- or high-elevation villages, and the other 11 are exactly 1.000. This floor makes ScoreA close to a binary variable.

Table 6 Village type counts and weighted ScoreA by city (prefecture)

City (prefecture)	VI	Mid-elevation hill	High-elevation mountain	Low-elevation valley	Total	ScoreA
Chengdu	0.0005	0	0	38	38	1.0000
Deyang	0.0000	0	0	22	22	1.0000
Mianyang	0.0147	11	0	61	72	1.1528
Suining	0.0048	0	0	14	14	1.0000
Leshan	0.0231	2	0	40	42	1.0476
Meishan	0.0111	2	0	16	18	1.1111
Ya'an	0.0179	19	0	37	56	1.3393
Ziyang	0.0076	0	0	6	6	1.0000
Zigong	0.0006	0	0	37	37	1.0000
Luzhou	0.0219	0	0	85	85	1.0000
Neijiang	0.0049	0	0	29	29	1.0000
Yibin	0.0091	0	0	50	50	1.0000
Guangyuan	0.0275	4	0	95	99	1.0404
Nanchong	0.0071	0	0	41	41	1.0000
Guang'an	0.0124	0	0	30	30	1.0000
Dazhou	0.0250	1	0	42	43	1.0233
Bazhong	0.0271	0	0	128	128	1.0000
Panzhihua	0.0238	5	0	4	9	1.5556
Liangshan	0.4280	25	9	0	34	2.2647
Aba	0.8291	70	99	1	170	2.5765

City (prefecture)	VI	Mid-elevation hill	High-elevation mountain	Low-elevation valley	Total	ScoreA
Ganzi	1.0000	56	245	1	302	2.8079
Total	**—**	**195**	**353**	**777**	**1325**	**—**

Table 7 Weighted ScoreB by city (prefecture)

City (prefecture)	Adjacent-linkage zone	Core-radiation zone	Peripheral-isolated zone	Total	ScoreB
Chengdu	2	36	0	38	1.0526
Deyang	1	21	0	22	1.0455
Mianyang	18	48	6	72	1.4167
Suining	1	13	0	14	1.0714
Leshan	14	28	0	42	1.3333
Meishan	0	18	0	18	1.0000
Ya'an	4	52	0	56	1.0714
Ziyang	2	4	0	6	1.3333
Zigong	3	34	0	37	1.0811
Luzhou	20	57	8	85	1.4235
Neijiang	0	29	0	29	1.0000
Yibin	18	31	1	50	1.4000
Guangyuan	33	66	0	99	1.3333
Nanchong	9	32	0	41	1.2195
Guang'an	8	22	0	30	1.2667
Dazhou	12	30	1	43	1.3256
Bazhong	30	98	0	128	1.2344
Panzhihua	6	3	0	9	1.6667
Liangshan	15	4	15	34	2.3235
Aba	72	80	18	170	1.6353
Ganzi	105	112	85	302	1.9106
Total	**373**	**818**	**134**	**1325**	**—**

This section carries out no test of the "independent effect" of accessibility, for two reasons. First, B2 is a constituent dimension of VI, so the two are definitionally linked. Second, ScoreB is derived by discretizing C8-C10 and correlates with B2 at 0.962, so it is almost the same source as a component of the dependent variable. If the three were placed in one regression, the coefficients would necessarily be dominated by the constructed components, and it would be impossible to distinguish "the effect is absorbed by terrain" from "the coefficient is compressed because the components are shared". Section 3.6 reports that regression as a descriptive decomposition only, not as inference. [Author note: as a non-circular alternative, please report the association between accessibility (B2 and ScoreB) and VI* = B1·B3/B4, which does not contain B2, as the primary association, and move Tables 9–10 to the Supplementary Material.]

The correlation structure itself is still informative (Table 8), and its visualization is shown in Fig. 11 [Author note: the heatmap must be regenerated from the final data; its values (e.g., B2–VI 0.754, B2–ScoreB 0.457) do not match Table 8]. The Pearson coefficient between B_2 and VI is 0.602 ($p=0.002$) and the rank correlation is 0.823; those of B_1 and B_3 are 0.713 and 0.801. More important is the relation between B_2 and ScoreA: their correlation coefficient is 0.698 ($p<0.001$), and the terrain score explains 48.8% of the accessibility score. This result is non-circular, because ScoreA and B2 do not contain each other, so the

judgment that "accessibility is distributed along the terrain gradient" is not contaminated by a definitional association.

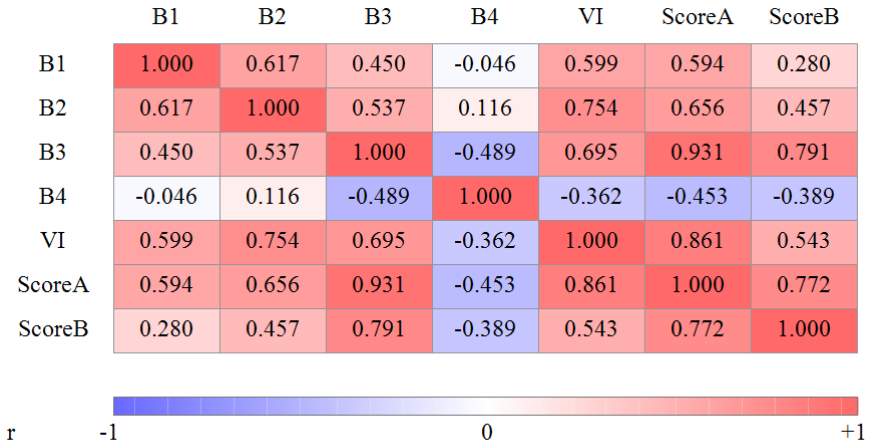


Fig. 11 Pearson correlation heatmap of the variables

The basin subsample provides a second test. In the 18 prefectures that remain after removing Ganzi, Aba and Liangshan, the correlation between distance and VI is 0.727, higher than the 0.602 of the full sample; the correlation between ScoreB and VI is 0.617. This shows that the association "the farther, the more vulnerable" does not rest on the extreme values of the three autonomous prefectures alone, and holds equally inside the basin, where tourism pressure is most concentrated and resources are densest.

