ARTICLE

Check for updates

Reimagining heritage villages' sustainability: machine learning-driven human settlement suitability in Hunan

Qikang Zhong[1], Liang Xie[1✉] & Jiade Wu[1]

Traditional villages play a crucial role in China's cultural, ecological, and social fabric, as they preserve indigenous knowledge and foster sustainable practices. However, rapid urbanization poses significant threats to their population stability and cultural integrity, creating challenges for both preservation and development. While existing studies often concentrate on architectural restoration or tourism, they tend to overlook a comprehensive approach to human settlement suitability (HSS). In this study, we assess the HSS of 688 traditional villages in Hunan Province by employing the entropy weight method, principal component analysis (PCA), and K-means clustering. To investigate the factors influencing HSS, we utilize Geodetector and Random Forest models. Our results indicate that 39.1% of the villages fall into the Critical Suitability category, while 29.8% belong to the General Suitability category, with overall suitability decreasing from the south to the northwest. Based on their characteristics, the villages are categorized into three types: environmentally stressed, moderately developed, and economically advantaged. Among the various influencing factors, economic variables—particularly Per Capita Disposable Income—emerge as the primary drivers of HSS variation. This research proposes an innovative framework for evaluating rural sustainability and offers strategic insights for transforming traditional villages in a sustainable manner.

[1] School of Architecture and Art, Central South University, Changsha, China. ✉email: 211301008@csu.edu.cn

Introduction

Traditional villages in China, as rich and diverse cultural heritage sites, embody a long history, unique regional cultures, and distinctive traditional residential styles (Bian et al. 2022; Gao and Wu 2017; Liu et al. 2022a; Xu and Wang 2021). Moreover, they reflect the social, religious, and cultural exchanges that have shaped Chinese civilization over thousands of years. However, the significance of traditional villages extends beyond their role as mere "living repositories of China's agricultural and communal heritage." These villages are crucial for sustaining ecological and social systems in rural environments, offering comprehensive benefits to local communities, such as maintaining biodiversity, fostering inter-generational social cohesion, and preserving cultural identity. These unique attributes highlight the importance of traditional villages not only as cultural relics but also as dynamic contributors to rural sustainability and development. In the context of rapid urbanization, traditional Chinese villages face increasing threats, including the erosion of their unique cultural characteristics and the significant degradation of human settlement systems (Chen et al. 2023; Wang et al. 2021; Zheng et al. 2021; Zhong et al. 2024). Notably, these vulnerabilities are not solely caused by urbanization itself but result from a series of intertwined processes that accompany it, such as rural-to-urban migration, agricultural abandonment, unsustainable development practices, ecological degradation, the loss of local livelihoods, and the weakening of traditional inter-generational transmission systems. For instance, rural-to-urban migration often leads to population decline, resulting in the abandonment of traditional agricultural practices, reduced eco-system maintenance, and the weakening of community networks. Similarly, modern development practices that prioritize economic benefits over environmental sustainability exacerbate ecological degradation, manifesting as deforestation, soil erosion, and water pollution, further undermining the viability of these villages. Consequently, these processes aggravate the cultural crises faced by traditional villages, as younger generations move to urban areas, leaving behind an aging population that struggles to maintain traditional lifestyles and knowledge systems. Together, these interconnected challenges pose significant threats to the survival and sustainability of traditional villages. Given this complexity, addressing the vulnerabilities of traditional villages requires more than isolated interventions or a singular focus on cultural heritage preservation. Instead, an integrated approach is essential—one that examines how environmental, social, and economic factors collectively influence the sustainability of these settlements. As a result, the study of traditional villages has emerged as a multidisciplinary focus, drawing attention from fields such as geography, tourism, history, urban planning, and architecture.

Traditional villages are not only invaluable cultural assets but also crucial nodes in broader strategies related to social cohesion, ecological sustainability, and rural revitalization. In recent years, historical studies have predominantly focused on the preservation and inheritance of the farming-reading culture characteristic of traditional villages in China (Chen et al. 2020; Gao et al. 2022; Zhang et al. 2021). In the field of architecture, researchers have emphasized comfort-oriented renovations of traditional dwellings, aiming to improve habitability while preserving cultural integrity (Hu et al. 2019; Liu et al. 2022b; Wang and Chiou 2019; Xu et al. 2021). In contrast, tourism research has primarily examined the activation and development models of traditional village tourism. Additionally, geographic studies have mainly focused on the spatial distribution of traditional village sites at the provincial, municipal, or county levels (Ma and Tong, 2022; Su et al. 2022; Wu et al. 2020; Xu et al. 2021; Yang et al. 2020), with limited attention to the village scale. In response to China's "Rural

Revitalization Strategy," the national government has implemented a series of protection policies, such as the "Chinese Traditional Village Protection Plan" and the "Historical Cultural Village Protection System," providing institutional support for the sustainable development of these villages. Current studies predominantly focus on spatial structure characteristics (Chu and Wu 2022; Ma and Tong 2022), pphysical space construction (Zhou et al. 2018), and space reconstruction (Duan et al. 2022; Xiong et al. 2022), and space reconstruction (Hu et al. 2019). However, much of this research centers on public space renewal, often neglecting the holistic characteristics of the human settlement environment in traditional villages, particularly the interconnections between environmental, social, and cultural factors.

The human settlement of traditional villages (TVHS) is not merely a reflection of a village's environmental, social, and historical contexts but also serves as the foundational spatial environment that sustains village life (Tong et al. 2022; Wang and Chiou 2019). Traditional village human settlements hold a unique and irreplaceable position in maintaining both the tangible and intangible aspects of rural life. From an ecological perspective, TVHS often represents a harmonious interaction between human activities and natural ecosystems, offering valuable insights for contemporary sustainable practices. Socially, the human settlement environment fosters community cohesion, intergenerational continuity, and cultural identity by preserving traditional lifestyles, customs, and collective memory. Economically, the quality of human settlements directly influences the viability of local livelihoods, including agriculture, handicrafts, and rural tourism —key elements for maintaining the economic resilience of these communities. Furthermore, the spatial characteristics of human settlements encode historical knowledge and adaptive strategies for resource management, disaster mitigation, and balancing ecological demands with human needs. Evidently, the quality of human settlement is not only crucial for human survival and societal progress but also constitutes a vital element in the protection of traditional villages. As such, the study of human settlements has long been a major academic concern, both domestically and internationally. Given the pressing need for sustainable development in TVHS, human settlement suitability (HSS) emerges as a theoretical framework capable of scientifically describing and assessing the quality of TVHS (Wang and Li 2018). Existing HSS research primarily focuses on urban settings (Guan et al. 2022; Jiaxing et al. 2021; Song et al. 2019), rural areas (Yu et al. 2022), and specific geographic regions, such as the Loess Plateau, the Yangtze River Delta, the Three Gorges Reservoir Area, and arid zones (Ding et al. 2021; Halik et al. 2013; Wei et al. 2022; Wei et al. 2013). Despite these advancements, there remains a significant gap in the study of traditional village HSS, particularly regarding the integration of environmental, social, and spatial factors into a comprehensive village-scale analysis. Regarding research perspectives, previous studies have predominantly focused on climate and environmental suitability (Lin et al. 2021), natural and social environmental assessments (Luo et al. 2021; Wang et al. 2017; Xiao et al. 2018), human-environment relationships (Liu et al. 2019), and spatial patterns of HSS (Guan et al. 2022; Luo et al. 2019; Xia and Dong, 2019; Xie et al. 2022). In terms of research methods, most studies have employed techniques such as the Analytic Hierarchy Process (Nandy 2021), principal component analysis (PCA) (Wang et al. 2015), entropy weight method (Wei et al. 2022; Xie et al. 2022; Zhao et al. 2013) and fuzzy comprehensive evaluation method (Xiong 2011), forming a robust set of assessment models and techniques. However, existing research on traditional villages exhibits notable limitations, primarily characterized by an over-reliance on point data paradigms, which often fail to capture the
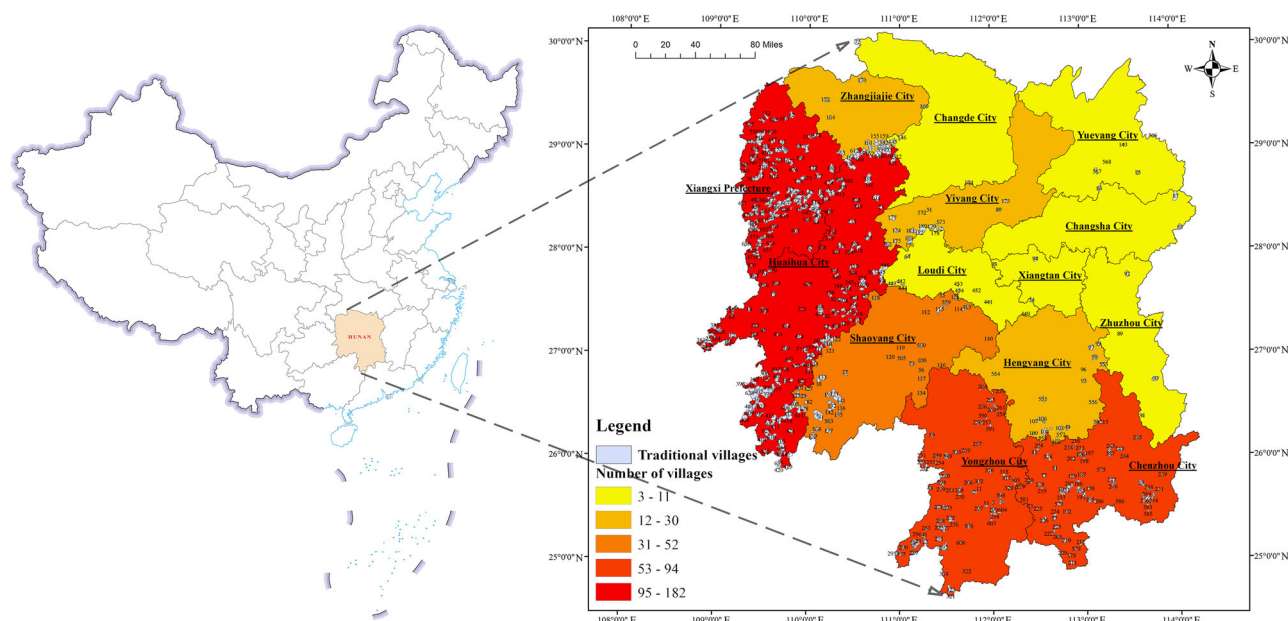
Fig. 1 Quantity distribution of traditional villages in Hunan Province (Source: author's drawing). Due to data availability, this paper only conducts research and analysis on the collected 688 traditional villages.

comprehensive environmental, social, and climatic context within traditional village spaces. In contrast, village-scale surface data offers a more holistic perspective, providing richer information to analyze the distinctive features of traditional villages and the multifaceted factors influencing them. Moreover, the unique challenges posed by the nonlinear interactions among environmental, social, and cultural dimensions in traditional villages necessitate more advanced analytical approaches beyond conventional linear methods. For instance, urban-rural issues often involve complex, nonlinear characteristics, requiring targeted strategies and nonlinear modeling techniques to address them effectively.

In recent years, studies in this field have primarily employed linear regression models, such as Ordinary Least Squares (OLS) (Jia et al. 2021; Li et al. 2023), Geographically Weighted Regression (GWR) (Li and Wu 2017; Ma et al. 2023), and quantile regression (Helbich et al. 2018), due to their simplicity and ease of interpretation. However, the use of nonlinear regression models has been on the rise, as they are more effective in capturing complex patterns and interactions between variables (Zhong et al. 2025). Particularly with the enhancement of computational capabilities and increased data availability, modeling techniques that incorporate machine learning algorithms have made significant advancements. Among these, the random forest model, a machine learning technique based on an ensemble of decision trees, has demonstrated considerable potential in addressing complex and nonlinear relationships. Unlike traditional regression models, it inherently manages nonlinear interactions between variables without the need for explicitly specifying interaction terms. Additionally, it exhibits strong resistance to overfitting and improves prediction accuracy by averaging the outcomes of multiple decision trees. Another notable advantage of the random forest model is its ability to provide variable importance measures, offering valuable insights into which predictors most significantly influence the prediction outcomes. For instance, Dumitrescu et al. (2022) illustrated how the random forest model, when combined with logistic regression, enhances prediction models by capturing the nonlinear effects of decision trees in credit scoring applications. This hybrid approach capitalizes on the strengths of both methods,

outperforming traditional logistic regression. Furthermore, Barboza et al. (2017) compared the predictive performance of the random forest model with that of logistic regression in bankruptcy prediction, reporting substantial improvements in accuracy owing to the random forest's capacity to manage nonlinear relationships and interaction variables. Despite the inherent complexity and nonlinear nature of urban-rural issues, existing research utilizing the random forest model to investigate the nonlinear impacts on traditional village human settlement suitability (TVHSS) remains insufficiently explored.