The result also addresses a competing explanation: if proximity to resources made villages more vulnerable through over-tourism, the correlation inside the basin should be negative. It is significantly positive, so that explanation gains no support. This is only correlational evidence, however, and heterogeneity cannot be excluded on this basis — a few villages adjacent to scenic areas may still be damaged by commercialization, which would require village-level data to identify.

Table 8 Pearson correlation matrix of the variables (n=21)

Variable	B1	B2	B3	B4	VI	ScoreA	ScoreB
B1	1						
B2	0.290	1					
B3	0.449*	0.595	1				
B4	−0.046	−0.362	−0.478*	1			
VI	0.713*	0.602	0.801*	−0.417	1		
ScoreA	0.594	0.698*	0.909*	−0.453*	0.956*	1	
ScoreB	0.280	0.962*	0.670*	−0.389	0.668*	0.772*	1

Note: * $p<0.001$, $p<0.01$, * $p<0.05$. The correlation matrix describes contemporaneous association only and is not interpreted causally.

The correlation between B_2 and ScoreB reaches 0.962, which shows that the two are almost the same variable in this sample: the former is a weighted synthesis of continuous distances, the latter a three-tier discretization of the same distances. This is a consequence of their shared definition, and it also means that any model with VI as the dependent variable and ScoreB as an independent variable cannot separate an "accessibility effect" from "index construction".

3.6 Sensitivity of index construction

This section reports two sets of results: a descriptive regression with VI as the dependent variable, and the changes in coefficients under five index constructions. Neither is used for hypothesis testing; both serve only to show how far the conclusions depend on the specification.

In the descriptive regression, model I has $R^2=0.926$, a standardized coefficient of 1.088 for ScoreA and −0.172 for ScoreB (Table 9). Four qualifications are necessary. First, the dependent variable contains a B2 component and the independent variable ScoreB correlates with B2 at 0.962, so the coefficients are affected by a definitional association. Second, with HC1 robust standard errors the t value of ScoreB is −1.37 (p≈0.19) and that of ScoreA is 9.17; the direction is unchanged but the strength is weaker. Third, the Breusch-Pagan statistic is 2.56, so homoskedasticity is not rejected; the Jarque-Bera statistic is 6.76, so residual normality is rejected (skewness −1.24, kurtosis 4.25), which indicates that extreme values still matter in a small sample. Fourth, the sample contains only 21 observations, and the Cook's distances of Ganzi and Liangshan are 1.124 and 0.981, so the regression line is largely determined by these two points. The fitted and observed values of model I are compared in Fig. 12 [Author note: please regenerate; the plotted observed VI of Aba and Liangshan does not match Table 3], where the two influential points lie at the two ends of the fitted line.

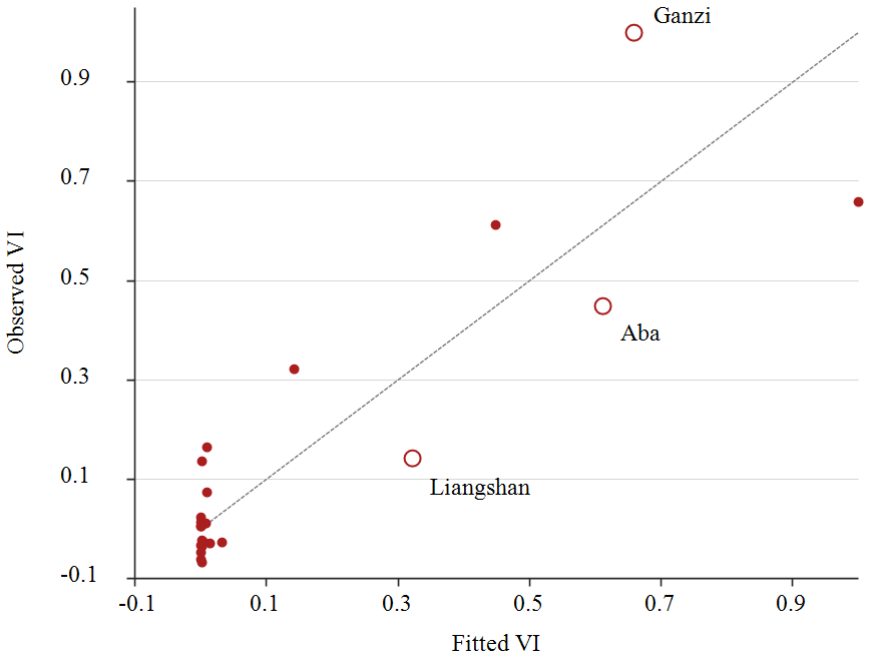


Fig. 12 Observed versus fitted persistence vulnerability

Table 9 Descriptive multiple regression of the persistence vulnerability index (no effect inference)

Model	Variable	B	SE	Robust SE	Standardized β	t (robust)	p	VIF
Model I (n=21)	Constant	−0.3908	0.0796	0.0952	—	−4.10	<0.001	—
	ScoreA	0.5514	0.0512	0.0601	1.088	9.17	<0.001	2.47
	ScoreB	−0.1471	0.0866	0.1076	−0.172	−1.37	0.188	2.47
	Diagnostics	R^2=0.926, adjusted R^2=0.917; BP=2.56; JB=6.76; max Cook's D 1.124 (Ganzi)						

Note: The dependent variable VI contains a B2 component and ScoreB correlates with B2 at 0.962, so the coefficients are affected by a definitional association and the table is descriptive only. Robust standard errors are HC1.

The comparison of the five index constructions is given in Table 10. The standardized coefficient of ScoreA ranges from 0.501 to 1.243 and is always significant; the coefficient of ScoreB changes sign between −0.455 and +0.486. It should be noted that these five are not perturbations of the data or the model but five

ways of constructing the index itself, so Table 10 reflects specification sensitivity, not robustness testing in the usual sense.