To address the aforementioned research gap, this study takes 688 traditional villages in Hunan Province as a case study and constructs a comprehensive model for assessing human settlement suitability (HSS) by employing the entropy weight method and principal component analysis (PCA). Furthermore, the K-means clustering technique is applied to categorize the main types of traditional villages. To analyze the overall influencing factors on HSS, the Geodetector model is utilized, while the Random Forest model is employed to examine how the nonlinear effects of various driving factors differ across HSS levels. This research is significant not only for the preservation and sustainable development of traditional villages in Hunan Province but also for providing valuable experience and references for village preservation in other regions, both domestically and internationally. By investigating the challenges faced by traditional villages amid modern societal development and exploring tailored strategies for different village types, this study offers actionable insights for preserving the dynamic cultural heritage inherent in traditional settlements.

Method
Study area. Hunan Province, situated in the central-southern region of China, was selected as the study area for its profound historical and cultural significance (Fig. 1). Known for its rich tapestry of cultural heritage and a long history of human settlement, Hunan's traditional villages stand as living symbols of China's rural legacy. The province is home to a diverse array of ethnic groups, including the Han, Miao, Dong, Tujia, and several other minority groups, each contributing to the unique cultural

landscapes of their respective villages. These ethnic communities have developed distinct architectural styles, social structures, and ways of life, rooted in their adaptation to the local environment over centuries. This diversity not only offers a window into the intricate socio-cultural dynamics of China's rural areas but also provides fertile ground for comparative analysis across different settlement types.

Hunan's traditional villages are unique in their ability to encapsulate both the ecological and cultural richness of the region. The diverse geographical features of the province—ranging from mountainous areas to fertile plains and river valleys—have shaped the settlements' spatial configurations and survival strategies. These villages reflect an interplay between human settlements and the natural environment, where traditional ecological knowledge has been passed down through generations, creating sustainable ways of life that are deeply integrated with the landscape. Thus, studying these villages offers a rare opportunity to examine how rural communities have historically managed their environmental resources while preserving cultural traditions.

Hunan Province serves as a valuable case study for examining how traditional villages navigate contemporary challenges, including rapid urbanization, economic development pressures, and the escalating impacts of climate change. Situated at the crossroads of central and southern China, the province's strategic location subjects its villages to varying degrees of external influence, making it an ideal site for exploring the resilience and adaptability of traditional settlement patterns. By analyzing the unique adaptive strategies employed by these villages, this research underscores the cultural and environmental resilience of rural communities while providing critical insights into broader efforts toward cultural preservation and ecological sustainability in the face of modern challenges. The historical significance of Hunan's traditional villages further elevates the importance of this research. Many of these settlements date back to the Ming and Qing dynasties, with some having been established even earlier. These villages are not only physical remnants of China's architectural past but also centers of intangible cultural heritage, where rituals, festivals, and communal activities continue to play a pivotal role in maintaining social cohesion. Hunan's villages serve as living museums, preserving ancient customs, traditional craftsmanship, and agricultural practices that are rapidly disappearing elsewhere. Exploring these villages sheds light on their enduring contribution to Chinese culture, highlighting their role as custodians of traditional knowledge systems and practices that have survived for centuries.

Furthermore, the study of Hunan Province's traditional villages offers significant theoretical contributions to the fields of geography, anthropology, rural development, and architecture. By examining the spatial organization, cultural practices, and ecological adaptations of these villages, this research helps fill critical gaps in our understanding of rural sustainability, cultural preservation, and human-environment interactions. The region's diversity provides an ideal setting for developing and refining theoretical models that can be applied to other geographical contexts, both within China and beyond. This research is poised to inform broader discussions on sustainable rural development, offering experiential knowledge and practical lessons that could benefit similar studies in other regions facing comparable challenges.

Ultimately, Hunan Province, with its blend of cultural depth and ecological complexity, represents a microcosm of China's broader rural landscape. The province's traditional villages, by standing at the intersection of history, culture, and the environment, present a rich canvas for exploring how traditional settlements can adapt and thrive amid modern pressures. This in-depth research not only contributes to the academic discourse on rural sustainability and cultural heritage but also provides a roadmap for future preservation efforts, ensuring that these invaluable cultural landscapes continue to flourish for generations to come.

Indicator selection and data sources. The HSS model can better reflect the quality of human settlement and the characteristics of regional limitations and is a scientific reference for regional economic development, resource protection, and human residential planning.

In this section, based on the HSS model proposed by previous studies (Feng et al. 2008; Tang et al. 2008; You et al. 2020), we establish the HSS model and its underlying indicator system. The indicators selected for the HSS model, as outlined in Table 1, reflect the multifaceted aspects of the environment that influence the quality of human settlements in traditional villages. These indicators are meant to assess the current state of settlement suitability by considering both natural and human-related factors. Examples of natural factors include topography, climate, hydrology, and ecology, while human-related factors cover economic activity, infrastructure, pollution, and social services. These indicators serve as a snapshot of the existing conditions within the settlements.

To avoid confusion, it is important to note that the HSS indicators reflect the concept of a settlement for human habitation. They do not capture the underlying causes of these conditions, but rather, they measure how suitable a particular area is for settlement based on existing factors. For instance, indicators such as average elevation, annual precipitation, road network density, and population density provide insight into the overall quality of the environment, which directly informs how suitable an area is for sustaining human life.

While these indicators offer a valuable cross-sectional snapshot, they are distinct from the driving factors that influence how settlement suitability evolves over time. The HSS indicators focus on measuring suitability at a specific moment in time, but they do not address the underlying causes that affect changes in this suitability. In addition, the detailed index calculation method can be seen in Appendix Eq. A.

The list of traditional villages was obtained from the Traditional Villages of China website (http://www.chuantongcunluo.com/). The administrative boundaries required for traditional villages were mainly obtained from the National Public Service Platform for Geographic Information (https://www.tianditu.gov.cn/). Data for natural environment factors (topography, climate, ecology, pollution, etc.), traffic (roads, railways, traffic nodes, etc.), social factors (population density, GDP, POI nodes, etc.), and river data were sourced from the Geographic Remote Sensing Ecology Network (http://www.gisrs.cn/) and the Resource and Environment Science Data Centre of the Chinese Academy of Sciences (https://www.resdc.cn/). Socioeconomic statistics were derived from the China County Statistical Yearbook and the Hunan Statistical Yearbook, supplemented by government reports and statistical bulletins on national and county-level economic and social development.

While the data used in this study is primarily sourced from regional-level datasets, efforts were made to aggregate and adapt the data to represent village-level conditions as accurately as possible. For example: Topographic, hydrological, and ecological data were sourced from high-resolution datasets (e.g., 100 m precision) and are considered suitable for village-level analysis; Climatic and pollution data, sourced from lower-resolution datasets (e.g., 1000 m precision), may present some limitations when applied to individual villages but still provide valuable

Table 1 TVHSS indicator system.

Criterion layer	Factor layer	Indicator layer	Indicator attribute
Nature category (A1)	Topographic index (B1)	Degree of relief (C1)	-
		Average elevation (C2)	-
		Slope (C3)	-
		Aspect (C4)	-
	Climatic index (B2)	Annual average temperature (C5)	-
		Annual average relative humidity (C6)	-
		Annual average wind speed (C7)	-
		Annual average air pressure (C8)	-
	Hydrological index (B3)	Precipitation (C9)	+
		Distance from water area (C10)	-
	Ecological index (B4)	NDVI (C11)	+
		Ecosystem service value (C12)	+
		Land use intensity (C13)	-
		Land use mix (C14)	+
Human category (A2)	Economic index (B5)	Commercial facility density (C15)	+
		Distance from commercial facility (C16)	-
		Primary industry added value (C17)	+
		Secondary industry added value (C18)	+
		Industrial enterprise above designated size (C19)	+
		GDP (C20)	+
	Traffic index (B6)	Distance from railway (C21)	-
		Road network density (C22)	+
		Distance from bus station (C23)	-
		Distance from road (C24)	-
	Pollution index (B7)	CO_2(C25)	-
		O_3(C26)	-
		SO_2(C27)	-
		PM10(C28)	-
		PM2.5 (C29)	-
	Social index (B8)	Nighttime lighting index (C30)	+
		Population density (C31)	-
		Hospital bed quantity in medical and health institutions (C32)	+
		Distance from medical institution (C33)	-
		Distance from educational institution (C34)	-

For the direction of effects, "+" indicates positive feedback, and the higher the value is, the higher the suitability index; "-" indicates negative feedback, and the higher the value is, the lower the suitability index.

insights at a broader level; Economic data are based on county-level statistics (e.g., GDP, commercial density), which are not directly representative of village-scale data but offer a general view of the regional economic context; Traffic and social index data are fully applicable at the village level, with data derived from local infrastructure and demographic statistics.

Furthermore, it is important to note that the data collected for this study covers the year 2020. As such, the analysis reflects the conditions in that year, and any temporal changes in village-level suitability since then are not captured in this study.

Driver variable setting. The driver variables, as described in this section, differ from the HSS indicators in their purpose and role within the research framework. These variables are dynamic in nature and reflect underlying social, economic, and cultural factors that drive changes in the settlement environment. Rather than evaluating the current state of suitability, the driving factors focus on variables that may affect settlement suitability in the future.

Inspired by existing research findings (Bao et al. 2021; Chen et al. 2023; Kang et al. 2021; Tang et al. 2017; Tian et al. 2023; Wang and Chiou 2019; Xin et al. 2020), seven variables, including Total Sown Area of Crops, Total Grain Production, Per Capita Disposable Income of Urban Residents, Per Capita Retail Sales of Consumer Goods, Distance to Intangible Cultural Heritage Sites, Housing Price, and Number of Social Welfare Adoption Units, are selected as influencing factors for the level of TVHSS. Total Sown Area of Crops and Total Grain Production reflect the

agricultural economic conditions of traditional villages. These indicators can reveal the agricultural production capacity and sustainability of villages, which significantly influence the HSS (Bao et al. 2021; Xin et al. 2020). Per Capita Disposable Income of Urban Residents and Per Capita Retail Sales of Consumer Goods focus on village residents' economic conditions and living standards. Per Capita Disposable Income of Urban Residents reflects the economic welfare of residents, while Per Capita Retail Sales of Consumer Goods can indicate the quality of life and purchasing power, guiding the HSS (Chen and Shen 2021; Wang et al. 2019). The intangible cultural heritage of traditional villages holds significant importance for the inheritance of local culture and the uniqueness of the human settlement. Distance to Intangible Cultural Heritage Sites can reflect the extent of traditional culture in the village, influencing residents' lifestyles and community identity (Chen et al. 2023). Housing Price reflects the state of the real estate market and can indirectly indicate the region's level of development and economic conditions. The correlation between Housing Price and the HSS is high, as it involves housing conditions and community infrastructure (Kang et al. 2021). The Number of Social Welfare Adoption Units includes various social service facilities such as schools, hospitals, and welfare institutions. These facilities have a crucial impact on residents' quality of life and the level of social services, closely related to the HSS (Müller 2020).

To clarify, the driving factors are the root causes that influence the changes in settlement suitability over time. They are more

5

Table 2 Driver variable setting.

Variable type	variable name	variable unit
Explained variable	Traditional village human settlement suitability	/
Explanatory variables	Total Sown Area of Crops	Thousand hectares
	Total Grain Production	Ton
	Per Capita Disposable Income of Urban Residents	Yuan
	Per Capita Retail Sales of Consumer Goods	Ten thousand yuan
	Distance to Intangible Cultural Heritage Sites	m
	Housing Price	Million yuan/m^2
	Number of Social Welfare Adoption Units	/

concerned with understanding how various factors interact and affect long-term outcomes. For example, housing prices may not immediately influence the quality of life in a settlement, but over time, rising prices may lead to changes in infrastructure development, migration patterns, and the overall economic landscape. Similarly, the availability of social welfare services may improve residents' quality of life and, in turn, drive improvements in the settlement's future suitability.

Unlike the HSS indicators, which assess the current quality of the settlement environment, the driving factors provide insight into how the settlement environment may evolve due to these underlying forces. While the indicators provide a static view, the driver variables offer a dynamic understanding of the processes that lead to changes in settlement suitability. By examining these factors, this study can better predict future trends and understand the forces at play in shaping human settlement dynamics Table 2.

Research methods

Weight assignment by entropy weight method and principal component analysis. In assessing Human Settlement Suitability (HSS), indicators are categorized into positive and negative types, reflecting their varying impacts on the overall suitability. To account for the disparate units of measurement across these indicators, a process of normalization is applied to render all data dimensionless and comparable.

Given the potential subjectivity of manually assigning weights to these indicators, this study employs two widely used and objective methods to determine the factor weights: the Principal Component Analysis (PCA) and the entropy weight method. These methods are chosen for their ability to reduce subjectivity and enhance the reliability of the results.