Table 10 Specification sensitivity of index construction

Specification	Description	Rank correlation with baseline	ScoreA β	p	ScoreB β	p	R^2
Baseline	Entropy weights, multiplicative VI, full sample	—	1.088	<0.001	−0.172	0.107	0.926
Additive aggregation	VIadd	0.955	0.617	<0.001	0.373	0.011	0.875
Geometric mean	VIgeo	0.936	0.501	0.002	0.486	0.002	0.862
B2 removed	VI*=B1B3/B4	0.844	1.243	<0.001	−0.455	0.002	0.879
B3 removed	VI^=B1B2/B4	0.922	1.016	<0.001	−0.075	0.485	0.920
Extreme cases removed	Ganzi and Liangshan excluded (n=19)	—	0.962	<0.001	−0.060	0.598	0.865

Note: The baseline is the multiplicatively aggregated VI; rank correlations are Spearman coefficients; all coefficients are standardized.

The value of this table lies in marking a boundary rather than in drawing a conclusion: the direction of the terrain score is stable, the direction of the accessibility score is not. This study therefore takes the terrain gradient as the main axis of zoning and accessibility as a secondary characterization within zones, and makes no assertion about its effect size or direction.

3.7 Two-dimensional zoning

ScoreA and ScoreB are used to build a zoning matrix. The thresholds are taken from the data themselves; no bootstrap interval estimation was performed and the stability of the classification under threshold shifts was not tested, so the zoning is positioned here as an exploratory scheme. ScoreA is split at 1.000 (the 11 prefectures without mid- or high-elevation villages score exactly 1.000) and ScoreB at the provincial median of 1.326. The four quadrants contain 7, 3, 3 and 8 prefectures respectively (Table 11). [Author note: please test the stability of quadrant membership by bootstrap resampling or threshold perturbation.]

Table 11 Two-dimensional strategy zoning matrix and rural revitalization measure packages

Quadrant	Criterion	Cities (prefectures)	Structural features	Rural revitalization measures
I High natural stress, high cultural vitality	ScoreA>1.0, ScoreB>1.326	Ganzi, Aba, Liangshan, Panzhihua, Mianyang, Leshan, Guangyuan (7)	Strong elevation stress; villages are numerous and the cultural network is dense; B4 is generally low (0.034-0.318)	Culture + organization: apply for cultural ecology protection areas with village clusters rather than single villages as the unit; embed ICH workshops and transmission studios in existing village networks instead of building new parks. Talent: create paired posts linking local bearers of transmission with resident village planners. Ecology: incorporate geological hazard and snow disaster avoidance into village spatial planning
II High natural stress, low cultural vitality	ScoreA>1.0, ScoreB≤1.326	Meishan, Ya'an, Dazhou (3)	Natural stress is evident; cultural accessibility is acceptable but inter-village coordination is weak (mean B4 0.143)	Industry + culture: develop study-tour and wellness businesses based on nearby cities. Organization: form cross-village protection alliances around central villages to avoid scattered resources

Quadrant	Criterion	Cities (prefectures)	Structural features	Rural revitalization measures
III Low natural stress, high cultural vitality	ScoreA≤1.0, ScoreB>1.326	Luzhou, Yibin, Ziyang (3)	Terrain constraints are weak and mean B4 is 0.220, but many villages fall in the adjacent-linkage zone (23.5%-36.0%)	Industry + culture: link villages and scenic areas with linear heritage corridors; develop combined products of ICH plus homestay and ICH plus study tours. Talent: train a workforce for rural operation and digital communication
IV Low natural stress, low cultural vitality	ScoreA≤1.0, ScoreB≤1.326	Chengdu, Deyang, Suining, Zigong, Neijiang, Nanchong, Guang'an, Bazhong (8)	Both natural and cultural stress are low, and B4 varies widely (0.129-0.817)	Ecology + culture: establish preventive monitoring and visitor capacity control. Organization: transfer cluster governance experience linking villages, ICH and scenic areas, and take on a demonstration role

The four quadrants face different constraints. Quadrant I is constrained by adaptive capacity, with B4 in the three prefectures falling to 0.034-0.093; quadrant II combines low adaptive capacity with moderate cultural accessibility, with a mean B4 of 0.143 and a mean B2 of 0.219; the weakness of quadrant III is conversion efficiency rather than allocative imbalance; quadrant IV has the highest adaptive capacity, with a mean B4 of 0.292.

Bazhong deserves separate mention. It lies in quadrant IV with a VI of only 0.027, yet its B_2 reaches 0.204, the highest among the five prefectures of northeastern Sichuan. This shows that in prefectures with lower terrain stress, cultural accessibility can still be a relative weakness; ranking by a single index tends to overlook such "low VI, high B2" combinations, and this is where two-dimensional zoning adds value.

4 Discussion

4.1 Dominant role of the terrain base

The most robust finding is that the physical-geographical base dominates the persistence vulnerability of traditional villages. Five pieces of evidence point in the same direction. C11 (mean elevation) has a weight of 45.68% within the sensitivity dimension, the highest of all indicators. The Pearson coefficient between ScoreA and VI is 0.956 and the rank correlation is 0.717, the highest among all explanatory variables. In the full-sample regression the standardized coefficient of ScoreA is 1.088 ($p<0.001$) and that of ScoreB is −0.172 ($p=0.107$). Across the five specifications the coefficient of ScoreA ranges from 0.501 to 1.243 and never loses significance. ScoreA alone explains 91.4% of the variance.

This is not a simple determinism of "the higher, the more vulnerable". ScoreA measures the elevation composition of villages within a prefecture, not mean elevation: 11 prefectures take the floor value of 1.000 because they contain no mid- or high-elevation villages. Its explanatory power comes from another fact — in some regions the settlement system as a whole is confined to high-elevation habitats, with no low-elevation space to fall back on. Of the 302 villages in Ganzi, 245 belong to the high-elevation mountain type, or 81.1%; of the 170 villages in Aba, 99 do, or 58.2%.

Adaptive capacity amplifies risk rather than determining rank. B_4 is only 0.034, 0.040 and 0.093 in Ganzi, Aba and Liangshan, less than half the provincial mean of 0.212. Because B_4 sits in the denominator of Eq. (1), a low value multiplies the given exposure and sensitivity several times over. This agrees with the established understanding of adaptive capacity as a conditioning factor of vulnerability [46,47,50]. An index decomposition gives a quantitative illustration: in the log-variance of *VI*, 31.2% comes from $1/B_4$ and 68.8% from the numerator. It must be stressed that this is an identity decomposition. It answers the question "how much of the extreme value comes from the denominator" and constitutes neither a test of a moderating effect nor a causal claim.

The policy implication is then clear. The terrain base cannot be changed, and adaptive capacity is the most operable lever. Scenario calculations show that raising the B_4 of Ganzi from 0.034 to the provincial mean of

0.212 would lower $VI_{(raw)}$ by about 84%, whereas lowering its B_2 from 0.573 to the provincial mean of 0.250 would lower $VI_{(raw)}$ by about 56%. The former is larger in magnitude, but the feasibility of the two cannot be judged from formula elasticity alone; fiscal and demographic realities must be taken into account (see Section 4.4).

4.2 Spatial coupling between accessibility and the terrain gradient

The second finding is that accessibility and vulnerability covary in the same direction, and that accessibility itself is distributed along the terrain gradient. The Pearson coefficient between B_2 and VI is 0.602 and the rank correlation is 0.823; more important, the correlation between B_2 and ScoreA reaches 0.698, and the terrain score explains 48.8% of the accessibility score. It should be noted that B_2 is a constituent component of VI and the two cannot be separated in this design — this is not a demonstration that "the effect is absorbed by terrain", but a statement that this design cannot answer that question.

The anthropological concept of the cultural niche offers an interpretation. A niche is not determined by physical distance alone but by the position of a village within the triple network of resources, markets and institutions [6,39,51]. In the Tibetan-Qiang-Yi corridor, the Miao frontier corridor and the Wuling corridor, cultural resources tend to be distributed along corridor axes and administrative centres, while villages are scattered in the tributary valleys on either side [7,8]. Accessibility differences are therefore not accidental misallocation but a spatial order shaped over the long term by human-land relations and administrative hierarchy in corridor regions.

A methodological choice in indicator construction deserves explanation. A centroid-based mismatch index, if its absolute value is used, codes “many villages, few resources” (Ganzi) and “many resources, few villages” (Chengdu) as pressure in the same direction. This study therefore uses village-level nearest distance, which has only one meaning for a village — the farther the worse — so the direction of the construct is unambiguous. Size imbalance is listed separately as RAI and is direction-neutral; its correlation with VI is −0.735, showing that prefectures whose resource share far exceeds their village share (Chengdu, Deyang) are in fact less vulnerable. This agrees with the judgment that “resource enrichment is not in itself stress” and explains why the two cannot be merged into one dimension.

The combination in Ganzi and Aba deserves attention: the two prefectures hold 23.7% of the provincial ICH items but only 9.2% of the heritage protection units. Much of their cultural resource exists in everyday practice rather than entering in the form of external investment [1,38,49]. As long as bearers of transmission, song masters and ritual specialists still live in the villages, the chain of cultural reproduction is not broken. This is the practical meaning of cultural subjectivity: persistence ultimately depends on whether villagers remain the holders and practitioners of culture [1,3].

A testable proposition follows: the persistence of villages in ethnic minority areas depends less on the distance to resources than on whether the chain of cultural reproduction is intact and whether community self-organization survives. The former rests on bearers, venues and practice; the latter on the density of social networks and the mode of institutional embedding [38,39]. The quantitative results of this study support the necessity of this proposition but cannot yet verify its sufficiency.

The focus of protection policy should therefore shift from "delivering resources" to "cultivating subjects": training bearers of transmission locally, keeping ritual venues in daily operation, and giving villagers a voice in the distribution of tourism revenue [4,38]. Sustained population outflow erodes the same network and pushes villages into a cycle of hollowing out, network thinning and broken transmission [52].

4.3 Five methodological reflections

First, the aggregation method can change the sign of a conclusion. The correlation between B_2 and VI is significant, but once VI is changed from multiplicative to additive or geometric aggregation, the standardized coefficient of ScoreB moves from −0.172 to +0.373 and +0.486, both significant; with B_2 removed it becomes −0.455. The sign of the coefficient reverses with the specification, which shows that 21 observations are insufficient to stably separate two highly covarying explanatory variables ($r=0.772$). This

study therefore makes no directional assertion about the independent effect of cultural location, and reports only that "no robust independent association was found".

Second, entropy weights are extremely sensitive to dispersion. The weight structure of B4 is a typical example: deposits per capita in Chengdu are about 59 times those in Ganzi, so C20 receives 37.91% of the within-dimension weight, while C22 (hospital beds), which relates directly to the level of public services, receives only 5.42%. Entropy weighting measures information rather than mechanistic importance; under alternative specifications (equal weights, geometric mean) the ranking of VI is preserved (rank correlation 0.936-0.955), so the conclusions do not depend on the weights, but the bias may be amplified in samples with greater dispersion [48].

Third, multiplicative aggregation compresses resolution at the low end. Of the 21 prefectures, 18 have a VI below 0.03 and two are close to 0. This is not a computational error but an inevitable result of treating vulnerability as a conditional state: multiplicative aggregation identifies situations in which several conditions hold at once and is not suited to continuous ranking across the whole sample. Table 3 therefore retains the original VI values without further processing, and Table 10 reports the additive and geometric-mean versions side by side, so that readers can choose according to purpose.

Fourth, the aggregation path requires explanation. Elevation, slope and nearest distance for the 1325 villages are village-level, whereas socioeconomic indicators are available only at prefecture level, and what finally enters the regression is 21 means. This compression from 1325 to 21 flattens differences within prefectures and leaves a scale gap between the statement "village vulnerability" and the evidence "prefecture means". The conclusions are therefore confined to provincial zoning and are not extrapolated to the village scale; village-level mechanisms would require village panels to identify.

Fifth, the indicator metric must match the scale of intervention. Cultural resource density, village density and the share of primary school students are all village-scale concepts, but data availability restricts them to the prefecture scale. At that scale, dense cultural resources often also mean a high degree of urbanization and high exposure, and these indicators no longer appear as protective factors. Accessibility has been changed to village-level distance to alleviate the problem, but socioeconomic indicators remain at prefecture level.