Entropy weight method calculates the information entropy of each indicator, which reflects the degree of variation in the data. Indicators with more variation provide more information and thus receive higher weights. This method allows for the automatic calculation of weights, ensuring that the most informative indicators are given more importance in the final assessment. The entropy weight method is applied at the indicator layer to determine the weights of the individual indicators. The rationale behind using this method is that indicators with higher entropy (greater variability) contribute more significantly to the model, as they provide more information for evaluating human settlement suitability. This ensures that indicators with low variability or less meaningful contribution are assigned lower weights, effectively eliminating arbitrary subjectivity in manual weight assignment.

Factor weights at the factor layer. PCA is a widely used statistical method that identifies the principal components of a dataset, which are linear combinations of the original indicators. The principal components account for the maximum variance in the data and are used to reduce the dimensionality of the dataset while retaining as much information as possible. The rationale for using PCA lies in its ability to combine correlated indicators into new variables (principal components), which can more effectively represent the

underlying patterns of the data. PCA helps to identify the most important factors (combinations of indicators) that explain the variance in HSS. This further contributes to a more robust and holistic assessment by simplifying the data structure without losing important information. The loading values of each principal component determine the contribution of the corresponding indicator to the overall factor, which are then used to compute the final weight for each indicator at the factor level.

In this approach, the combination of entropy weights and PCA ensures a comprehensive and data-driven weighting system, where the entropy method ensures that more informative indicators receive appropriate weights, while PCA identifies the core factors that explain the variability in the data. This dual-method approach minimizes subjectivity and provides a more objective and transparent way of determining the weights for the Total HSS Evaluation Model. Therefore, this approach establishes a comprehensive evaluation model, to calculate the Total Value of HSS Table 3. Look the Appendix Eq B for the detailed steps.

Cluster analysis of HSS using the K-means clustering and principal component analysis. This study first explored the HSS of various regions by clustering them based on a set of HSS indicators. The methodology combines the power of K-Means clustering with Principal Component Analysis (PCA) for dimensionality reduction. K-Means is a popular unsupervised machine learning algorithm for partitioning a dataset into a set of distinct, non-overlapping subgroups. One of its main advantages is the simplicity of its underlying assumptions that clusters are spherical and equally sized, enabling the algorithm to be robust, reliable, and fast. This makes K-Means especially suited for our needs to efficiently segment a large dataset (Kim and Kim 2023). By synergizing the Elbow method, three distinct categories are established for classification. In our analysis, the K-Means algorithm from the scikit-learn library of Python was used to categorize traditional villages based on their HSS indicators outlined in Table 1 (Pedregosa et al. 2011).

Principal Component Analysis is a dimensionality reduction technique that transforms high-dimensional data into a lower-dimensional form, ensuring that the variance in the data is maximized in the first few dimensions (Liu et al. 2003). This is particularly useful in visualizing high-dimensional data and understanding the major underlying patterns. Another advantage of PCA is data compression without a significant loss of information, hence facilitating faster computation with reduced resources (Jolliffe and Cadima 2016). In our study, PCA was employed to reduce the 8-dimensional factor layers to two dimensions, making it conducive for visualization on a 2D plane. The reduced components were then used to visualize the K-Means clustering results.

Influencing factors analysis by the geodetector model and random forest model. This study primarily uses the geodetector model to reveal the extent to which driving factors affect HSS. The

Table 3 TVHSS indicator weight system.

Criterion layer	Factor layer	Factor weight	Indicator layer	Indicator weight
Nature category (A1)	Topographic index (B1)	0.1543	Degree of relief (C1)	0.0101
			Average elevation (C2)	0.0062
			Slope (C3)	0.0109
			Aspect (C4)	0.0079
	Climatic index (B2)	0.1288	Annual average temperature (C5)	0.0073
			Annual average relative humidity (C6)	0.0088
			Annual average wind speed (C7)	0.0073
			Annual average air pressure (C8)	0.0127
	Hydrological index (B3)	0.1263	Precipitation (C9)	0.0069
			Distance from water area (C10)	0.0236
	Ecological index (B4)	0.1182	NDVI (C11)	0.0113
			Ecosystem service value (C12)	0.0594
			Land use intensity (C13)	0.0133
			Land use mix (C14)	0.023
Human category (A2)	Economic index (B5)	0.1192	Commercial facility density (C15)	0.0876
			Distance from commercial facility (C16)	0.0233
			Primary industry added value (C17)	0.0933
			Secondary industry added value (C18)	0.0907
			Industrial enterprise above designated size (C19)	0.0266
			GDP (C20)	0.034
	Traffic index (B6)	0.1318	Distance from railway (C21)	0.0204
			Road network density (C22)	0.054
			Distance from bus station (C23)	0.0227
			Distance from road (C24)	0.0394
	Pollution index (B7)	0.1138	CO_2(C25)	0.0923
			O_3(C26)	0.0123
			SO_2(C27)	0.0168
			PM10(C28)	0.0111
			PM2.5 (C29)	0.0112
	Social index (B8)	0.1076	Nighttime lighting index (C30)	0.0594
			Population density (C31)	0.0511
			Hospital bed quantity in medical and health institutions (C32)	0.0142
			Distance from medical institution (C33)	0.0146
			Distance from educational institution (C34)	0.0163

Geodetector model can simultaneously analyze changes across multiple scales, revealing the extent of the impact of different-scale factors. By assessing the significance of interactions and overall effects, the Geodetector model can help uncover the influencing mechanisms and identify interactions among different variables, further exploring how variables interact spatially. This study calculates the impact level (q-value) of the detection factors (the single indicator scores) X on HSS Y using the factor detection module. Correlation analysis is used to determine the direction of influence for each detection factor, and if the detection factor is positively correlated with HSS, the factor is considered to be positively influenced with a positive q value and vice versa (Cao et al. 2013; Zhang et al. 2022). The model calculation formula is detailed in Appendix Eq C.

Random Forest model is an ensemble learning method for classification and regression that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by the most trees. In regression tasks, it returns the mean or average prediction of the individual trees. Random Forest model corrects for decision trees' habit of overfitting to their training set, providing a more generalized and robust model (Biau 2012; Speiser et al. 2019). In this study, the Random Forest model was employed to analyze the impact of seven external factors on the Traditional Villages Human Settlement Suitability (TVHSS). The model's capability to process large datasets with multiple input variables, combined with its robustness against overfitting, makes it particularly suitable for addressing the complex and nonlinear relationships inherent in socio-economic data. By leveraging the

Random Forest model, the study not only identified the importance ranking of these factors but also revealed their threshold effects on TVHSS. These insights offer valuable and reliable predictions, contributing to a deeper understanding of the dynamics influencing traditional villages and potentially guiding effective developmental policies.

Results
HSS analysis
Spatial distribution of overall HSS. The critical suitability area is the region with the highest proportion in the assessment of the TVHSS, accounting for 39.10% of the total number of traditional villages in Hunan Province (Table 4). The general suitability area ranks second in suitability assessment, with 205 traditional villages, accounting for 29.80% of the total. Although the unsuitable area is relatively small in size, it still includes 109 traditional villages, accounting for 15.84% of the total. The moderately suitable area is relatively small both in terms of area and the number of traditional villages, with 734.77 km^2 and 62 traditional villages, respectively, accounting for 9.01% of the total. The highly suitable area is the smallest both in terms of area and the number of traditional villages, with 372.70 km^2 and 43 traditional villages, respectively, accounting for 6.25% of the total.

As shown in Fig. 2, among the traditional villages participating in the assessment, the composite human settlement suitability index (HSSI) mainly ranges from 0.1 to 0.5, with some traditional villages having a higher index, with an average value of 0.2823 and a poorer overall quality. Hejiatian Village in Huaihua City has the highest score, while Shangbao Village in Shaoyang City

Table 4 Statistics of TVHSS evaluation.

Suitability grading	Area/km²	Proportion/%	Quantity of villages/EA	Percentage of villages/%
Unsuitable	1589.33	20.71%	109	15.84%
Critically suitable	3132.17	40.81%	269	39.10%
General suitable	1845.30	24.05%	205	29.80%
More suitable	734.77	9.57%	62	9.01%
Highly suitable	372.70	4.86%	43	6.25%

Fig. 2 Distribution of TVHSS scores. All suitability indices are standardized.

has the lowest score, with obvious differences in TVHSS. The distribution is shown as "oscillation-type" in the sense of "low overall and high partly". This is probably due to the rugged terrain, inconvenient transportation, relatively poor economic development, and sparse population distribution. The humanity suitability index is similar to the composite HSSI. The natural suitability index mainly ranges from 0.1 to 0.4, with an average value of 0.2842. The overall index is low, and the natural environment is poor. Hejiatian Village in Huaihua has the highest score, while Tiexulin Village in Shaoyang City has the lowest score, with a difference of 4.5 times. The difference in the natural suitability of traditional villages is extremely obvious, with an overall "olive-shaped" distribution of "small at the two ends and large in the middle".

The overall low natural HSS indicates that traditional villages are mostly in unsuitable or critically suitable areas in terms of topography, climate, hydrology, and ecology, or there are at least two types of unsuitable factors in combination (as shown in Fig. 3a). The high values show a certain clustering, mainly in the plain areas or areas with gentle terrain, and the natural suitability has a strong limiting effect on the TVHS. From the evaluation of humanity suitability (Fig. 3b), the high values are mainly

distributed in the vicinity of water areas, traffic arteries, and areas with good social conditions. Water resources are the basic natural resources for survival and development, and the hydrological conditions are crucial to the HSS of an area; traffic accessibility has a certain direction for the development of HSS; social urbanization plays a double-edged role in HSS, which can make the development of social urbanization have a double-edged effect on HSS, which may cause some traditional villages to increase or decrease in HSS.

The results of the synthetically superposed factors of HSS are shown in Fig. 3c. The overall suitability decreases from south to northwest, from river valleys and basins to mountains and hills. Traditional villages in this region are also plain areas in terms of topography; the more suitable areas account for 9.01% and are mainly located in the northern end of Huaihua City and the central-southern part of Huaihua City. As a potential area for the improvement of human settlement, it can develop green industries and build eco-towns based on its natural advantages and energy resources; the highest percentage of generally suitable areas is 29.80%, which is the transition area between good and bad suitability, and should reduce human interference and gradually restore its ecological functions; the percentage of
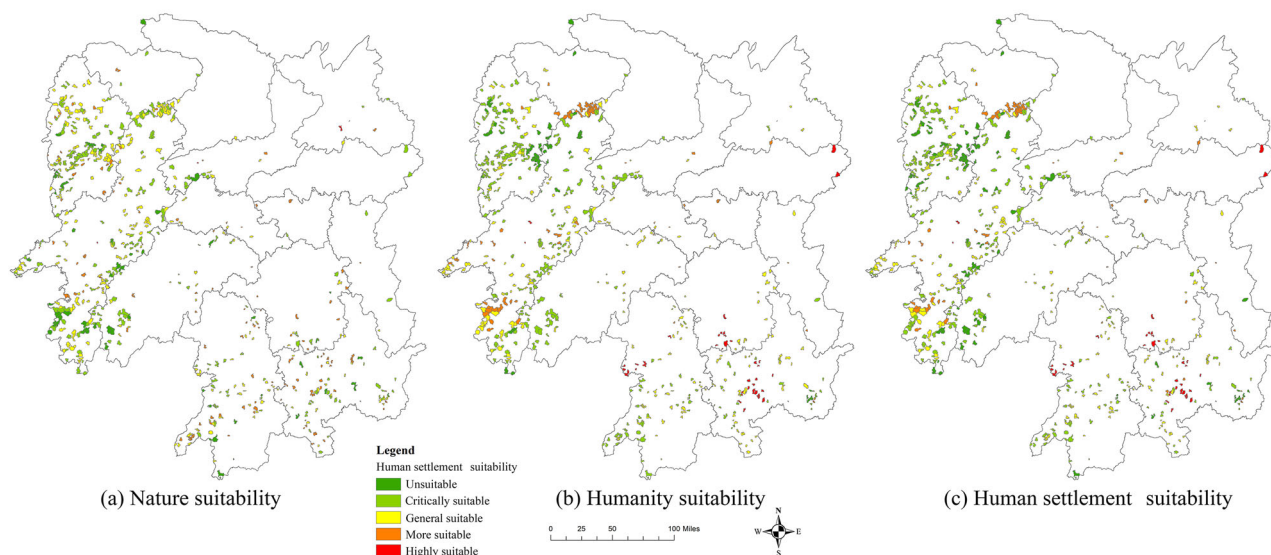
Fig. 3 Spatial distribution of TVHSS evaluation (Source: author's drawing).

critically suitable and unsuitable areas is more than half, reaching 54.94%, which are ecologically deplorable or need key protection areas, and should avoid any development and construction activities to prevent ecological deterioration. The spatial differentiation in the suitability of traditional villages in Hunan Province for human settlement is notably significant. To address this issue, Hunan Province needs to adapt strategies to local conditions, emphasize the unique characteristics of different regions, and implement rational planning. Efforts should focus on continuously optimizing the spatial development patterns of rural areas to promote sustainable rural development. County towns play a critical bridging and linking role between urban and rural areas, enhancing the functionality and vitality of traditional villages.