4.4 Differentiated measures and governance linkages

Differentiated measures presuppose that policy priorities correspond to the constraints of each zone. The four quadrants face different constraints, so the measures should differ as well.

Quadrant I is represented by Ganzi, Aba and Liangshan, where the constraint is adaptive capacity. Culture and organization should come first: apply for cultural ecology protection areas with village clusters rather than single villages as the unit, and embed ICH workshops and transmission studios in existing mutual-aid networks instead of building new parks [4,49]; create paired posts for local bearers of transmission and resident village planners in order to retain talent [6]; incorporate geological hazard and snow disaster avoidance into village spatial planning [41].

Quadrant II is represented by Meishan, Ya'an and Dazhou, where natural stress is evident and cultural networks are sparse, so "protection at single points, decline of the whole" easily occurs. Industry and organization should come first: develop study tours and wellness businesses based on nearby cities, and form cross-village protection alliances around central villages to avoid fragmentation of protection resources [19,53].

Quadrant III is represented by Luzhou, Yibin and Ziyang, where terrain constraints are weak and resources are near but not reachable. Industry and culture should come first: link villages and scenic areas with linear heritage corridors, develop combined products of ICH with homestays and study tours [21], and provide supporting talent for rural operation and digital communication.

Quadrant IV is represented by Chengdu, Deyang and Suining, where vulnerability is generally low and the main risk is loss of authenticity through urbanization and over-tourism. Ecology and culture should come

first: establish preventive monitoring and capacity control, and transfer cluster governance experience linking villages, ICH and scenic areas [6,20].

Three statements need to be reconciled here. First, terrain determines the ordering of vulnerability; this is structural and cannot be changed in the short term. Second, cultural location is a spatial projection of the terrain gradient, and at the regional scale moving resources directly can hardly change the ordering. Third, accessibility is itself an intervenable mediator: road networks, delivery logistics, digital infrastructure and mobile services can reduce the actual travel cost from villages to resources, and thus compress B_2 without changing terrain. Policy therefore targets "lowering travel cost" rather than "moving resources", which does not contradict conclusion (4).

Zoning is not fixed. As transport and digital infrastructure improve, provincial differences in B_4 will narrow and the quadrant boundaries will shift. Differentiated measures require a monitoring loop that tracks VI and its dimensions year by year, forming a closed cycle of assessment, zoning, intervention and reassessment. For ethnic minority areas this loop has an additional meaning: maintaining the continuity of cultural transmission in the course of development is the practical basis for forging a strong sense of community for the Chinese nation.

4.5 Implications for intelligent spatial decision support

Beyond its substantive findings, the study shows how a transparent, data-driven pipeline can serve as a decision-support layer for heritage-oriented rural governance. Because every step — distance computation, standardisation, entropy weighting, aggregation, clustering and zoning — is specified explicitly, the assessment can be re-run whenever village registers, resource lists or yearbook statistics are updated, which matches the cycle of assessment, zoning, intervention and reassessment proposed in Section 4.4. In SDSS terms, the vulnerability index and its four dimension scores form the analytical model, the zoning matrix provides the classification and scenario module, and the maps and measure packages constitute the user-facing output [24].

Three extensions would strengthen this decision-support function. First, Euclidean distances could be replaced by road-network travel times or DEM-based cost distances, which can be computed routinely from open map services and would make B2 responsive to transport investment. Second, prefecture-level socioeconomic inputs could be complemented by village-scale proxies derived from remote sensing (built-up area, night-time light and land-use change) and from mobile-signalling data, narrowing the scale gap discussed in Section 4.3. Third, spatially explicit learning methods, such as spatially constrained clustering of multivariate village-level features, random-forest models interpreted with Shapley additive explanations, or graph-based models of village–resource networks, could identify vulnerable village groups and their drivers directly from village-level data [22,23,54,55], and digital-twin and smart-village platforms could link these outputs to field monitoring of bearers of transmission and ritual spaces [25,26]. These extensions lie beyond the present data, but the pipeline proposed here provides the scaffolding into which they can be inserted.

The limits identified above apply with greater force once results are automated. A routinely updated index can convey a precision that the data do not support: with 21 prefectures, rank orders outside the three autonomous prefectures are compressed and sensitive to aggregation choices, and the zoning thresholds have not yet been tested for stability. An operational system should therefore report uncertainty intervals alongside index values [34], keep human judgment — including that of village communities — in the loop, and treat the zoning as an input to policy dialogue rather than as an automated allocation rule.

In system terms, the pipeline can be deployed as a lightweight web-based SDSS with four components: (i) a data-ingestion module that harvests updated village and heritage registers, yearbook statistics and weather records through scheduled extract–transform–load routines; (ii) an indicator engine that recomputes distances, standardised indicators and entropy weights; (iii) a model module that recalculates the index, village typologies and zoning together with sensitivity diagnostics; and (iv) a visual-analytics interface, comprising map dashboards and ranked lists, through which planners can query prefectures, compare scenarios such as

the changes in B4 and B2 discussed in Section 4.1, and export measure packages [24]. Because the components exchange well-defined tables, individual modules can be replaced without rebuilding the whole system, for example by swapping Euclidean distances for network travel times or entropy weights for expert-elicited weights.

Deploying such a system in ethnic minority areas also raises questions of responsible computing. Fine-grained village data, especially on ritual spaces and bearers of transmission, are culturally sensitive and should be governed with community consent and appropriate access control. Data-driven indices can also encode structural bias: remote villages with sparse statistics may be systematically scored as more vulnerable or, conversely, overlooked, so data gaps should be reported rather than imputed silently. Finally, the outputs should remain explainable to non-specialists; the transparent, formula-based design adopted here, in which every score can be traced back to its indicators, is a deliberate choice of auditability over predictive complexity.

4.6 Limitations

Five limitations need to be stated. First, identification: B2 is a constituent dimension of VI and ScoreB correlates with B2 at 0.962, so any regression with VI as the dependent variable contains a definitional component; this study therefore makes no directional assertion about the independent effect of accessibility. Second, scale: socioeconomic indicators are available only at prefecture level, and information on 1325 villages is finally compressed into 21 means, so the conclusions are limited to provincial zoning and would fall into the ecological fallacy if extrapolated to the village scale [56]. Third, the distance metric: C8-C10 are Euclidean straight-line distances without road network or terrain resistance, which underestimates the actual travel cost in mountain areas; the centroid mismatch index has the same limitation and is therefore used only for regional-scale description. Fourth, the sample: n=21, the Cook's distances of Ganzi and Liangshan are 1.124 and 0.981, the residuals are tested and do not follow a normal distribution, and R2 should not be claimed as explanatory power; the correlation between ScoreA and ScoreB is 0.772, so the two effects cannot be fully separated. Fifth, mechanism: the chain of cultural reproduction and community self-organization are theoretical inferences that still lack field evidence.

Three directions can be pursued: paired village-level panels that bring distance, population and the number of bearers of transmission down to the village scale; road network and travel-cost data to replace Euclidean distance; and mixed-method fieldwork to verify qualitative judgments such as "bearers remain in place and ritual spaces keep operating".

5 Conclusions

This study takes 1325 traditional villages at or above the provincial level in the 21 cities (prefectures) of Sichuan Province as its sample, 506 of them in the three prefectures, and within the VSD three-dimensional framework splits exposure into a disturbance block and a structural block, measures persistence vulnerability and proposes a zoning scheme. Seven conclusions follow.

(1) Accessibility to cultural resources has a clear magnitude and direction. The mean distances from the 1325 villages to the nearest heritage protection unit, ICH item and A-class scenic area are 14.6, 18.7 and 9.2 km respectively; they are highest in Liangshan (43.5, 40.4 and 19.3 km) and lowest in Chengdu (8.0, 13.1 and 6.1 km). Centroid separations (178.6, 114.2 and 196.5 km) reflect only distributional deviation at the regional scale and cannot be read as travel distances from villages to resources.

(2) Vulnerability shows a steep gradient, and provincial differences are dominated by between-group differences. Ganzi 1.000, Aba 0.829 and Liangshan 0.428, while the other 18 cities are all below 0.028. The Theil index is 1.536, with between-group differences accounting for 92.6%. Differences within the three prefectures (Ganzi is 1.21 times Aba) are clearly smaller than differences between the three prefectures and the basin (Liangshan is 15.6 times Guangyuan), and the three prefectures form a high-risk group with clear boundaries.

(3) Accessibility and vulnerability covary in the same direction, but its independent effect is not identifiable in this design. Their correlation coefficient is 0.602 and the rank correlation is 0.823; however, B2 is a constituent component of VI and ScoreB correlates with B2 at 0.962, so the regression coefficients are governed by a definitional association. This study accordingly claims neither that "accessibility is absorbed by terrain" nor that "the resource distance argument does not hold" — both assertions lie beyond what the data support.

(4) Accessibility and the terrain gradient are highly coupled in space, and this is the part that can be confirmed. The correlation between B_2 and ScoreA is 0.698, and the terrain score explains 48.8% of the accessibility score; the two variables do not contain each other and are not affected by a definitional association. In the basin subsample of 18 cities that excludes the three prefectures, the correlation between distance and VI is still 0.727, which shows that the association is not carried by extreme cases.

(5) Adaptive capacity determines the magnitude of extreme cases, but the relevant conclusions depend on index construction. B4 in the three prefectures is only 0.034-0.093, which amplifies vulnerability by about 2 to 6 times under the multiplicative structure; under additive or geometric aggregation, the coefficient of the accessibility score changes sign between −0.455 and +0.486. Index construction affects the conclusions more than the data do, and this must be stated whenever the results are used.

(6) The four-zone classification is an exploratory scheme. The 21 prefectures fall into four strategy zones by ScoreA and ScoreB (7:3:3:8), corresponding to four measure packages: culture-organization, industry-organization, industry-culture and ecology-culture. The thresholds are taken from the data and no interval estimation was performed, so the zoning should serve as a starting point for policy dialogue rather than a conclusion; the VI of Bazhong is only 0.027 while its B_2 reaches 0.204, which suggests that monitoring should track both indicators at the same time.

(7) Methodologically, the assessment is organised as a reproducible pipeline that runs from open registers and statistics to a zoning matrix. It can be updated annually and extended with network travel times, remote-sensing proxies and spatially explicit machine-learning methods, so that it can serve as a decision-support layer for heritage-oriented rural governance.

Declarations

Funding: Humanities and Social Sciences Planning Project of the Ministry of Education, "Coexistence of Multi-ethnic ICH Symbols in the Wuling Ethnic Corridor and the Construction of the Chinese National Community" (25YJA850007); Chongqing Municipal Education Science Planning Project, "Vocational Education Promoting Local ICH Craft Transmission and Innovation from the Perspective of Rural Revitalization" (25K25YG3130163); Chongqing Educational Science Planning Project "Research on Pathways for Vocational Colleges to Empower Rural Revitalization through the Integration of Education, Culture, and Tourism" (K22YG3203353);Humanities and Social Sciences Planning Project of the Chongqing Municipal Education Commission, "Multiple Paths of Digital Topological Aesthetics in ICH Transmission under the Rural Revitalization Strategy" (25SKGH329).

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

Ethics Statement and Informed Consent Statement: Not applicable. This study used published registers, statistical yearbooks and meteorological records and did not involve human participants or animals.

Data Availability Statement: [Author note: please state where the village-level coordinates, distances, indicator data and analysis scripts can be accessed, e.g., a public repository with a DOI or the Supplementary Material.]

Author Contributions: Z.M. designed the study (Methodology) and coordinated manuscript preparation (Project administration). Y.T. conducted field research (Investigation) and recorded video footage (Visualization). X.J. compiled the research data (Data curation). F.G. prepared the maps for the study (Visualization). Y.L. drafted portions of the manuscript (Writing – original draft).

References

[1] Wang M, Chen XG. An integrated multi-perspective approach to the anthropological study of traditional villages. *J Yunnan Minzu Univ (Philos Soc Sci Ed)*. 2023;40(5):79–88. Chinese.