Spatial distribution of HSS in each city. As seen from Fig. 4, it is evident that the TVHSS in Hunan Province exhibits significant regional differences, showing a distribution pattern of "protrusion in the central-eastern region and depression in the southwestern-western region." Overall, the traditional villages in Xiangdong (Yueyang, Changsha, Xiangtan, Zhuzhou) and Xiangnan (Hengyang, Chenzhou) regions receive higher suitability assessments for human settlements, while traditional villages in Xiangbei (Changde) and Xiangxi (Xiangxi Prefecture, Huaihua, Shaoyang, Yongzhou) regions receive lower suitability assessments. Among them, traditional villages in Changsha and Hengyang, belonging to their respective cities, receive relatively higher assessments and become dual core areas in terms of the TVHSS in Hunan Province. Among the highly suitable traditional villages, those in Changsha account for the highest proportion, reaching 66.67%. This phenomenon may be attributed to the fact that highly suitable traditional villages typically require strong economic support, government planning, and policy support. As a sub-provincial city in Hunan Province, Changsha has a relatively high level of economic development, providing a solid foundation for improving the TVHS.

The research results reveal differences in the TVHSS among different cities (Fig. 5). In Chenzhou City, the suitability of traditional villages for human settlement is higher compared to the other eight cities, demonstrating a relatively balanced performance with smaller disparities. Hengyang City also exhibits

a high overall suitability level, although it is slightly lower than that of Chenzhou City, with a trend of gradual decline. The situations in Yongzhou City and Hengzhou City are similar, but there is a noticeable difference between the eighth and ninth-ranked villages, showing a declining trend with a gap of 0.239. As for Huaihua City, despite having traditional villages with relatively high suitability levels, it exhibits an overall significant decreasing trend. The highest index can reach 1.523, while the index for the tenth-ranked village is only 0.704, with a gap of up to 0.819. The situations in Zhangjiajie City and Shaoyang City are similar, and stable but not outstanding. Shaoyang City's development is slightly behind that of Huaihua City and Zhangjiajie City, but the gap is small, and both show a relatively stable development trend. As for Xiangxi Prefecture, despite having one traditional village with a high suitability level, the overall suitability level is lower, showing a continuous declining trend. The situations in Loudi City and Yiyang City are similar, with only a few traditional villages having suitability index levels higher than the average (0.5489). Their development levels are relatively low, and they exhibit a more significant declining trend.

According to the ranking of the top ten and bottom ten traditional villages in HSS (Table 5), there is a high level difference in HSS among the traditional villages. The study found that Hejiatian village has the highest HSS score (1.5228), and Shangbao village has the lowest (0.1848), with a difference of more than eight times. The average score of the whole province (0.5489) is closer to the score of the bottom ten traditional villages; Changsha has only three traditional villages, but two of them are ranked highly in terms of HSS. There are significant differences in the levels of TVHSS across Hunan Province, and overall, the development of traditional villages is extremely uneven and low, but Changsha has a high level of development and protection for traditional villages, which can be emulated by other cities.

Spatial distribution analysis of the HSS components. Based on the results of cluster analysis, traditional villages can be categorized into three distinct types: environmentally stressed villages ($n = 373$), moderately developed villages ($n = 252$), and economically and ecologically advantaged villages ($n = 63$), each
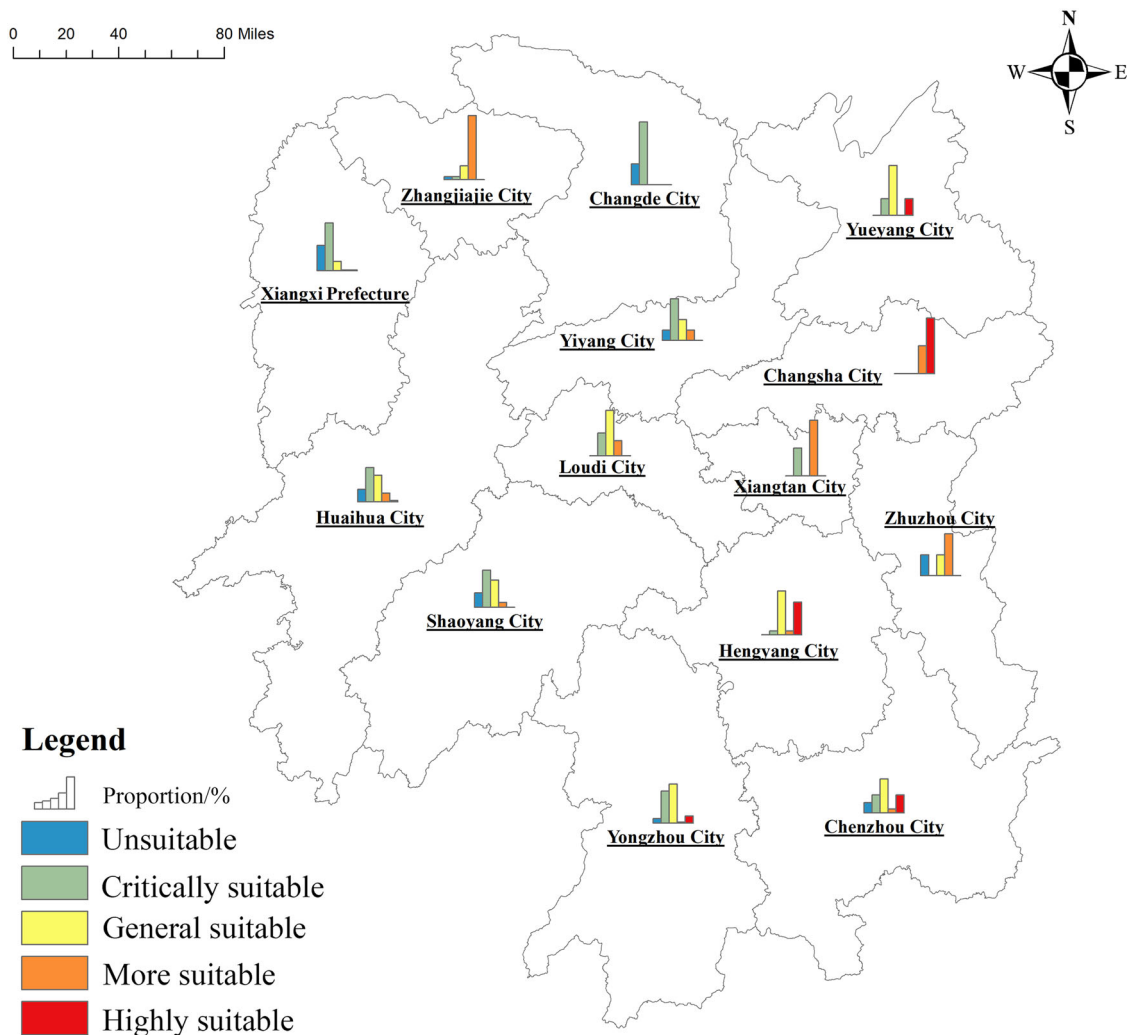
Fig. 4 Spatial distribution of the proportion of different grades of traditional villages (Source: author's drawing).

with its unique characteristics (Fig. 6). Environmentally stressed villages are characterized by flat topography, mild climate, and non-significant ecological features, coupled with limited economic and social activities, but facing high environmental pollution pressures (Table 6). Villages in this cluster have relatively low scores in standardized topography (0.51 ± 0.11), climate (0.59 ± 0.16), and ecology (0.05 ± 0.02), suggesting they may be located in areas with flatter terrain, more temperate climatic conditions, and less pronounced ecological features. The economic and social indicators (0.12 ± 0.10 and 0.52 ± 0.10 respectively) are low, implying limited economic activities and social development. The high standardized pollution indicator (0.96 ± 0.05) may reflect significant environmental pollution or ecological stress. Moderately Developed Villages exhibit moderate to high diversity in topography and climate, moderate ecological characteristics, and a medium level of social and economic activities, with relatively good environmental conditions. These villages score moderately high in standardized topography (0.70 ± 0.10), climate (0.75 ± 0.11), and ecology (0.07 ± 0.02), showing varied terrain and climatic conditions along with moderate ecological features. The social and economic indicators (0.62 ± 0.09 and 0.13 ± 0.11) indicate that these areas have more active social and economic activities. The slightly lower pollution index (0.93 ± 0.07) indicates a comparatively cleaner environment. Economically and Ecologically Advantaged Villages excels in topography, climate, and ecological aspects, along with robust

economic growth and good transportation conditions. They exhibit the highest scores across all indicators. These villages have the highest scores in standardized topography (0.74 ± 0.13), climate (0.57 ± 0.22), and ecology (0.13 ± 0.09), indicating regions with distinctive terrain, climate, and ecological features. They also score highest in standardized economy (0.42 ± 0.28) and transportation (0.54 ± 0.12), suggesting strong economic development and favorable transportation conditions. Social indicators are also relatively high (0.67 ± 0.16), but the pollution index is the lowest (0.72 ± 0.20), reflecting good environmental conditions and lower levels of pollution.

Ecology, Topography, and Pollution indexes are the three most significant contributors to the clustering results, with respective values of 1.00, 0.66, and 0.63. Conversely, the Hydrology, Society, and Climate indexes have relatively minor impacts on the clustering outcome, with contribution values of 0.01, 0.24, and 0.27, respectively. The average value of the topographic index is 0.6002, most traditional village areas have large topographic undulations and a relatively harsh topographic environment (Fig. 7). There are obvious differences in the topographic values of the traditional villages, with large fluctuations, and the range of differences is mainly concentrated between 0.4 and 0.9. The topographic unsuitability area accounts for the largest proportion, mainly in the mountainous areas, and the highly suitable area accounts for the smallest proportion. The highest score for the topographic index is recorded for the Changxin community
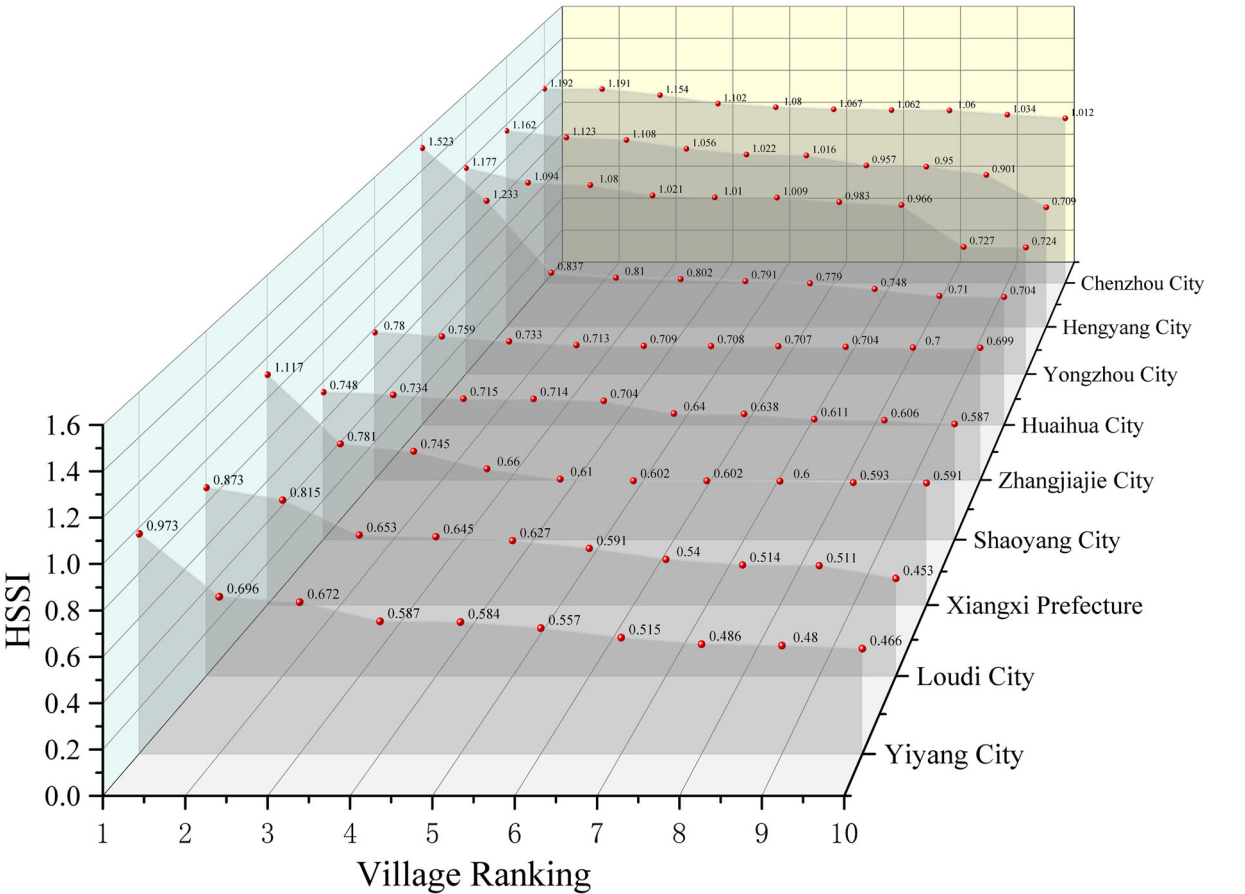
Fig. 5 Comparison chart of HSS of the top ten traditional villages in each city (Source: author's drawing). The quantity of traditional villages in some cities is less than 10, so they are not included in this comparison.