[2] Li Y, Shu Y. The value of living protection and development for traditional villages in Southwest China. *J Orig Ecol Natl Cult*. 2025;17(1):135–142,156. Chinese.

[3] Guo W. Sacred space in traditional villages: form, contemporary value and a reappraisal of research paradigms. *Hum Geogr*. 2020;35(6):1–8. Chinese.

[4] Wu XZ, Zhang YX. A study of the neo-endogenous development model for traditional villages in China. *Guizhou Ethn Stud*. 2024;45(4):95–101. Chinese.

[5] Sun JX, Zhou M, Wang YN. A theoretical construction of the generation of traditional village resource value: based on the framework of "technology–institution–subject". *J Nat Resour*. 2024;39(8):1924–1939. https://doi.org/10.31497/zrzyxb.20240811. Chinese.

[6] Cheng G, Wang C, Dou H, Wang Z, Lei X. Toward a livable, workable, beautiful, and harmonious countryside: a neo-endogenous development framework for revitalizing traditional villages. *Chin Geogr Sci*. 2026;36(7):1246–1264. https://doi.org/10.1007/s11769-026-1628-4

[7] Shi Y, He M, Liu L, Li J. Holistic protection and sustainable planning: exploring the spatial association characteristics of traditional villages in Tibetan-Yi Ethnic Corridor, China. *npj Herit Sci*. 2026;14(1):325. https://doi.org/10.1038/s40494-026-02582-9

[8] Zhao W, Xiao D, Li J, Xu Z, Tao J. Research on traditional village spatial differentiation from the perspective of cultural routes: a case study of 338 villages in the Miao Frontier Corridor. *Sustainability*. 2024;16(13):5298. https://doi.org/10.3390/su16135298

[9] Li Y, Fan W, Yuan X, Li J. Spatial distribution characteristics and influencing factors of traditional villages based on geodetector: Jiarong Tibetan in Western Sichuan, China. *Sci Rep*. 2024;14(1):11700. https://doi.org/10.1038/s41598-024-62666-z

[10] He XQ, Li F, Gao JB, et al. Spatial distribution and influencing factors of traditional villages in China based on geographic grids. *Sci Geogr Sin*. 2024;44(6):995–1004. Chinese.

[11] Yin LC, Liu PL, Li BH, et al. A landscape morphology gene atlas of traditional settlements: a case study of the Xiangjiang River Basin. *Sci Geogr Sin*. 2023;43(6):1053–1065. Chinese.

[12] Hu Z, Strobl J, Min Q, Tan M, Chen F. Visualizing the cultural landscape gene of traditional settlements in China: a semiotic perspective. *Herit Sci*. 2021;9(1):115. https://doi.org/10.1186/s40494-021-00589-y

[13] Zou J, Liu Y, Tan FH, et al. Landscape vulnerability of traditional villages and its quantitative evaluation: a case study of Xintian County, Hunan Province. *Sci Geogr Sin*. 2018;38(8):1292–1300. Chinese.

[14] Zou J, Liu Y, Liu PL. A comparative study of vulnerability across different types of traditional villages. *Hum Geogr*. 2020;35(4):56–63. Chinese.

[15] He YB, Ma Y, Guo YH, et al. A comparative study of objective and subjective assessments of historical environment vulnerability in traditional villages: a case study of Nanyang City, Henan Province. *Econ Geogr*. 2023;43(9):208–218. Chinese.

[16] Wang X, Cui K. Resilience assessment, type identification and spatial zoning of traditional villages from a tripartite attribute perspective: a case study of Jincheng City, Shanxi Province, China. *Land*. 2026;15(7):1229. https://doi.org/10.3390/land15071229

[17] Li Y, Zhai B, Wang P, Villa D, Ventura E. Spatial resilience differentiation and governance strategies of traditional villages in the Qinba Mountains, China. *Land*. 2025;14(9):1852. https://doi.org/10.3390/land14091852

[18] Ma R, Liu L, Xu J, Yang M, Xu G, Yang H, et al. Vulnerability-based analysis of traditional village distribution in Northwest Yunnan as an ecologically fragile area. *Sci Rep*. 2025;15(1):43217. https://doi.org/10.1038/s41598-025-27342-w

[19] Liu S, Ge J, Bai M, Yao M, He L, Chen M. Toward classification-based sustainable revitalization: assessing the vitality of traditional villages. *Land Use Policy*. 2022;116:106060. https://doi.org/10.1016/j.landusepol.2022.106060

[20] Gao J, Yang J, Chen C, Chen W. From 'forsaken site' to 'model village': unraveling the multi-scalar process of rural revitalization in China. *Habitat Int*. 2023;133:102766. https://doi.org/10.1016/j.habitatint.2023.102766

[21] Zhou Z, Zheng X. A cultural route perspective on rural revitalization of traditional villages: a case study from Chishui, China. *Sustainability*. 2022;14(4):2468. https://doi.org/10.3390/su14042468

[22] Janowicz K, Gao S, McKenzie G, Hu Y, Bhaduri B. GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *Int J Geogr Inf Sci*. 2020;34(4):625–636. https://doi.org/10.1080/13658816.2019.1684500

[23] Liu P, Biljecki F. A review of spatially-explicit GeoAI applications in urban geography. *Int J Appl Earth Obs Geoinf*. 2022;112:102936. https://doi.org/10.1016/j.jag.2022.102936

[24] Keenan PB, Jankowski P. Spatial decision support systems: three decades on. *Decis Support Syst*. 2019;116:64–76. https://doi.org/10.1016/j.dss.2018.10.010

[25] Zavratnik V, Kos A, Stojmenova Duh E. Smart villages: comprehensive review of initiatives and practices. *Sustainability*. 2018;10(7):2559. https://doi.org/10.3390/su10072559

[26] Jouan P, Hallot P. Digital twin: research framework to support preventive conservation policies. *ISPRS Int J Geo-Inf*. 2020;9(4):228. https://doi.org/10.3390/ijgi9040228

[27] Li Z, Li D, Lai W, Liu Z, Cai Q, Liu C, et al. A deep learning-based method for identifying traditional villages across cultural types. *npj Herit Sci*. 2025;13(1):671. https://doi.org/10.1038/s40494-025-02187-8

[28] Tao H, Wen YZ, Wang L, Chen L. Spatiotemporal change and place-making of traditional villages from the perspective of digital media. *Prog Geogr*. 2025;44(9):1911–1924. https://doi.org/10.18306/dlkxjz.2025.09.011. Chinese.

[29] Li R, Feng S, Wang J, Peng W, Tan C. Optimizing environmental comfort and landscape visibility in traditional villages via digital platforms: a case study of Dazhai Village, Chengbu County, Hunan. *Sustainability*. 2025;17(24):11147. https://doi.org/10.3390/su172411147

[30] Shannon CE. A mathematical theory of communication. *Bell Syst Tech J*. 1948;27(3):379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[31] Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognit Lett*. 2010;31(8):651–666. https://doi.org/10.1016/j.patrec.2009.09.011

[32] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math*. 1987;20:53–65. https://doi.org/10.1016/0377-0427(87)90125-7

[33] Davies DL, Bouldin DW. A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell*. 1979;PAMI-1(2):224–227. https://doi.org/10.1109/TPAMI.1979.4766909

[34] Saisana M, Saltelli A, Tarantola S. Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *J R Stat Soc Ser A Stat Soc*. 2005;168(2):307–323. https://doi.org/10.1111/j.1467-985X.2005.00350.x

[35] Gobillon L, Selod H, Zenou Y. The mechanisms of spatial mismatch. *Urban Stud*. 2007;44(12):2401–2427. https://doi.org/10.1080/00420980701540937

[36] Kain JF. Housing segregation, Negro employment, and metropolitan decentralization. *Q J Econ*. 1968;82(2):175–197. https://doi.org/10.2307/1885893

[37] Wang W, Zhang X, Han Y, Fu C. Spatial differentiation and mechanisms of spatial mismatch between traditional villages and intangible cultural heritage: collaborative conservation zoning and strategies in Anhui Province, China. *Land*. 2026;15(7):1232. https://doi.org/10.3390/land15071232

[38] Luo CJ. From lineage and territorial identity to a community of shared destiny: the construction of inter-embedded ethnic relations in traditional villages. *Guangxi Ethn Stud*. 2020;(1):32–40. Chinese.

[39] Zhang HC, Shu BY. Institutional embeddedness: the evolutionary logic of tourism development models in ethnic traditional villages. *J Yunnan Minzu Univ (Philos Soc Sci Ed)*. 2019;36(3):88–94. Chinese.

[40] Fan D, Maliki NZB, He C, Bi Y, Yu S. Cultural gene characterization and mapping of traditional Tibetan village landscapes in Western Sichuan, China. *npj Herit Sci*. 2025;13(1):317. https://doi.org/10.1038/s40494-025-01877-7

[41] Liu Q, Shu P, Zhou F, Tang M. Cultural-ecological integrated governance: quantifying approaches to reducing landslide risk in traditional Tibetan villages of Western Sichuan, China. *J Mt Sci*. 2026;23(6):2794–2812. https://doi.org/10.1007/s11629-025-9832-8

[42] Shi B, Liu H, Huang L, Zhang Y, Xiang Z. Increasing vulnerability of village heritage: evidence from 123 villages in Aba Prefecture, Sichuan, China. *Land*. 2023;12(11):2048. https://doi.org/10.3390/land12112048

[43] Polsky C, Neff R, Yarnal B. Building comparable global change vulnerability assessments: the vulnerability scoping diagram. *Glob Environ Change*. 2007;17(3–4):472–485. https://doi.org/10.1016/j.gloenvcha.2007.01.005

[44] Füssel HM. Vulnerability: a generally applicable conceptual framework for climate change research. *Glob Environ Change*. 2007;17(2):155–167. https://doi.org/10.1016/j.gloenvcha.2006.05.002

[45] Li RY, Shi ZY. Spatial relationship and formation mechanism between traditional villages and intangible cultural heritage in the Yellow River Basin. *Econ Geogr*. 2022;42(8):205–212. Chinese.

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

[47] Gallopín GC. Linkages between vulnerability, resilience, and adaptive capacity. *Glob Environ Change*. 2006;16(3):293–303. https://doi.org/10.1016/j.gloenvcha.2006.02.004

[48] de Sherbinin A, Bukvic A, Rohat G, Gall M, McCusker B, Preston B, et al. Climate vulnerability mapping: a systematic review and future prospects. *WIREs Clim Change*. 2019;10(5):e600. https://doi.org/10.1002/wcc.600

[49] Lu JC, Liu AL. A symbiosis model for the development of intangible cultural heritage in traditional villages: taking Hanxin Village in Jiangxi Province as an example. *Resour Sci*. 2023;45(7):1396–1409. https://doi.org/10.18402/resci.2023.07.09. Chinese.

[50] Folke C. Resilience: the emergence of a perspective for social–ecological systems analyses. *Glob Environ Change*. 2006;16(3):253–267. https://doi.org/10.1016/j.gloenvcha.2006.04.002

[51] Wu HX. Research on traditional village protection from the perspective of cultural ecology. *J Orig Ecol Natl Cult*. 2017;9(1):95–100. Chinese.

[52] Li BH, Huang CY, Dou YD. Characteristics and formation mechanism of hollowing out of traditional villages in ethnic minority areas: taking Pingtan Village in Hunan Province as an example. *Prog Geogr*. 2025;44(2):343–358. https://doi.org/10.18306/dlkxjz.2025.02.010. Chinese.

[53] Gao J, Wu B. Revitalizing traditional villages through rural tourism: a case study of Yuanjia Village, Shaanxi Province, China. *Tour Manag*. 2017;63:223–233. https://doi.org/10.1016/j.tourman.2017.04.003

[54] Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32. https://doi.org/10.1023/A:1010933404324

[55] Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell*. 2020;2(1):56–67. https://doi.org/10.1038/s42256-019-0138-9

[56] Openshaw S. The modifiable areal unit problem. Norwich: Geo Books; 1984.