Ranking	Yiyang City	Loudi City	Xiangxi Prefecture	Shaoyang City	Zhangjiajie City	Huaihua City	Yongzhou City	Hengyang City	Chenzhou City
1	Huangshaping community	Huishui village	Liye community	Qingshui village	Qiaobianhe village	Hejiatian village	Shanmuqiao village	Zhongtian village	Gangjiao village
2	Huayuandong village	Sanjia village	Xima community	Kuli village	Matouxi village	Jingping village	Daqingping community	Shuangwan village	Jinjiang village
3	Tangshi community	Louxia village	Liandong village	Yangjiajiang village	Taiyangshan village	Jinyuan village	tianjiawan village	Guangrong village	Zhujiawan village
4	Daan village	Zhengcun village	Lahao village	liuli village	Zijingta village	shangping village	Fujiangzai village	Dayu village	Duishang village
5	Dongshi community	Tongpen village	Dongma village	wenquan village	Laopeng village	Huangxi village	Fenxiang village	Xuantang village	Wushan village
6	Sutuan village	Shangxi village	Aizhai community	Tangtian village	Shuangxiqiao village	qiaotou village	Dapikou village	Shipan village	Aoshang village
7	Huanghuaxi village	Xianghua village	Zixia village	Yanglin village	Hepinggang village	Tuqiao village	Xingmuyuan village	Liutu village	Changchong village
8	Shuangyan village	Xiatuan village	Ximenkou village	Liujia Village	Jiucaiya village	Shaotian village	Jianyantou village	Dahong village	Zhengyuan village
9	Meishan village	Wuxing village	Molao village	Xicunfang village	Mushan village	Liaoyuan village	Baijia village	Wuchong village	Gaoyaling village
10	Xintan village	Shangtuan village	Hongping village	Sanmen village	Huangjiahe village	Dongyuesi village	Bachi village	Baogai village	Liangwandong village

(number 568), while the lowest scores are recorded for Quanping Village (number 72) and Liaocheng Village (number 170). The average value of the ecological index is 0.0620, with an overall extremely low value and poor ecological quality. There is a slight difference in values among traditional villages, and the range differences are mainly concentrated between 0.01 and 0.3. the ecological indices of the unsuitable areas and the critically suitable areas of human settlement account for more than 85% of the total, which is almost distributed in all traditional villages, with unsuitable areas concentrated in the central part of Xiangxi Prefecture and the southern to the northern part of Huaihua City. The highest ecological index score is for Hejiatian Village (number 62), and the lowest scores are for Lingxitang Village (number 363) and Yanshi Village (number 138). The average value of the pollution index is 0.9269, and the pollution environment quality is very poor. There are slight differences in the values of the traditional villages, with the range of differences mainly concentrated between 0.6 and 1. The highly suitable areas are mainly located in Xiangxi Prefecture and Huaihua City, where the rural revitalization and industrialization processes have led to a significant increase in atmospheric pollution. The relatively higher suitability areas are primarily found in the central part of

Chenzhou City, where human perception is relatively comfortable. Pollution levels in these regions exhibit an increasing trend from northwest to south. The highest score on the pollution index is for Xiqi Village (number 681), and the lowest scores are for Yunshan Village (number 234) and Hejiatian Village (number 62).

Driving factors of HSS

Driving factors of HSS by geodetector. The spatial differentiation of the TVHSS is mainly influenced by economic factors, particularly consumer behavior and agricultural output (Table 7). The results indicated that Per Capita Retail Sales of Consumer Goods (X4) and Total Grain Production (X3) emerged as the most influential factors, with q statistics of 0.395 and 0.373 respectively. This suggests a substantial correlation between the economic activities of a region, particularly in retail and agriculture, and the suitability of its traditional villages for human habitation. The high q statistic for retail sales underscores the link between consumer spending habits and the quality of living environments, possibly reflecting the economic prosperity and availability of amenities in these areas. Similarly, the significance of grain production highlights the impact of agricultural productivity on rural

living conditions, possibly due to its implications for food security and rural income. Conversely, the number of Social Welfare Adoption Units (X5) and Total Sown Area of Crops (X2) exhibited the least influence, with q statistics of 0.044 and 0.108, respectively. This indicates a relatively minor direct impact of these factors on the TVHSS, compared to more dominant economic indicators.

Driving factor of different levels in HSS by random forest model. This study employs a random forest model to determine the nonlinear relationships of driving factors on TVHSS. The importance results of the random forest model indicate that X3 is the most important, ranking as follows: X3 > X4 > X6 > X5 > X1 > X2 > X7. The last four indicators, although significant,

Table 5 The top ten and bottom ten TVHSS scores in Hunan Province.

City	Village	Serial number	HSSI	Ranking
Huaihua City	Hejiatian Village	62	1.5228	1
Huaihua City	Jingping Village	325	1.2332	2
Changsha City	Tanwan Village	88	1.2317	3
Changsha City	Chudong Village	87	1.2105	4
Chenzhou City	Gangjiao Community	198	1.1922	5
Chenzhou City	Jincun Village	233	1.1908	6
Yongzhou City	Shanmuqiao Village	250	1.1771	7
Hengyang City	Zhongtian Village	558	1.1619	8
Chenzhou City	Zhujiawan Village	197	1.1536	9
Hengyang City	Shuangwan Village	107	1.1225	10
Xiangxi Prefecture	Pojiao Village	486	0.2303	679
Huaihua City	Huabi Village	417	0.2273	680
Huaihua City	Pingtou Village	632	0.2254	681
Huaihua City	Laozhai Village	419	0.2181	682
Yongzhou City	Niulu Community	321	0.2163	683
Xiangxi Prefecture	Wulong Village	66	0.2158	684
Shaoyang City	Tieshanlin Village	9	0.2114	685
Changde City	Quanping Village	72	0.2096	686
Shaoyang City	Baijiao Village	122	0.2070	687
Shaoyang City	Shangbao Village	81	0.1848	688
The average score of 688 villages	-	-	0.5489	-

have a multiple difference in importance values compared to the top three indicators (Fig. 8).

Additionally, according to Fig. 9, X3 shows a continuous positive influence on TVHSS, but when X3 reaches around 41,000 yuan, its positive impact tends to plateau; X4 has a continuously rising positive effect on TVHSS, with a sudden increase when X4 reaches about 25,000 yuan, and the positive impact tends to level off when X4 reaches about 28,000 yuan; X6 shows a fluctuating positive impact on TVHSS, with a pronounced positive effect when X6 is between 4000 and 5000 yuan, a sudden increase when exceeding 6000 yuan, and a leveling off near or at 7000 yuan. Contrarily, X5 has a continuous negative impact on TVHSS. It is noteworthy that X1's positive impact on HSS shows an irregular fluctuating trend, especially when X1 reaches 100–110 thousand hectares, where it causes a sudden surge and peaks in its positive effect. X2's positive impact on TVHSS shows a band-like upward trend. When X7 approaches 100 units, its positive impact on TVHSS reaches its peak, and the level of impact almost remains unchanged thereafter.

Discussion

Important findings on the TVHSS. In this study, a comprehensive assessment of TVHSS was conducted, revealing significant spatial variations among traditional villages in Hunan Province. As previously discussed, traditional villages in Chenzhou City demonstrate a higher level of suitability, primarily attributable to strong economic development driven by effective government policies and strategic regional planning. In contrast, villages in Huaihua City face considerable challenges, characterized by a notable decline in suitability. This disparity underscores the need for more targeted attention and policy support in areas exhibiting low suitability levels.

These findings are consistent with previous research, which highlights the pivotal role of economic factors in determining human settlement suitability. However, the implications of these results extend beyond merely identifying spatial disparities. To effectively address the needs of these villages, it is crucial to situate them within the broader context of rural sustainability and the preservation of traditional villages. Specifically, it is essential to recognize that regions with differing suitability levels require customized strategies for sustainable development.

Sustainable village transformation strategies should not only address economic factors but also encompass the cultural and ecological dimensions of these villages. For instance, in high-suitability areas, priority should be given to preserving cultural heritage and socio-ecological systems. Policies in these regions should aim to maintain ecological balance, enhance cultural practices, and promote green development. Conversely, in low-
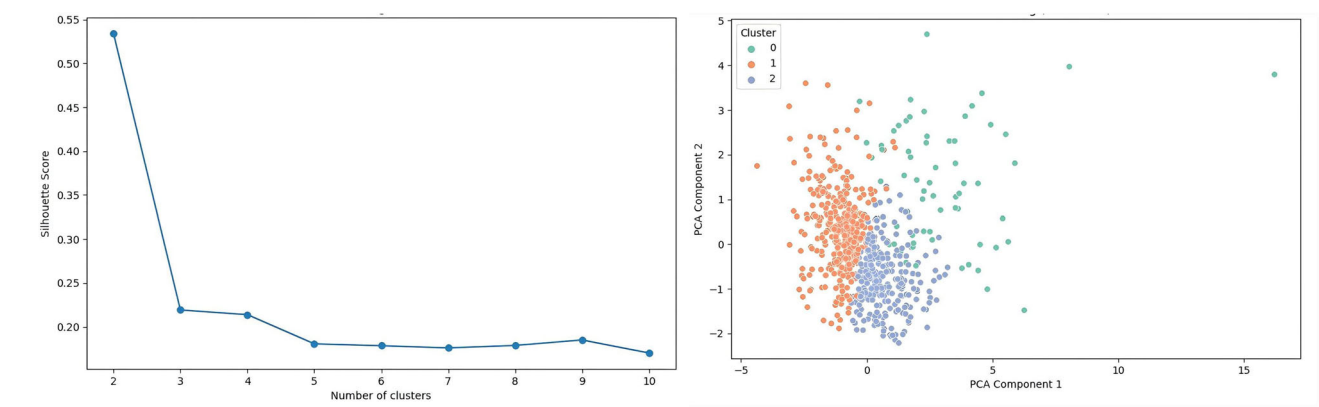

Fig. 6 Cluster results of the traditional villages (Source: author's drawing).

Table 6 Descriptive analysis of the three types of traditional villages.

	Cluster			F	p
	cluster_2($n = 373$)	cluster_1($n = 252$)	cluster_0($n = 63$)		
Topographic index (B1)	0.51 ± 0.11	0.70 ± 0.10	0.74 ± 0.13	203.016	0.000**
Climatic index (B2)	0.59 ± 0.16	0.75 ± 0.11	0.57 ± 0.22	64.172	0.000**
Hydrological index (B3)	0.72 ± 0.17	0.75 ± 0.18	0.77 ± 0.13	2.85	0.037*
Ecological index (B4)	0.05 ± 0.02	0.07 ± 0.02	0.13 ± 0.09	391.498	0.000**
Economic index (B5)	0.12 ± 0.10	0.13 ± 0.11	0.42 ± 0.28	109.613	0.000**
Traffic index (B6)	0.35 ± 0.09	0.46 ± 0.07	0.54 ± 0.12	134.539	0.000**
Pollution index (B7)	0.52 ± 0.10	0.62 ± 0.09	0.67 ± 0.16	72.514	0.000**
Social index (B8)	0.96 ± 0.05	0.93 ± 0.07	0.72 ± 0.20	211.989	0.000**

*$p < 0.05$ **$p < 0.01$.

suitability areas, the focus should shift towards improving economic conditions and developing infrastructure, while simultaneously supporting ecological restoration and the preservation of cultural identity. This underscores the need for greater attention and policy support to enhance human settlement suitability (HSS) in these regions.

Our study further categorized traditional villages into three distinct clusters: environmentally stressed villages, moderately developed villages, and economically and ecologically advantaged villages. The unique characteristics of each type reflect the diversity and complexity inherent in rural environments. Environmentally stressed villages, predominantly characterized by flat topography, mild climate, and limited ecological features, face significant challenges. These villages are burdened by high environmental pollution pressures, as evidenced by elevated standardized pollution indicators. Their low scores in economic and social indicators indicate underdeveloped areas grappling with not only economic and social issues but also severe environmental degradation. This highlights the need for targeted interventions focusing on environmental rehabilitation and sustainable economic development.

Moderately developed villages exhibit a more balanced profile, characterized by moderate diversity in topography, climate, and ecological features. Compared to environmentally stressed villages, these areas demonstrate a higher level of social and economic activity. Additionally, the slightly lower pollution index indicates comparatively better environmental conditions. Nevertheless, there remains a need for further improvement, particularly in narrowing the gap between these villages and the more advantaged ones. Policies promoting economic growth while safeguarding environmental health would be particularly beneficial for these areas.

Economically and ecologically advantaged villages exhibit the optimal combination of robust economic growth and favorable environmental conditions. These villages score highly in standardized topography, climate, ecology, economy, and transportation indicators, indicating that they are well-positioned in terms of natural resources and have effectively leveraged these advantages for sustainable economic development. The primary challenge for these villages is to maintain this balance while serving as models of sustainable rural development.

Analysis of the influence mechanism on the TVHSS. The analysis of the nonlinear influence mechanisms of external factors on traditional village human settlement suitability (TVHSS) further substantiates the notion that economic factors significantly shape settlement suitability. However, the findings also indicate that while economic prosperity is essential, its positive effects diminish beyond certain thresholds. For instance, income levels (X3) exhibit a plateau effect, suggesting that once a certain

income level is reached, further increases do not substantially enhance settlement suitability. This implies that policies aimed at improving living standards should not solely focus on raising income but should also prioritize enhancing other social and ecological factors that contribute to the overall quality of life. This observation aligns with the findings of Chen and Liu (2016), who identified economic factors as significant drivers of urban residents' behavior and settlement choices. X4 and X6, which reflect economic prosperity and market conditions, respectively, also emerge as crucial elements. Bian and Chen (2024) similarly reported that economic activities influence environmental and social conditions in various Chinese provinces. Although X5 and other factors are relatively less influential, they still play a role by encapsulating cultural and social aspects that shape lifestyle and community identity. Additionally, while X1 and X2 show positive impacts at certain stages, their overall importance remains comparatively low. X7, after surpassing a specific threshold, exhibits a stabilizing effect on TVHSS, indicating that its influence becomes steady over time. Furthermore, the influence of cultural proximity (X5) and social security (X7) underscores the importance of integrating cultural and social dimensions into rural development strategies. These factors not only enhance individual well-being but also shape collective identity and strengthen community resilience in rural areas. Future research should explore how cultural and social factors intersect with economic development to influence long-term sustainability.

Regarding specific factors, X3 exhibits a positive impact on TVHSS, which plateaus at approximately 41,000 CNY. This suggests that beyond this income level, additional income does not substantially enhance HSS, likely due to the saturation of basic needs fulfillment. Sun and Shang (2015) investigated the relationship between income and health status in China, observing significant impacts up to a certain income level, beyond which the effects plateau—a pattern that parallels the findings of this study. Similarly, X4 shows a sharp increase in influence around 25,000 CNY, plateauing at 28,000 CNY, possibly indicating a threshold where increased accessibility and diversity of consumer goods no longer significantly improve lifestyle quality. Wei et al. (2007) also noted that after a specific point, economic growth and consumer spending do not continue to enhance quality of life, mirroring the observed pattern of per capita retail sales of consumer goods in this study. X6 demonstrates variable impacts on TVHSS, suggesting that real estate values are associated with both economic prosperity and potential gentrification effects, which may alter the original community fabric. Lu et al. (2022) examined how housing prices can reflect various tangible and intangible community aspects, emphasizing the role of real estate in shaping economic conditions and social dynamics. X5, however, uniquely exhibits a consistently negative impact on TVHSS, likely due to the
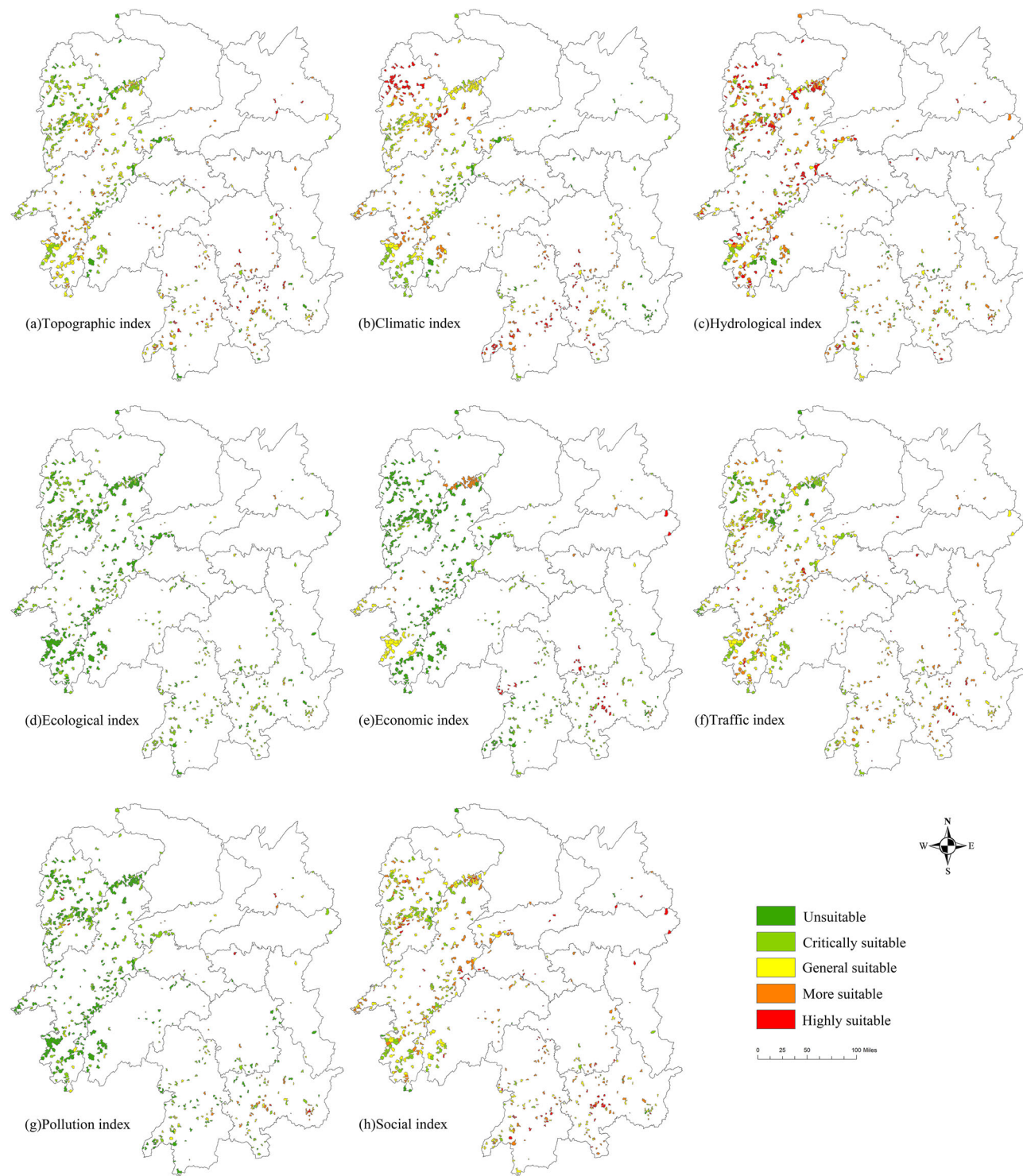
Fig. 7 Spatial distribution of the single-factor suitability evaluation of the TVHS (Source: author's drawing).

Table 7 Driving factor analysis of the Geodetector.

Variable	X1	X2	X3	X4	X5	X6	X7
q statistic	0.134	0.108	0.373	0.395	0.044	0.319	0.164
p-value	<0.001	<0.001	<0.001	<0.001	<0.01	<0.001	<0.001

commercialization or neglect of cultural heritage areas that adversely affect traditional lifestyles. Niemczewska (2021) explored the sociocultural impacts of the adaptive reuse of immovable cultural heritage, highlighting that commercial functions can lead to insufficient sociocultural benefits and limited local community engagement. This finding aligns with the present study, where the negative impact of distance to intangible cultural heritage sites is attributed to commercialization, reducing local access and affecting traditional practices. X1 and X2 demonstrate irregular and phased positive impacts, implying that agricultural productivity influences living standards cyclically, potentially linked to economic cycles and policy changes affecting

rural areas. Barbier and Di Falco (2021) reviewed how land as a productive asset impacts living standards in developing countries, emphasizing the relationship between agricultural practices, economic cycles, and socio-economic conditions—a perspective consistent with the current findings. The positive impact of X7 on TVHSS peaks when it reaches approximately 100, after which the effect stabilizes. This result may indicate that once a certain number of social service facilities are provided, the primary needs of the local population are effectively met, and further increases yield diminishing marginal benefits. This observation aligns with Bock (2016), who explored rural marginalization and the role of social innovation in rural development. The study underscored that beyond a specific threshold, increasing the quantity of social services does not necessarily lead to substantial incremental benefits, emphasizing the need for strategic, context-specific social innovations rather than merely expanding service provision.

These results suggest a complex interplay of economic, cultural, and social factors affecting the TVHSS. Economic factors (X3, X4, X6) play significant roles with distinct thresholds indicating saturation points beyond which their positive impacts plateau. Cultural proximity (X5), social security (X7) and agricultural productivity (X1, X2) also contribute uniquely, reflecting both opportunities and challenges in integrating traditional village settings into contemporary socio-economic frameworks. Comparing these results with existing literature, similar nonlinear and threshold effects have been observed in various contexts, emphasizing the need for tailored, context-sensitive approaches in urban planning and rural development.

Policies and sustainable development. Based on our research findings, we propose several actionable recommendations for policymakers to support the sustainable development of traditional villages:

1. Region-Specific Policies: To address the clear spatial differences in settlement suitability, the government can formulate region-specific policies. For areas with high suitability, the focus should be on sustainability, cultural heritage preservation, and ecological conservation. In areas
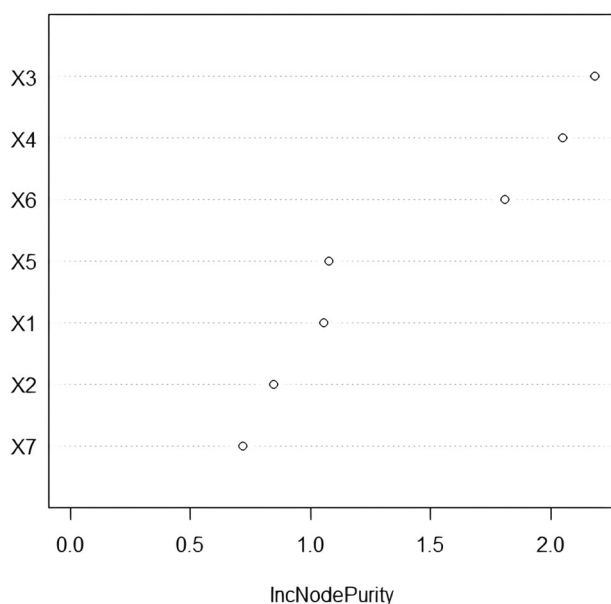

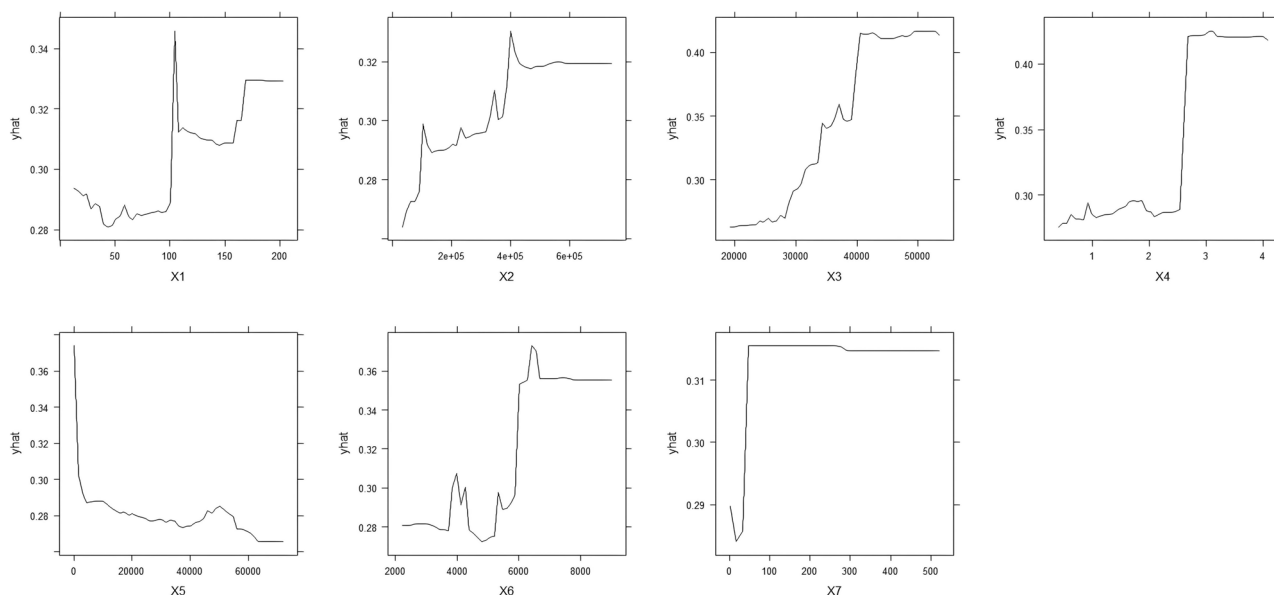**Fig. 8 Driving Factors importance ranking of the Random Forest Model** (Source: author's drawing).

Fig. 9 Partial dependence plot of driving factors (Source: author's drawing). X1, Total Sown Area of Crops; X2, Total Grain Production; X3, Per Capita Disposable Income of Urban Residents; X4, Per Capita Retail Sales of Consumer Goods; X5, Distance to Intangible Cultural Heritage Sites; X6, Housing Price; X7, Number of Social Welfare Adoption Units.

with lower suitability, policies should aim at improving economic development, enhancing infrastructure, and addressing environmental degradation.

2. Targeted Development Strategies: Different villages with varying suitability levels require tailored development strategies. For economically and ecologically advantaged villages, policies should encourage the preservation of natural resources while fostering innovation and technology-driven economic growth. For environmentally stressed villages, environmental rehabilitation and economic diversification should be prioritized, alongside the promotion of community resilience.

3. Cultural and Ecological Preservation: Policies in traditional villages should not only focus on economic development but also on the preservation of cultural practices and traditional ecological knowledge. Encouraging community-driven conservation efforts can enhance both cultural identity and ecological sustainability. Integrating traditional ecological knowledge with modern conservation techniques can provide a more holistic approach to sustainable development.

4. Infrastructure Development: For low-suitability villages, improving land use efficiency, enhancing transportation infrastructure, and providing better access to basic services such as education, healthcare, and clean energy are crucial for promoting rural revitalization.

5. Public Participation and Governance: Future policies should foster public participation, management fairness, and policy inclusiveness, ensuring that local communities have a voice in the development process. Incorporating bottom-up initiatives, as well as supporting social innovation, can empower local communities and ensure that development strategies align with their needs and aspirations.

6. Cross-Regional Cooperation: Economic core areas can provide support to surrounding regions with lower suitability levels, helping to balance development across the province. Interregional cooperation can facilitate resource sharing and knowledge exchange, contributing to more balanced and sustainable regional development.

In conclusion, the findings of this study underscore the need for a context-sensitive approach to rural development. Policy-makers should consider the economic, cultural, and ecological dimensions of traditional village life when formulating development strategies. Tailored, integrated strategies that account for the unique characteristics of each village will be critical in fostering long-term sustainability.

Limitations and future research. While this study provides a comprehensive analysis of the Human Settlement Suitability (HSS) in traditional villages, several key aspects of traditional village life were not fully captured in the current framework. These aspects are vital to understanding the unique challenges and opportunities faced by traditional villages. We acknowledge that these factors, while not directly incorporated into our analysis, represent important dimensions for future research.

1. Geographical Constraints and Generalizability: This study was conducted in Hunan Province, China, and the results are influenced by specific geographical, cultural, and economic characteristics of the region. For instance, the agricultural systems, climatic conditions, and infrastructure in Hunan may differ from those in other countries or regions with distinct socioeconomic structures and environmental factors. These differences could limit the generalizability of the findings to areas with different agricultural practices, climatic zones, or levels of urbanization. Future research should examine the applicability of the HSS framework to comparable regions in other countries with similar rural contexts but differing geographies, economic conditions, and cultural backgrounds. Studies comparing regions with diverse agricultural systems, climatic zones, or urbanization levels will help validate and refine the model's broader applicability. This will allow for a better understanding of how the HSS model can be adapted to different contexts globally. 2. Agrobiodiversity and Agriculture-Related Income: These factors are essential for the long-term sustainability of traditional villages, yet they were not directly represented in our indicators. The complex relationships between agricultural productivity, biodiversity, and economic sustainability warrant further investigation. Future research could integrate these dimensions to better capture the role of agriculture in shaping the suitability of human settlements. Traditional ecological knowledge, as part of agrobiodiversity, should also be considered, as it plays a critical role in sustaining rural livelihoods and maintaining ecological balance. 3. Tourism-Driven Market Demand: The economic impact of tourism on traditional villages is significant, as it often serves as a key driver of local economies. However, tourism's effect on settlement suitability was not explicitly included in our analysis. Future studies could examine how tourism development influences economic vitality, infrastructure, and social cohesion in rural settings. Research could focus on the socio-cultural implications of tourism, including how tourism affects local culture, traditional practices, and community identity. 4. Industrial Diversity and Employment Opportunities: While our study considers economic factors broadly, the specific role of industrial diversity and employment opportunities, especially in non-agricultural sectors, was not fully explored. These factors are crucial for assessing the viability of traditional villages as they transition from agrarian to more diversified economies. Future studies could explore how the expansion of secondary and tertiary industries—such as manufacturing, services, and technology—affects settlement suitability and the economic resilience of these communities. 5. Intergenerational Transmission: The loss of traditional knowledge and cultural practices, as younger generations migrate to urban centers, presents a significant challenge for the sustainability of these villages. Although we discuss social and cultural factors in broad terms, the role of intergenerational transmission in preserving cultural identity and local practices is an important area for future research. This aspect is fundamental to understanding the cultural resilience of traditional villages and their ability to adapt to modern challenges without losing their identity. 6. Population Migration: The shift of young and middle-aged labor forces to urban areas has profound impacts on the socio-economic structure of traditional villages. While migration patterns were not explicitly incorporated into the current study, they significantly affect the vitality and future development of rural settlements. Future research should incorporate migration data to explore its effects on demographic structures, economic sustainability, and community cohesion. This will help better understand the long-term implications of migration on rural settlements, especially in terms of labor availability and community dynamics. 7. Government Policy: In the context of China's rural revitalization efforts, government policies related to economic development, rural land use, and tourism promotion play a critical role in shaping the suitability of human settlements. These policies have not been fully captured in the current study, but their impact on traditional villages is undeniable. Future studies could examine the influence of policy interventions on settlement suitability and how local and national policies interact with the socio-economic dynamics of rural communities. The role of policy inclusiveness, management fairness, and public participation in governance should also be explored, as these

factors significantly shape the overall social context of traditional villages.m 8. Platform Data and Resident Engagement: Another important dimension for future research is the integration of platform data that more explicitly represents the spontaneous activities and grassroots initiatives of residents. Incorporating bottom-up data could provide insights into how resident participation, local initiatives, and community engagement impact the vitality and sustainability of rural communities. Such data could reveal how traditional villages adapt to modern challenges through community-driven efforts.

These factors, though not fully represented in our indicator system, are essential for a more complete understanding of the human settlement suitability of traditional villages. We recommend that future research not only consider these dimensions but also explore the qualitative aspects of traditional village life. Many of these factors—such as cultural heritage, community resilience, and local governance—are difficult to quantify but are fundamental to long-term sustainability.

Finally, the inclusion of longitudinal data would provide deeper insights into the temporal changes in settlement suitability, allowing for a better understanding of the dynamics and trends over time. Future studies should aim to extend the geographic scope and temporal coverage to gain a more comprehensive understanding of the evolving human settlement suitability in traditional villages.

Conclusions

This study, conducted at the village scale, has developed an indicator system suitable for evaluating human settlement suitability (HSS) in traditional villages. Through weighted comprehensive analysis and geographical detector models, the study has systematically examined the HSS index of each traditional village and identified external driving factors. The main conclusions are as follows:

1. The overall quality of HSS in traditional villages is relatively poor, with most villages falling within the Critical Suitability Area and General Suitability Area. There is significant spatial heterogeneity in HSS among these villages. Hejiatian Village exhibits the highest HSS score (1.5228), while Shangbao Village has the lowest (0.1848), with a disparity exceeding eightfold. In highly suitable areas, efforts should concentrate on maintaining existing advantages while promoting sustainable development and preserving cultural and ecological balance. Developing eco-tourism can capitalize on natural and cultural resources to attract visitors, while careful management is required to prevent over-commercialization. Policy incentives and financial support should be utilized to encourage investment in heritage protection.

2. Geographically, HSS shows a decreasing trend from south to northwest, with regional differences evident across traditional villages. Cluster analysis results categorize traditional villages into three distinct types: environmentally stressed villages, moderately developed villages, and economically and ecologically advantaged villages, each with unique characteristics. Ecology, Topography, and Pollution indexes emerge as the most significant contributors to clustering, with respective values of 1.00, 0.66, and 0.63. In contrast, the Hydrology, Society, and Climate indexes have relatively minor impacts, with contribution values of 0.01, 0.24, and 0.27, respectively. These findings provide critical guidance for government resource allocation and policy enhancement. In moderately suitable areas, targeted interventions are needed to address resource limitations and development challenges. Strategic efforts should focus on fostering economic growth, improving infrastructure, and enhancing ecological management through initiatives such as vegetation planting, soil conservation, and pollution mitigation.

3. Analyzing the nonlinear influence of various driving factors enhances the understanding of mechanisms underlying HSS formation. The Geodetector results indicate that the economic environment is the primary driving force behind TVHSS variations, while social welfare and the total sown area of crops exert less influence. The Random Forest model further reveals the ranking of influencing factors as X3 > X4 > X6 > X5 > X1 > X2 > X7. Except for X5, all influencing factors demonstrate significant positive effects and exhibit specific threshold points. Critically suitable and unsuitable areas face the most severe challenges, necessitating urgent measures to prioritize ecological protection while gradually enhancing infrastructure and economic opportunities.

Overall, this study provides valuable insights into the sustainable development of traditional villages in Hunan Province and serves as a case study reference for the international community regarding the preservation and development of traditional villages. The research highlights the inherent complexity of traditional villages, offering substantial support and guidance for scholars and policymakers engaged in related fields. Furthermore, given the global impacts of urbanization and modernization on traditional villages, this study has universal relevance. By adopting strategies tailored to varying HSS levels, policymakers can effectively balance development and preservation, thereby ensuring the long-term sustainability of traditional villages.

Data availability

No datasets were generated or analysed during the current study.

References

Bao BF, Jiang AL, Jin ST, Zhang R (2021) The Evolution and Influencing Factors of Total Factor Productivity of Grain Production Environment: Evidence from Poyang Lake Basin, China. Land 10(6):606

Barbier EB, Di Falco S (2021) Rural Populations, Land Degradation, and Living Standards in Developing Countries. Rev Environ Econ Policy 15(1):115–133

Barboza F, Kimura H, Altman E (2017) Machine learning models and bankruptcy prediction. Expert Syst Appl 83:405–417

Bian JJ, Chen WX, Zeng J (2022) Spatial Distribution Characteristics and Influencing Factors of Traditional Villages in China. Int J Environ Res Public Health 19(8):4627

Bian N, Chen YG (2024) Testing for stationary of housing prices in China: An examination using efficient unit root tests. Heliyon 10(1):e23891

Biau G (2012) Analysis of a Random Forests Model. J Mach Learn Res 13:1063–1095

Bock BB (2016) Rural Marginalisation and the Role of Social Innovation; A Turn Towards Nexogenous Development and Rural Reconnection. Sociologia Ruralis 56(4):552–573

Cao F, Ge Y, Wang JF (2013) Optimal discretization for geographical detectors-based risk assessment. Giscience Remote Sens 50(1):78–92

Chen L, Zhong Q, Li Z (2023) Analysis of spatial characteristics and influence mechanism of human settlement suitability in traditional villages based on multi-scale geographically weighted regression model: A case study of Hunan province. Ecol Indic 154:110828

Chen LL, Shen W (2021) Spatiotemporal differentiation of urban-rural income disparity and its driving force in the Yangtze River Economic Belt during 2000-2017. Plos One 16(2):e0245961

Chen SW, Liu ZL (2016) What determines the settlement intention of rural migrants in China? Economic incentives versus sociocultural conditions. Habitat Int 58:42–50

Chen XH, Xie WZ, Li HB (2020) The spatial evolution process, characteristics and driving factors of traditional villages from the perspective of the cultural ecosystem: A case study of Chengkan Village. Habitat Int 104:102250

Chu KF, Wu MY (2022) The traditional settlement planning and the renovation of residential buildings based on spatial syntax analysis. Soft Comput 26(16):7809–7815

Ding YP, Shi B, Su GJ, Li QQ, Meng J, Jiang YJ, Qin Y, Dai LW, Song S (2021) Assessing Suitability of Human Settlements in High-Altitude Area Using a Comprehensive Index Method: A Case Study of Tibet, China. Sustainability 13(3):1485

Duan YP, Yan LQ, Lai ZL, Chen QT, Sun YY, Zhang L (2022) The spatial form of traditional villages in Fuzhou area of Jiangxi Province determined via GIS methods. Front Earth Sci

Dumitrescu E, Hué S, Hurlin C, Tokpavi S (2022) Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. Eur J Operational Res 297(3):1178–1192

Feng ZM, Tang Y, Yang YZ, Zhang D (2008) Establishment and Application of Human Settlements Environment lndex Model (HEl) Based on GlS. Acta Geographica Sin 63(12):1327–1336

Gao CZ, Hou JP, Gong J (2022) Study on the symbiotic development path of traditional villages in the county based on the grey relation analysis: the case of Xunxian County in Henan, China. Grey Systems-Theory and Application

Gao J, Wu BH (2017) Revitalizing traditional villages through rural tourism: A case study of Yuanjia Village, Shaanxi Province, China. Tour Manag 63:223–233

Guan YY, Li XM, Yang J, Li SB, Tian SZ (2022) Spatial differentiation of comprehensive suitability of urban human settlements based on GIS: a case study of Liaoning Province, China. Environ Dev Sustainability 24(3):4150–4174

Halik W, Mamat A, Dang JH, Deng BS, Tiyip T (2013) Suitability analysis of human settlement environment within the Tarim Basin in Northwestern China. Quat Int 311:175–180

Helbich M, Klein N, Roberts H, Hagedoorn P, Groenewegen PP (2018) More green space is related to less antidepressant prescription rates in the Netherlands: A Bayesian geoadditive quantile regression approach. Environ Res 166:290–297

Hu XL, Li HB, Zhang XL, Chen XH, Yuan Y (2019) Multi-dimensionality and the totality of rural spatial restructuring from the perspective of the rural space system: A case study of traditional villages in the ancient Huizhou region, China. Habitat Int 94:102062

Jia ZY, Gao MY, Xu SJ, Lyu YP, Feng JC, Zhou ZX, Yu TF, Wu WJ (2021) Sociocultural vitality versus regulation policy and tourism development in preservation of traditional rural landscape: a case from Guizhou, China. Int J Sustain Dev World Ecol 28(2):179–192

Jiaxing Z, Lin L, Hang L, Dongmei P (2021) Evaluation and analysis on suitability of human settlement environment in Qingdao. Plos One 16(9):e0256502

Jolliffe IT, Cadima J (2016) Principal component analysis: a review and recent developments. Philos Trans R Soc a-Math Phys Eng Sci 374(2065):20150202

Kang YH, Zhang F, Gao S, Peng WZ, Ratti C (2021) Human settlement value assessment from a place perspective: Considering human dynamics and perceptions in house price modeling. Cities 118:103333

Kim Y, Kim Y (2023) Global regionalization of heat environment quality perception based on K-means clustering and Google trends data. Sustain Cities Soc 96:104710

Li CX, Wu KN (2017) Driving forces of the villages hollowing based on geographically weighted regression model: a case study of Longde County, the Ningxia Hui Autonomous Region, China. Nat Hazards 89(3):1059–1079

Li T, Li CK, Zhang R, Cong Z, Mao Y (2023) Spatial Heterogeneity and Influence Factors of Traditional Villages in the Wuling Mountain Area, Hunan Province, China Based on Multiscale Geographically Weighted Regression. Buildings 13(2):294

Lin YM, Li P, Feng ZM, Yang YZ, You Z, Zhu FX (2021) Climate Suitability Assessment of Human Settlements for Regions along the Belt and Road. Chin Geographical Sci 31(6):996–1010

Liu CL, Wang FH, Xu YP (2019) Habitation environment suitability and population density patterns in China: A regionalization approach. Growth Change 50(1):184–200

Liu RX, Kuang J, Gong Q, Hou XL (2003) Principal component regression analysis with SPSS. Computer Methods Prog Biomedicine 71(2):141–147

Liu S, Ge J, Bai M, Yao M, He LY, Chen M (2022a) Toward classification-based sustainable revitalization: Assessing the vitality of traditional villages. Land Use Policy 116:106060

Liu XQ, Li YW, Wu YF, Li CR (2022b) The Spatial Pedigree in Traditional Villages under the Perspective of Urban Regeneration-Taking 728 Villages in Jiangnan Region, China as Cases. Land 11(9):1561

Lu S, Rao XF, Duan PX (2022) The Rural Gentrification and Its Impacts in Traditional Villages?A Case Study of Xixinan Village, in China. Sustainability 14(16):10077

Luo JQ, Zhou TG, Du PJ, Xu ZG (2019) Spatial-temporal variations of natural suitability of human settlement environment in the Three Gorges Reservoir AreaA case study in Fengjie County, China. Front Earth Sci 13(1):1–17

Luo X, Yang J, Sun W, He BJ (2021) Suitability of human settlements in mountainous areas from the perspective of ventilation: A case study of the main urban area of Chongqing. J Clean Prod 310:127467

Ma HD, Tong YQ (2022) Spatial differentiation of traditional villages using ArcGIS and GeoDa: A case study of Southwest China. Ecol Inform 68:101416

Ma XY, Zhang SD, Guo L, Yang JB (2023) The Influence of Traditional Ethnic Villages on Forest Structure Based on PLS-SEM: A Case Study of Miao Inhabited Area. Forests 14(10):2011

Müller A (2020) External Migrants under Mainland China's Informal Welfare Regime: Risk Shifts, Resource Environments, and the Urban Employees' Social Insurance. Pac Aff 93(2):281–303

Nandy S (2021) Assessment of terrain stability zones for human habitation in Himalayan Upper Pindar River Basin, Uttarakhand using AHP and GIS. Environ Earth Sci 80(9):356

Niemczewska ZE (2021) The sociocultural impact of adaptive reuse of immovable cultural heritage from the perspective of direct users and the local community. J Cultural Herit Manag Sustain Dev 11(3):240–261

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine Learning in Python. J Mach Learn Res 12:2825–2830

Song F, Yang XH, Wu FF (2019) Suitable Pattern of the Natural Environment of Human Settlements in the Lower Reaches of the Yangtze River. Atmosphere 10(4):200

Speiser JL, Miller ME, Tooze J, Ip E (2019) A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst Appl 134:93–101

Su HR, Wang YW, Zhang Z, Dong W (2022) Characteristics and Influencing Factors of Traditional Village Distribution in China. Land 11(10):1631

Sun YH, Shang JX (2015) Factors affecting the health of residents in China: A perspective based on the living environment. Ecol Indic 51:228–236

Tang LS, Ruth M, He QY, Mirzaee S (2017) Comprehensive evaluation of trends in human settlements quality changes and spatial differentiation characteristics of 35 Chinese major cities. Habitat Int 70:81–90

Tang Y, Feng ZM, Yang YZ (2008) Evaluation of Climate Suitability for Human Settlement in China. Res Sci (05), 648–653

Tian SZ, Li XM, Yang J, Wang H, Guo JK (2023) Spatiotemporal evolution of pseudo human settlements: case study of 36 cities in the three provinces of Northeast China from 2011 to 2018. Environ Dev Sustainability 25(2):1742–1772

Tong SY, Zhu YF, Li Z (2022) Correlation Study between Rural Human Settlement Health Factors: A Case Study of Xiangxi, China. Computational Intell Neurosci 2022:2484850

Wang DG, Zhu YJ, Zhao MF, Lv QY (2021) Multi-dimensional hollowing characteristics of traditional villages and its influence mechanism based on the micro-scale: A case study of Dongcun Village in Suzhou, China. Land Use Policy 101:105146

Wang HF, Chiou SC (2019) Study on the Sustainable Development of Human Settlement Space Environment in Traditional Villages. Sustainability 11(15):4186

Wang M, Li XM (2018) Urban human settlement environment suitability assessment: A case study of four districts in Dalian City. J Hum Settl West China 33(04):48–53

Wang Y, Ding Q, Zhuang DF (2015) An eco-city evaluation method based on spatial analysis technology: A case study of Jiangsu Province, China. Ecol Indic 58:37–46

Wang Y, Jin C, Lu MQ, Lu YQ (2017) Assessing the suitability of regional human settlements environment from a different preferences perspective: A case study of Zhejiang Province, China. Habitat Int 70:1–12

Wang ZB, Liang LW, Sun Z, Wang XM (2019) Spatiotemporal differentiation and the factors influencing urbanization and ecological environment synergistic effects within the Beijing-Tianjin-Hebei urban agglomeration. J Environ Manag 243:227–239

Wei HJ, Gao YY, Han Q, Li L, Dong XB, Liu MX, Meng QX (2022) Quality Evaluation and Obstacle Identification of Human Settlements in the Qinghai-Tibet Plateau Based on Multi-Source Data. Land 11(9):1479

Wei W, Shi PJ, Zhou JJ, Feng HC, Wang XF, Wang XP (2013) Environmental suitability evaluation for human settlements in an arid inland river basin: A case study of the Shiyang River Basin. J Geographical Sci 23(2):331–343

Wei YM, Liu LC, Fan Y, Wu G (2007) The impact of lifestyle on energy use and CO_2 emission:: An empirical analysis of China's residents. Energy Policy 35(1):247–257

Wu C, Chen MM, Zhou L, Liang XJ, Wang W (2020) Identifying the Spatiotemporal Patterns of Traditional Villages in China: A Multiscale Perspective. Land 9(11):449

Xia JS, Dong PL (2019) Spatial characteristics of physical environments for human settlements in Jinsha River watershed (Yunnan section), China. Geomat Nat Hazards Risk 10(1):544–561

Xiao CW, Feng ZM, Li P, You Z, Teng JK (2018) Evaluating the suitability of different terrains for sustaining human settlements according to the local elevation range in China using the ASTER GDEM. J Mt Sci 15(12):2741–2751

Xie TC, Liu XY, Nie PJ (2022) Study on Spatial-Temporal Patterns and Factors Influencing Human Settlement Quality in Beijing. Sustainability 14(7):3752

Xin FF, Xiao XM, Dong JW, Zhang GL, Zhang Y, Wu XC, Li XG, Zou ZH, Ma J, Du GM, Doughty RB, Zhao B, Li B (2020) Large increases of paddy rice area, gross primary production, and grain production in Northeast China during 2000-2017. Sci Total Environ 711:135183

Xiong Y (2011) Uncertainty evaluation of the coordinated development of urban human settlement environment and economy in Changsha city. J Geographical Sci 21(6):1123–1137

Xiong Y, Zhang JP, Yan Y, Sun SB, Xu XY, Higueras E (2022) Effect of the spatial form of Jiangnan traditional villages on microclimate and human comfort. Sustain Cities Soc 87:104136

Xu J, Yang MS, Hou CP, Lu ZL, Liu D (2021) Distribution of rural tourism development in geographical space: a case study of 323 traditional villages in Shaanxi, China. Eur J Remote Sens 54:318–333

Xu Q, Wang J (2021) Recognition of Values of Traditional Villages in Southwest China for Sustainable Development: A Case Study of Liufang Village. Sustainability 13(14):1–21

Yang R, Pan YX, Xu Q (2020) Space diversification process and evolution mechanism of typical village in the suburbs of Guangzhou: A case study of Beicun. J Geographical Sci 30(7):1155–1178

You Z, Feng ZM, Yang YZ, Shi H, Li P (2020) Evaluation of human settlement environmental suitability in Tibet based on gridded data. Resour Sci 42(02):394–406

Yu L, Xie DL, Xu X (2022) Environmental Suitability Evaluation for Human Settlements of Rural Residential Areas in Hengshui, Hebei Province. Land 11(12):2112

Zhang L, Lu P, Lau R, Yan LJ, Li X, Yang RX, Leung HH, Chen PP, Wang X (2021) Unique traditional villages on the Loess Plateau of China: historic evolution and challenges to sustainable development of silo-caves. Herit Sci 9(1):118

Zhang ZH, Song YZ, Wu P (2022) Robust geographical detector. Int J Appl Earth Observation Geoinf 109:102782

Zhao J, Xu M, Lu SL, Cao CX (2013) Human settlement evaluation in mountain areas based on remote sensing, GIS and ecological niche modeling. J Mt Sci 10(3):378–387

Zheng XY, Wu JH, Deng HB (2021) Spatial Distribution and Land Use of Traditional Villages in Southwest China. Sustainability 13(11):6326

Zhong Q, Fu H, Yan J, Li Z (2024) How does energy utilization affect rural sustainability development in traditional villages? Re-examination from the coupling coordination degree of atmosphere-ecology-socioeconomics system. Build Environ 257:111541

Zhong Q, Li Z, Zhu J, Yuan C (2025) Revealing Multiscale and Nonlinear Effects of Urban Green Spaces on Heat Islands in High-Density Cities: Insights from MSPA and Machine Learning. Sustain Cities Soc 120:106173

Zhou ZX, Jia ZY, Wang N, Fang M (2018) Sustainable Mountain Village Construction Adapted to Livelihood, Topography, and Hydrology: A Case of Dong Villages in Southeast Guizhou, China. Sustainability 10(12):4619

Acknowledgements

Author contributions

Conceptualization: QZ and LX; methodology: QZ and LX; formal analysis: QZ; resources: QZ; writing—original draft preparation: QZ, JW, and LX; writing—review and editing: QZ, JW, and LX; visualisation: QZ and LX; supervision: QZ; funding acquisition: QZ; All authors have read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Ethics approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-025-04971-0.

Correspondence and requests for materials should be addressed to Liang Xie.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.